

# Notes and exercises from *Linear Algebra and Multilinear Algebra*

John Peloquin

## Introduction

This document contains notes and exercises from [1] and [2].

**Unless otherwise stated,  $\Gamma$  denotes a field of characteristic 0 over which all vector spaces are defined.**

# Linear Algebra

## Chapter I

### § 1

*Remark.* The free vector space  $C(X)$  is intuitively the space of all “formal linear combinations” of  $x \in X$ .

### § 2

**Exercise** (5 - Universal property of  $C(X)$ ). Let  $X$  be a set and  $C(X)$  the free vector space on  $X$  (subsection 1.7). Recall

$$C(X) = \{f : X \rightarrow \Gamma \mid f(x) = 0 \text{ for all but finitely many } x \in X\}$$

The inclusion map  $i_X : X \rightarrow C(X)$  is defined by  $a \mapsto f_a$  where  $f_a$  is the “characteristic function” of  $a$ :  $f_a(a) = 1$  and  $f_a(x) = 0$  for all  $x \neq a$ . For  $f \in C(X)$ ,  $f = \sum_{a \in X} f(a)f_a$ .

- (i) If  $F$  is a vector space and  $f : X \rightarrow F$ , there is a unique *linear*  $\varphi : C(X) \rightarrow F$  “extending  $f$ ” in the sense that  $\varphi \circ i_X = f$ :

$$\begin{array}{ccc} X & \xrightarrow{i_X} & C(X) \\ & \searrow f & \downarrow \varphi \\ & & F \end{array}$$

- (ii) If  $\alpha : X \rightarrow Y$ , there is a unique *linear*  $\alpha_* : C(X) \rightarrow C(Y)$  which makes the following diagram commute:

$$\begin{array}{ccc} X & \xrightarrow{\alpha} & Y \\ i_X \downarrow & & \downarrow i_Y \\ C(X) & \xrightarrow{\alpha_*} & C(Y) \end{array}$$

If  $\beta : Y \rightarrow Z$ , then  $(\beta \circ \alpha)_* = \beta_* \circ \alpha_*$ .

- (iii) If  $E$  is a vector space, there is a unique linear map  $\pi_E : C(E) \rightarrow E$  such that  $\pi_E \circ i_E = \iota_E$  (where  $\iota_E : E \rightarrow E$  is the identity map):

$$\begin{array}{ccc} E & \xrightarrow{i_E} & C(E) \\ & \searrow \iota_E & \vdots \pi_E \\ & & E \end{array}$$

- (iv) If  $E$  and  $F$  are vector spaces and  $\varphi : E \rightarrow F$ , then  $\varphi$  is linear if and only if  $\pi_F \circ \varphi_* = \varphi \circ \pi_E$ :

$$\begin{array}{ccc} E & \xrightarrow{\varphi} & F \\ \pi_E \uparrow & & \uparrow \pi_F \\ C(E) & \xrightarrow{\varphi_*} & C(F) \end{array}$$

- (v) Let  $E$  be a vector space and  $N(E)$  the subspace of  $C(E)$  generated by all elements of the form

$$f_{\lambda a + \mu b} - \lambda f_a - \mu f_b \quad (a, b \in E \text{ and } \lambda, \mu \in \Gamma)$$

Then  $\ker \pi_E = N(E)$ .

*Proof.*

- (i) By Proposition II, since  $i_X(X)$  is a basis of  $C(X)$ .  
(ii) By (i), applied to  $i_Y \circ \alpha$ . Note  $\beta_* \circ \alpha_*$  is linear such that

$$(\beta_* \circ \alpha_*) \circ i_X = i_Z \circ (\beta \circ \alpha)$$

so  $\beta_* \circ \alpha_* = (\beta \circ \alpha)_*$  by uniqueness:

$$\begin{array}{ccccc} X & \xrightarrow{\alpha} & Y & \xrightarrow{\beta} & Z \\ i_X \downarrow & & i_Y \downarrow & & i_Z \downarrow \\ C(X) & \xrightarrow{\alpha_*} & C(Y) & \xrightarrow{\beta_*} & C(Z) \end{array}$$

(iii) By (i), applied to  $\iota_E$ .

(iv) If  $\varphi$  is linear, then  $\varphi \circ \pi_E : C(E) \rightarrow F$  is linear and extends  $\varphi$  in the sense that  $\varphi \circ \pi_E \circ i_E = \varphi \circ \iota_E = \varphi$ . However,  $\pi_F \circ \varphi_* : C(E) \rightarrow F$  is also linear and extends  $\varphi$  since

$$\pi_F \circ \varphi_* \circ i_E = \pi_F \circ i_F \circ \varphi = \iota_F \circ \varphi = \varphi$$

By uniqueness, these two maps must be equal. Conversely, if these two maps are equal, then  $\varphi$  is linear since  $\pi_F \circ \varphi_*$  is linear and  $\pi_E$  is surjective.

(v) By (iii),

$$\begin{aligned} \pi_E(f_{\lambda a + \mu b} - \lambda f_a - \mu f_b) &= \pi_E(f_{\lambda a + \mu b}) - \lambda \pi_E(f_a) - \mu \pi_E(f_b) \\ &= \lambda a + \mu b - \lambda a - \mu b \\ &= 0 \end{aligned}$$

for all  $a, b \in E$  and  $\lambda, \mu \in \Gamma$ . It follows that  $N(E) \subseteq \ker \pi_E$  since  $N(E)$  is the *smallest* subspace containing these elements and  $\ker \pi_E$  is a subspace.

On the other hand, it follows from the fact that  $N(E)$  is a subspace that

$$\sum \lambda_i f_{a_i} - f_{\sum \lambda_i a_i} \in N(E)$$

for all (finite) linear combinations. Now if  $g = \sum_{a \in E} g(a) f_a \in \ker \pi_E$ , then

$$0 = \pi_E(g) = \sum_{a \in E} g(a) \pi_E(f_a) = \sum_{a \in E} g(a) a$$

This implies  $f_{\sum_{a \in E} g(a) a} = f_0 \in N(E)$ . But by the above,  $g - f_0 \in N(E)$ , so  $g \in N(E)$ . Therefore also  $\ker \pi_E \subseteq N(E)$ .  $\square$

*Remark.* Note (i) shows that  $C(X)$  is a universal (initial) object in the category of “vector spaces with maps of  $X$  into them”. In this category, the objects are maps  $X \rightarrow F$ , for vector spaces  $F$ , and the arrows are *linear* (i.e. structure-preserving) maps  $F \rightarrow G$  between the vector spaces which respect the mappings of  $X$ :

$$\begin{array}{ccc} X & \xrightarrow{\quad} & F \\ & \searrow & \downarrow \\ & & G \end{array}$$

By (i), every object  $X \rightarrow F$  in this category can be obtained from the inclusion map  $X \rightarrow C(X)$  in a unique way. This is why  $C(X)$  is called “universal”. This is only possible because  $C(X)$  is free from any nontrivial relations among the elements of  $X$ , so any relations among the images of those elements in  $F$  can be obtained starting from  $C(X)$ . This is why  $C(X)$  is called “free”. It is immediate from the universal property that  $C(X)$  is unique up to isomorphism: if  $X \rightarrow U$  is also universal, then the composites  $\psi \circ \varphi$  and  $\varphi \circ \psi$  of the induced linear maps  $\varphi : C(X) \rightarrow U$  and  $\psi : U \rightarrow C(X)$  are linear and extend the inclusion maps, so must be the identity maps on  $C(X)$  and  $U$  by uniqueness; that is,  $\varphi$  and  $\psi$  are mutually inverse and hence *isomorphisms*. In fact they are also unique by the universal property.

Now (ii) shows that we have a *functor* from the category of sets into the category of vector spaces, which sends sets  $X$  and  $Y$  to the vector spaces  $C(X)$  and  $C(Y)$ , and which sends a set map  $\alpha : X \rightarrow Y$  to the linear map  $\alpha_* : C(X) \rightarrow C(Y)$ . The functor preserves the category structure of composites of arrows.

In (iii), we are “forgetting” the linear structure of  $E$  when forming  $C(E)$ . For example, if  $E = \mathbb{R}^2$ , then  $(1, 1) = (1, 0) + (0, 1)$  in  $E$ , but *not* in  $C(E)$ . The “formal” linear combination

$$(1, 1) - (1, 0) - (0, 1)$$

is not zero in  $C(E)$  because the pairs are unrelated elements (symbols) which are *linearly independent*. Note  $\pi_E$  is surjective (since  $\iota_E$  is), so  $E$  is a projection of  $C(E)$ . In (iv), we see that  $\varphi : E \rightarrow F$  is linear if and only if it is a “projection” of  $\varphi_* : C(E) \rightarrow C(F)$ .

In (v), we see that  $\pi_E$  just recalls the linear structure of  $E$  that was forgotten in  $C(E)$ . In particular,  $C(E)/N(E) \cong E$ . In other words, if you start with  $E$ , then forget about its linear structure, then recall that linear structure, you just get  $E$  again.

### § 3

*Remark.* If

$$E_1 \oplus E_2 = E = F_1 \oplus F_2$$

are two direct sum decompositions of  $E$  with  $E_i \subseteq F_i$  for  $i = 1, 2$ , then  $E_i = F_i$  for  $i = 1, 2$ .

*Proof.* If  $x \in F_i$ , then  $x = x_1 + x_2$  with  $x_j \in E_j$ . But then we must have  $x = x_i$  since  $x_j \in F_j$ , so  $x \in E_i$ . □

## § 4

**Exercise (11).** Let  $E$  be a real vector space and  $E_1$  a vector hyperplane in  $E$  (that is, a subspace of codimension 1). Define an equivalence relation on  $E^1 = E - E_1$  as follows: for  $x, y \in E^1$ ,  $x \sim y$  if the segment

$$x(t) = (1 - t)x + ty \quad (0 \leq t \leq 1)$$

is disjoint from  $E_1$ . Then there are precisely two equivalence classes.

*Proof.* Fix  $e \in E^1$  with  $E = E_1 \oplus \langle e \rangle$  and define  $\alpha : E \rightarrow \mathbb{R}$  by  $x - \alpha(x)e \in E_1$  for all  $x \in E$ . It is clear that  $\alpha$  is linear, and  $x \in E_1$  if and only if  $\alpha(x) = 0$ . For  $x, y \in E^1$ , it follows that  $x \sim y$  if and only if

$$0 \neq \alpha(x(t)) = \alpha((1 - t)x + ty) = (1 - t)\alpha(x) + t\alpha(y)$$

for all  $0 \leq t \leq 1$ . But this is just equivalent to  $\alpha(x)\alpha(y) > 0$ .

Now if  $x \in E^1$ , then  $\alpha(x) \neq 0$ , so  $\alpha(x)^2 > 0$  and  $x \sim x$ . If  $x \sim y$ , then  $\alpha(y)\alpha(x) = \alpha(x)\alpha(y) > 0$ , so  $y \sim x$ . If also  $y \sim z$ , then  $\alpha(y)\alpha(z) > 0$ , so  $\alpha(x)\alpha(z) > 0$  and  $x \sim z$ . In other words, this is indeed an equivalence relation.

Note there are at least two equivalence classes since  $\alpha(e) = 1$  and  $\alpha(-e) = -1$ , so  $\alpha(e)\alpha(-e) = -1 < 0$  and  $e \not\sim -e$ . On the other hand, there are at most two classes since if  $x \in E^1$ , then either  $\alpha(x) > 0$  and  $x \sim e$  or  $\alpha(x) < 0$  and  $x \sim -e$ .  $\square$

*Remark.* This result shows that the hyperplane separates the vector space into two disjoint half-spaces.

## Chapter II

### § 2

*Remark.* In subsection 2.11, in the second part of the proof of Proposition I, just let  $\psi : E \leftarrow F$  be any linear mapping extending  $\varphi_1^{-1} : E \leftarrow \text{Im } \varphi$ .<sup>1</sup>

### § 4

*Remark.* The direct sum  $E \oplus F$  is a coproduct in the category of vector spaces in the following sense: if  $\varphi : E \rightarrow G$  and  $\psi : F \rightarrow G$  are linear maps, there is a

---

<sup>1</sup> See Corollary I to Proposition I in subsection 1.15.

unique linear map  $\chi : E \oplus F \rightarrow G$  such that  $\varphi = \chi \circ i_E$  and  $\psi = \chi \circ i_F$ , where  $i_E$  and  $i_F$  are the canonical injections:

$$\begin{array}{ccccc}
 E & \xrightarrow{i_E} & E \oplus F & \xleftarrow{i_F} & F \\
 & \searrow \varphi & \downarrow \chi & \swarrow \psi & \\
 & & G & & 
 \end{array}$$

Indeed,  $\chi$  is given by  $\chi(x + y) = \varphi(x) + \psi(y)$  for  $x \in E$ ,  $y \in F$ . It is the unique linear map “extending” both  $\varphi$  and  $\psi$ . This property makes  $E \oplus F$  unique up to a unique isomorphism.

Dually,  $E \oplus F$  is a product in the following sense: if  $\varphi : G \rightarrow E$  and  $\psi : G \rightarrow F$  are linear maps, there is a unique linear map  $\chi : G \rightarrow E \oplus F$  such that  $\varphi = \pi_E \circ \chi$  and  $\psi = \pi_F \circ \chi$ :

$$\begin{array}{ccccc}
 & & G & & \\
 & \searrow \varphi & \downarrow \chi & \swarrow \psi & \\
 E & \xleftarrow{\pi_E} & E \oplus F & \xrightarrow{\pi_F} & F
 \end{array}$$

Indeed,  $\chi$  is given by  $\chi(x) = \varphi(x) + \psi(x)$ , and “combines”  $\varphi$  and  $\psi$ . This property also makes  $E \oplus F$  unique up to a unique isomorphism. An infinite direct sum is also a coproduct, but *not* a product, essentially because it has no infinite sums of elements.

In the proof of Proposition I,  $\sigma$  is the product map and  $\tau$  is the coproduct map. If  $\varphi_1 : E_1 \rightarrow F_1$  and  $\varphi_2 : E_2 \rightarrow F_2$  are linear maps, then  $\varphi = \varphi_1 \oplus \varphi_2$  is both a coproduct and product map:

$$\begin{array}{ccccc}
 E_1 & \xrightleftharpoons{\quad} & E_1 \oplus E_2 & \xrightleftharpoons{\quad} & E_2 \\
 \downarrow \varphi_1 & & \downarrow \varphi & & \downarrow \varphi_2 \\
 F_1 & \xrightleftharpoons{\quad} & F_1 \oplus F_2 & \xrightleftharpoons{\quad} & F_2
 \end{array}$$

The structure of  $\varphi$  is completely determined by the structures of  $\varphi_1$  and  $\varphi_2$ . In particular,  $\varphi$  is injective (surjective, bijective) if and only if  $\varphi_1$  and  $\varphi_2$  are.

*Remark.* If  $E = E_1 \oplus E_2$  and  $F = F_1 \oplus F_2$ , and  $\varphi : E \rightarrow F$  is defined by  $\varphi = \varphi_1 \oplus \varphi_2$  where  $\varphi_1 : E_1 \rightarrow F_1$  and  $\varphi_2 : E_2 \rightarrow F_2$ , then the following diagrams commute:

$$\begin{array}{ccc} E_1 & \xrightarrow{\varphi_1} & F_1 \\ \cong \downarrow & & \downarrow \cong \\ E/E_2 & \xrightarrow{\overline{\varphi}} & F/F_2 \end{array} \quad \begin{array}{ccc} E_2 & \xrightarrow{\varphi_2} & F_2 \\ \cong \downarrow & & \downarrow \cong \\ E/E_1 & \xrightarrow{\overline{\varphi}} & F/F_1 \end{array}$$

Here the isomorphisms are those induced by the canonical projections.

## § 5

*Remark.* The definition of dual space is fundamentally *symmetrical* between  $E$  and  $E^*$ , as is the definition of dual mapping between  $\varphi$  and  $\varphi^*$ . This symmetry often allows us to use bidirectional reasoning and derive two theorems from one proof. For example, (2.48) actually follows from (2.47) by symmetry of  $\varphi$  and  $\varphi^*$ . The proof of Proposition I in subsection 2.23 exploits symmetry, as do other proofs in the book. Many other books simply *define* the dual space of  $E$  to be  $L(E)$  (no doubt in light of Proposition I of this subsection), at the expense of this symmetry.

*Remark.* The definition of duality can be generalized to modules. Let  $R$  be a ring (with unit),  $M$  a *left*  $R$ -module, and  $N$  a *right*  $R$ -module. A binary mapping  $\Phi : M \times N \rightarrow R$  is *bilinear* if

$$\Phi(\lambda x_1 + \mu x_2, y) = \lambda \Phi(x_1, y) + \mu \Phi(x_2, y)$$

and

$$\Phi(x, y_1 \lambda + y_2 \mu) = \Phi(x, y_1) \lambda + \Phi(x, y_2) \mu$$

for all  $x, x_1, x_2 \in M$ ,  $y, y_1, y_2 \in N$  and  $\lambda, \mu \in R$ . The map is *nondegenerate* if

$$\{x \in M \mid \Phi(x, y) = 0 \text{ for all } y \in N\} = 0$$

and

$$\{y \in N \mid \Phi(x, y) = 0 \text{ for all } x \in M\} = 0$$

Now  $M$  is *dual* to  $N$  if there is a nondegenerate bilinear map  $\langle -, - \rangle : M \times N \rightarrow R$ . Note this relation is not in general symmetric unless  $R$  is commutative.



*Remark.* The results in subsection 2.23 show that quotient spaces are dual to subspaces.

*Remark.* If  $E, E^*$  and  $F, F^*$  are pairs of dual spaces and  $\varphi : E \rightarrow F$  is linear, then  $\varphi^* : E^* \leftarrow F^*$  is dual to  $\varphi$  if and only if the following diagram commutes:

$$\begin{array}{ccc} F^* \times E & \xrightarrow{\varphi^* \times \iota_E} & E^* \times E \\ \downarrow \iota_{F^*} \times \varphi & & \downarrow \langle , \rangle \\ F^* \times F & \xrightarrow{\langle , \rangle} & \Gamma \end{array}$$

*Remark.* Let  $E$  be a vector space and  $(x_\alpha)_{\alpha \in A}$  be a basis of  $E$ . For each  $x \in E$ , write  $x = \sum_{\alpha \in A} f_\alpha(x) x_\alpha$ . Then  $f_\alpha \in L(E)$  for each  $\alpha \in A$ . The function  $f_\alpha$  is called the  $\alpha$ -th coordinate function for the basis.

Coordinate functions can be used in an alternative proof of Proposition IV. If  $E_1$  is a subspace of  $E$ , let  $B_1$  be a basis of  $E_1$  and extend it to a basis  $B$  of  $E$ . For each  $x_\alpha \in B - B_1$ , we have  $f_\alpha \in E_1^\perp$ . If  $x \in E_1^{\perp\perp}$ , then  $f_\alpha(x) = \langle f_\alpha, x \rangle = 0$  for all such  $\alpha$ , so  $x \in E_1$ . In other words,  $E_1^{\perp\perp} \subseteq E_1$ .

*Remark.* In the corollary to Proposition V, for  $f \in L(E)$  let  $f_k = f \circ i_k \circ \pi_k$  where  $i_k : E_k \rightarrow E$  is the  $k$ -th canonical injection and  $\pi_k : E \rightarrow E_k$  is the  $k$ -th canonical projection. Then  $f = \sum_k f_k$  and  $f_k \in F_k^\perp$  for all  $k$ , so  $L(E) = \sum_k F_k^\perp$ . The sum is direct since if  $f \in F_k^\perp \cap \sum_{j \neq k} F_j^\perp$ , then  $f$  kills  $\sum_{j \neq k} E_j$  and  $E_k$ , so  $f = 0$ . A scalar product is induced between  $E_k, F_k^\perp$  since  $E_k \cap F_k = 0$  and  $F_k^\perp \cap E_k^\perp = 0$ .<sup>2</sup> The induced injection  $F_k^\perp \rightarrow L(E_k)$  is surjective since every linear function on  $E_k$  can be extended to a linear function on  $E$  which kills  $F_k$ .

*Remark.* For  $\varphi : E \rightarrow F$  a linear map, let  $L(\varphi) : L(E) \leftarrow L(F)$  be the dual map given by  $L(\varphi)(f) = f \circ \varphi$  (2.50). Then  $L$  linearly embeds  $L(E; F)$  in  $L(L(F); L(E))$ , by (2.43) and (2.44). Also,  $L(\psi \circ \varphi) = L(\varphi) \circ L(\psi)$  and  $L(\iota_E) = \iota_{L(E)}$ . This shows that  $L$  is a contravariant functor in the category of vector spaces. This functor preserves exactness of sequences (see 2.29), and finite direct sums, which are just (co)products in the category (see 2.30), among other things.

**Exercise (10).** If  $\varphi : E \rightarrow F$  is a linear map with restriction  $\varphi_1 : E_1 \rightarrow F_1$  and dual map  $\varphi^* : E^* \leftarrow F^*$ , then  $\varphi^*$  can be restricted to  $F_1^\perp, E_1^\perp$  and the induced map  $\overline{\varphi^*} : E^*/E_1^\perp \leftarrow F^*/F_1^\perp$  is dual to  $\varphi_1$ .

<sup>2</sup>See subsection 2.23.

*Proof.* If  $y^* \in F_1^\perp$  and  $x \in E_1$ , then

$$\langle \varphi^* y^*, x \rangle = \langle y^*, \varphi x \rangle = 0$$

so  $\varphi^*$  maps  $F_1^\perp$  into  $E_1^\perp$ . We know that the pairs  $E_1, E^*/E_1^\perp$  and  $F_1, F^*/F_1^\perp$  are dual under the induced scalar products. For  $\overline{y^*} \in F^*/F_1^\perp$  and  $x \in E_1$ ,

$$\begin{aligned} \langle \overline{\varphi^* y^*}, x \rangle &= \langle \overline{\varphi^* y^*}, x \rangle \\ &= \langle \varphi^* y^*, x \rangle \\ &= \langle y^*, \varphi_1 x \rangle \\ &= \langle \overline{y^*}, \varphi_1 x \rangle \end{aligned} \quad \square$$

*Remark.* This result shows that quotient maps are dual to restriction maps. The examples in subsections 2.24 and 2.27 are special cases.

## § 6

*Remark.* If  $E$  is finite-dimensional, then since  $E$  is dual to  $L(E)$  and  $L(E)$  is dual to  $L(L(E))$ , it follows from Corollary II that there is a natural isomorphism  $E \rightarrow L(L(E))$  given by

$$x \mapsto (f \mapsto f(x))$$

This means that the vector  $x$  may be naturally *identified* with the linear function “evaluation at  $x$ ”.

*Remark.* If  $E$  is finite-dimensional, then every basis of a dual space  $E^*$  is a dual basis. Indeed, if  $f_1, \dots, f_n$  is a basis of  $E^*$ , let  $e_1, \dots, e_n$  be its dual basis in  $E$ . Then  $\langle f_i, e_j \rangle = \delta_{ij}$  by (2.62), so  $f_1, \dots, f_n$  is the dual basis of  $e_1, \dots, e_n$ , again by (2.62).

Alternatively, with  $E^* = L(E)$ , let  $f_1^*, \dots, f_n^*$  be the dual basis of  $f_1, \dots, f_n$  in  $E^{**} = L(L(E))$ , so  $\langle f_j^*, f_i \rangle = \delta_{ij}$  by (2.62). Let  $e_1, \dots, e_n \in E$  be defined by  $\langle f_j^*, f \rangle = \langle f, e_j \rangle$  for all  $f \in E^*$ .<sup>3</sup> Then  $\langle f_i, e_j \rangle = \langle f_j^*, f_i \rangle = \delta_{ij}$ , so  $f_1, \dots, f_n$  is the dual basis of  $e_1, \dots, e_n$ .

The first proof here uses the symmetry between  $E$  and  $E^*$ , while the second uses the natural isomorphism  $E \cong E^{**}$ .

*Remark.* In Proposition III,

$$\Phi(x^*, x) = \langle x^*, \varphi x \rangle = \langle \varphi^* x^*, x \rangle$$

---

<sup>3</sup>See problem 3 in § 5 or the prior remark.

so there is also a natural isomorphism  $B(E^*, E) \cong L(E^*; E^*)$ . We see that  $\varphi^*$  is “left-dual” to  $\Phi$  while  $\varphi$  is “right-dual” to  $\Phi$ . It might be said that  $\Phi, \varphi, \varphi^*$  enter into a holy trinity.<sup>4</sup>

**Exercise (9).** If  $E$  and  $F$  are finite-dimensional, then the mapping

$$\Phi : L(E; F) \rightarrow L(F^*; E^*)$$

defined by  $\varphi \mapsto \varphi^*$  is a linear isomorphism.

*Proof.* By the remark in § 5 above, and the fact that  $\varphi^{**} = \varphi$ . □

## Chapter III

*Warning.* Greub’s notational choices in this chapter are insane. In particular, although he uses left-hand mapping notation (writing  $\varphi x$  instead of  $x\varphi$ , and  $\varphi\psi$  to mean  $\varphi$  after  $\psi$ ), and follows the usual “row-by-column” convention for matrix multiplication, his convention for the matrix of a linear mapping is the transpose of that normally used in these circumstances. This has the following unfortunate consequences:

- The matrix of the linear mapping naturally associated with a system of linear equations has the coefficients from each equation appear *vertically in columns*.
- If  $M(x)$  is the *column vector* representing  $x$ , then  $M(\varphi x) = M(\varphi)^* M(x)$ , and if  $M(x)$  is the *row vector* representing  $x$ , then  $M(\varphi x) = M(x)M(\varphi)$ .
- $M(\varphi\psi) = M(\psi)M(\varphi)$

Compounding the insanity, Greub (inspired by tensor notation) indexes over columns instead of rows when working in dual spaces. This further increases the risk of confusion and error, as we see below. Greub says that “it would be very undesirable...to agree once and for all to always let the subscript count the rows”, but we couldn’t disagree more in this context.

---

<sup>4</sup>No one says this.

### § 3

*Remark.* In subsection 3.13, although  $\beta_v^\theta = \check{\alpha}_v^\theta$ , we must remember that  $v$  indexes the *columns* of the matrix of the dual basis transformation  $x^{*v} \mapsto \bar{x}^{*v}$  by (3.24), whereas  $v$  indexes the *rows* of the matrix of the basis transformation  $x_v \mapsto \bar{x}_v$  by (3.22). In other words, the matrix of the dual basis transformation  $x^{*v} \mapsto \bar{x}^{*v}$  is the *transpose* of the inverse of the matrix of the basis transformation  $x_v \mapsto \bar{x}_v$ , contrary to what the book says.<sup>5</sup> It's easier to remember that the matrix of  $x^{*v} \mapsto \bar{x}^{*v}$  (arrow reversed!) is the transpose of the matrix of  $x_v \mapsto \bar{x}_v$ .

*Remark.* In subsection 3.13, we see that if a basis transformation is effected by  $\tau$ , then the corresponding *coordinate* transformation is effected by  $\tau^{-1}$ . The coordinates of a vector are transformed “exactly in the same way” as the vectors of the dual basis, despite the previous remark, because of Greub's notational choices.

## Chapter IV

*Remark.* In this chapter, it is implicitly assumed that all vector spaces have dimension  $n \geq 1$ , except in the definition of intersection number (subsection 4.31) where  $n = 0$ . Here we summarize results for the case  $n = 0$ :

- For a set  $X$ ,  $X^0 = \{\emptyset\}$ . Therefore maps  $\Phi : X^0 \rightarrow Y$  can be identified with elements of  $Y$ .
- For vector spaces  $E$  and  $F$ , a map  $\Phi : E^0 \rightarrow F$  is vacuously 0-linear. Since the only permutation in  $S_0$  is the identity  $\iota = \emptyset$ ,  $\Phi$  is also trivially skew symmetric.

In particular if  $E = 0$ , the following results hold:

- Determinant functions in  $E$  are just scalars in  $\Gamma$ , and dual determinant functions are just reciprocal scalars.
- The only transformation of  $E$  is the zero transformation, which is also the identity transformation. It has determinant 1, trace 0, and constant characteristic polynomial 1. It has no eigenvalues or eigenvectors. Its adjoint is also the zero transformation. Its matrix on the empty basis is empty.

---

<sup>5</sup>Compare this to the proof of equation (3.4) in subsection 3.3.

- If  $E$  is real ( $\Gamma = \mathbb{R}$ ), the orientations in  $E$  are represented by the scalars  $\pm 1$ , and determine whether the empty basis is positive or negative. The zero transformation is orientation preserving. The empty basis is deformable into itself.

## § 1

*Remark.* To see why (4.1) holds, observe by definition of  $\tau(\sigma\Phi)$  that

$$(\tau(\sigma\Phi))(x_1, \dots, x_p) = (\sigma\Phi)(x_{\tau(1)}, \dots, x_{\tau(p)})$$

Let  $y_i = x_{\tau(i)}$ . Then by definition of  $\sigma\Phi$  and  $(\tau\sigma)\Phi$ ,

$$\begin{aligned} (\sigma\Phi)(x_{\tau(1)}, \dots, x_{\tau(p)}) &= (\sigma\Phi)(y_1, \dots, y_p) \\ &= \Phi(y_{\sigma(1)}, \dots, y_{\sigma(p)}) \\ &= \Phi(x_{\tau(\sigma(1))}, \dots, x_{\tau(\sigma(p))}) \\ &= \Phi(x_{(\tau\sigma)(1)}, \dots, x_{(\tau\sigma)(p)}) \\ &= ((\tau\sigma)\Phi)(x_1, \dots, x_p) \end{aligned}$$

Therefore  $\tau(\sigma\Phi) = (\tau\sigma)\Phi$ .

*Remark.* By Proposition I(iii) and Proposition II, a determinant function  $\Delta \neq 0$  “determines” linear independence in the sense that  $\Delta(x_1, \dots, x_n) \neq 0$  if and only if  $x_1, \dots, x_n$  are linearly independent. By (4.8), it follows that  $\det \varphi$  “determines” whether a linear transformation  $\varphi$  preserves linear independence, i.e. whether or not  $\varphi$  is invertible.

Geometrically,  $\Delta(x_1, \dots, x_n)$  measures the oriented (signed) volume of the  $n$ -dimensional parallelepiped determined by the vectors  $x_1, \dots, x_n$ . Therefore  $\det \varphi$  is the factor by which  $\varphi$  changes oriented volume. Since a small change in the vectors  $x_1, \dots, x_n$  results in a small change in the oriented volume,  $\Delta$  is continuous.

*Remark.* We provide an alternative proof of Proposition IV. First note

$$(-1)^{j-1} \Delta(x, x_1, \dots, \widehat{x_j}, \dots, x_n) = \Delta(x_1, \dots, x, \dots, x_n)$$

where  $x$  is in the  $j$ -th position on the right.<sup>6</sup> Therefore

$$\sum_{j=1}^n (-1)^{j-1} \Delta(x, x_1, \dots, \widehat{x_j}, \dots, x_n) x_j = \Delta(x, x_2, \dots, x_n) x_1 + \dots + \Delta(x_1, \dots, x_{n-1}, x) x_n$$

---

<sup>6</sup> $\widehat{x_j}$  denotes deletion of  $x_j$  from the sequence on the left.

Viewing this as a function of  $x_1, \dots, x_n$  (that is, a set map from  $E^n \rightarrow L(E; E)$ ), it is obviously multilinear and skew symmetric (by Proposition I(ii)). Therefore if  $x_1, \dots, x_n$  are linearly dependent, it is zero (by Proposition I(iii)). If  $x_1, \dots, x_n$  are linearly independent (and hence a basis), then viewing it as a function of  $x$ , its value at  $x_i$  is just  $\Delta(x_1, \dots, x_n) x_i$  (by Proposition I(ii)), so it agrees on a basis with  $\Delta(x_1, \dots, x_n) x$  and hence is equal to it.

*Remark.* Let  $E$  be a vector space with  $\dim E = n > 1$  and  $E_1$  a subspace with  $\dim E_1 = 1$ . Let  $\Delta$  be a determinant function in  $E$  with  $\Delta(e_1, \dots, e_n) = 1$  where  $e_1 \in E_1$ . Then  $\Delta$  induces a determinant function  $\Delta_1$  in  $E/E_1$  by

$$\Delta_1(\overline{x_2}, \dots, \overline{x_n}) = \Delta(e_1, x_2, \dots, x_n)$$

with  $\Delta_1(\overline{e_2}, \dots, \overline{e_n}) = 1$ . Define  $D : E^n \rightarrow \Gamma$  by

$$D(x_1, \dots, x_n) = \sum_{j=1}^n (-1)^{j-1} \Delta_1(\overline{x_1}, \dots, \widehat{\overline{x_j}}, \dots, \overline{x_n}) \cdot \pi_1(x_j)$$

where  $\pi_1 : E \rightarrow \Gamma$  is the coordinate function for  $e_1$ . Then  $D$  is skew symmetric and  $n$ -linear with  $D(e_1, \dots, e_n) = 1$ , so  $D = \Delta$  by uniqueness (Proposition III). Therefore

$$\Delta(x_1, \dots, x_n) = \sum_{j=1}^n (-1)^{j-1} \Delta_1(\overline{x_1}, \dots, \widehat{\overline{x_j}}, \dots, \overline{x_n}) \cdot \pi_1(x_j)$$

This result, which can also be obtained by applying  $\pi_1$  to both sides of (4.6) with  $x = e_1$ , expresses a fundamental relationship between an  $n$ -dimensional determinant function and an  $(n-1)$ -dimensional one. The cofactor expansion formulas for the determinant (subsection 4.15) follow immediately.

Note that this relationship can also be exploited in reverse to *define* an  $n$ -dimensional determinant function in terms of an  $(n-1)$ -dimensional one. In fact, if  $\Phi : E^{n-1} \rightarrow \Gamma$  is a skew symmetric  $(n-1)$ -linear function in  $E$ , then there is a vector  $a \in E$  such that

$$\Phi(x_1, \dots, x_{n-1}) = \Delta(x_1, \dots, x_{n-1}, a) \quad (1)$$

To see this, define  $\varphi : E^n \rightarrow E$  by

$$\varphi(x_1, \dots, x_n) = \sum_{j=1}^n (-1)^{n-j} \Phi(x_1, \dots, \widehat{x_j}, \dots, x_n) \cdot x_j$$

Then  $\varphi$  is skew symmetric and  $n$ -linear, so  $\varphi = \Delta a$  where  $a = \varphi(e_1, \dots, e_n)$ . This means

$$\Delta(x_1, \dots, x_n)a = \sum_{j=1}^n (-1)^{n-j} \Phi(x_1, \dots, \widehat{x_j}, \dots, x_n) \cdot x_j \quad (2)$$

If  $x_1, \dots, x_{n-1}$  are linearly dependent, then (1) holds trivially, so we may assume that they are linearly independent. If  $a$  is not in their span, then (1) follows from (2) with  $x_n = a$ ; if  $a$  is in the span, then (1) follows from (2) with  $x_n$  chosen so that  $\Delta(x_1, \dots, x_n) = 1$ .

Geometrically, this result expresses a relationship between  $n$ -dimensional volume and  $(n-1)$ -dimensional volume. We also see from (1) that the latter is just the former measured *relative to a fixed vector*. The fact that the volume of an  $n$ -dimensional parallelepiped is equal to the product of the volume of any  $(n-1)$ -dimensional “base” and the corresponding “height” (subsection 7.15) is a special case.

## § 2

*Remark.* In subsection 4.6, we want a transformation  $\psi$  with  $\psi\varphi = (\det\varphi)\iota$ . We can choose a basis  $x_1, \dots, x_n$  in  $E$  with  $\Delta(x_1, \dots, x_n) = 1$ , for which we want

$$\begin{aligned} (\psi\varphi)x_i &= \psi(\varphi x_i) = (\det\varphi)x_i \\ &= (\det\varphi)\Delta(x_1, \dots, x_n)x_i \\ &= \Delta(\varphi x_1, \dots, \varphi x_n)x_i \end{aligned}$$

To obtain this, we can define

$$\psi(x) = \sum_{j=1}^n \Delta(\varphi x_1, \dots, x, \dots, \varphi x_n)x_j$$

where  $x$  is in the  $j$ -th position on the right.<sup>7</sup> Then  $\psi$  obviously satisfies the above properties, by multilinearity and skew symmetry of  $\Delta$ .

To obtain  $\psi$  in a “coordinate-free” manner (without choosing a basis), we observe that the construction on the right is multilinear and skew symmetric in  $x_1, \dots, x_n$  when viewed as a mapping  $\Phi : E^n \rightarrow L(E; E)$ . By the universal property of  $\Delta$  (Proposition III), there is a unique  $\psi \in L(E; E)$  satisfying the above; this  $\psi$  is also seen to be independent of the choice of  $\Delta$ .

---

<sup>7</sup>See the remark on Proposition IV above.

*Remark.* In subsection 4.7, observe that

$$\Delta(x_1, \dots, x_p, y_1, \dots, y_q)$$

induces a determinant function on  $E_2$  when  $x_1, \dots, x_p \in E$  are fixed, and induces a determinant function on  $E_1$  when  $y_1, \dots, y_q \in E$  are fixed. Now let  $a_1, \dots, a_p$  be a basis of  $E_1$ , so  $a_1, \dots, a_p, b_1, \dots, b_q$  is a basis of  $E$ . Then by (4.8),

$$\begin{aligned} \det \varphi \cdot \Delta(a_1, \dots, a_p, b_1, \dots, b_q) &= \Delta(\varphi_1 a_1, \dots, \varphi_1 a_p, \varphi_2 b_1, \dots, \varphi_2 b_q) \\ &= \det \varphi_1 \cdot \Delta(a_1, \dots, a_p, \varphi_2 b_1, \dots, \varphi_2 b_q) \\ &= \det \varphi_1 \cdot \det \varphi_2 \cdot \Delta(a_1, \dots, a_p, b_1, \dots, b_q) \end{aligned}$$

Since  $\Delta(a_1, \dots, a_p, b_1, \dots, b_q) \neq 0$ , it follows that  $\det \varphi = \det \varphi_1 \cdot \det \varphi_2$ . Note this result shows that

$$\det(\varphi_1 \oplus \varphi_2) = \det \varphi_1 \cdot \det \varphi_2$$

**Exercise (2).** Let  $\varphi : E \rightarrow E$  be linear with  $E_1$  a stable subspace. If  $\varphi_1 : E_1 \rightarrow E_1$  and  $\overline{\varphi} : E/E_1 \rightarrow E/E_1$  are the induced maps, then

$$\det \varphi = \det \varphi_1 \cdot \det \overline{\varphi}$$

*Proof.* Let  $e_1, \dots, e_n$  be a basis of  $E$  where  $e_1, \dots, e_p$  is a basis of  $E_1$ . Let  $\Delta \neq 0$  be a determinant function in  $E$ . First observe that

$$\Delta_1(x_1, \dots, x_p) = \Delta(x_1, \dots, x_p, \varphi e_{p+1}, \dots, \varphi e_n) \quad (1)$$

is a determinant function in  $E_1$  and

$$\Delta_2(\overline{x_{p+1}}, \dots, \overline{x_n}) = \Delta(e_1, \dots, e_p, x_{p+1}, \dots, x_n) \quad (2)$$

is a well-defined determinant function in  $E/E_1$ . Now

$$\det \overline{\varphi} \cdot \Delta_2(\overline{x_{p+1}}, \dots, \overline{x_n}) = \Delta_2(\overline{\varphi x_{p+1}}, \dots, \overline{\varphi x_n}) = \Delta_2(\overline{\varphi x_{p+1}}, \dots, \overline{\varphi x_n}) \quad (3)$$

It follows from (2) and (3) that

$$\det \overline{\varphi} \cdot \Delta(e_1, \dots, e_p, x_{p+1}, \dots, x_n) = \Delta(e_1, \dots, e_p, \varphi x_{p+1}, \dots, \varphi x_n) \quad (4)$$

Now

$$\begin{aligned} \det \varphi \cdot \Delta(e_1, \dots, e_n) &= \Delta(\varphi e_1, \dots, \varphi e_n) \\ &= \Delta_1(\varphi_1 e_1, \dots, \varphi_1 e_p) && \text{by (1)} \\ &= \det \varphi_1 \cdot \Delta_1(e_1, \dots, e_p) \\ &= \det \varphi_1 \cdot \det \overline{\varphi} \cdot \Delta(e_1, \dots, e_n) && \text{by (1), (4)} \end{aligned}$$

Since  $\Delta(e_1, \dots, e_n) \neq 0$ , the result follows.  $\square$



## § 4

*Remark.* If  $A$  is an  $n \times n$  matrix of the form

$$A = \begin{pmatrix} A_1 & \\ * & A_2 \end{pmatrix}$$

where  $A_1$  is  $p \times p$  and  $A_2$  is  $(n-p) \times (n-p)$ , then

$$\det A = \det A_1 \cdot \det A_2 \quad (1)$$

Indeed, let  $E$  be an  $n$ -dimensional vector space and  $\varphi : E \rightarrow E$  be defined by  $M(\varphi; e_1, \dots, e_n) = A$ , so  $\det A = \det \varphi$ . Let  $E_1 = \langle e_1, \dots, e_p \rangle$  and  $E_2 = \langle e_{p+1}, \dots, e_n \rangle$ . Then  $E = E_1 \oplus E_2$  and  $E_1$  is stable under  $\varphi$ . If  $\varphi_1 : E_1 \rightarrow E_1$  is the induced map, then  $A_1 = M(\varphi_1)$ , so  $\det A_1 = \det \varphi_1$ . Dually,  $E^* = E_1^* \oplus E_2^*$  where  $E_1^* = \langle e_1^*, \dots, e_p^* \rangle$  and  $E_2^* = \langle e_{p+1}^*, \dots, e_n^* \rangle$ , and  $E_2^*$  is stable under  $\varphi^*$  since

$$M(\varphi^*; e_1^*, \dots, e_n^*) = A^* = \begin{pmatrix} A_1^* & * \\ & A_2^* \end{pmatrix}$$

If  $\varphi_2^* : E_2^* \rightarrow E_2^*$  is the induced map, then  $A_2^* = M(\varphi_2^*)$  and  $\det A_2 = \det A_2^* = \det \varphi_2^*$ . So we must prove that  $\det \varphi = \det \varphi_1 \cdot \det \varphi_2^*$ .

Let  $\Delta \neq 0$  be a determinant function in  $E$  and  $\Delta^*$  its dual in  $E^*$ . We claim

$$\begin{aligned} \Delta^*(e_1^*, \dots, e_n^*) \cdot \Delta(e_1, \dots, e_p, \varphi e_{p+1}, \dots, \varphi e_n) \\ = \Delta(e_1, \dots, e_n) \cdot \Delta^*(e_1^*, \dots, e_p^*, \varphi^* e_{p+1}^*, \dots, \varphi^* e_n^*) \end{aligned} \quad (2)$$

Indeed, by (4.26) the left side of (2) is a determinant of the form

$$\begin{vmatrix} J & \\ * & B \end{vmatrix}$$

where  $J$  is the  $p \times p$  identity matrix and  $B = (\beta_i^j)$  with  $\beta_i^j = \langle e_j^*, \varphi e_i \rangle = \langle \varphi^* e_j^*, e_i \rangle$ . However, since the determinant is multilinear in its rows,<sup>8</sup> it is equal to

$$\begin{vmatrix} J & \\ & B \end{vmatrix}$$

---

<sup>8</sup>See subsection 4.9, item 4.

A similar argument shows that the same is true of the right side of (2). Now

$$\begin{aligned}
\det \varphi &= \det \varphi \cdot \Delta(e_1, \dots, e_n) \cdot \Delta^*(e_1^*, \dots, e_n^*) \\
&= \Delta(\varphi_1 e_1, \dots, \varphi_1 e_p, \varphi e_{p+1}, \dots, \varphi e_n) \cdot \Delta^*(e_1^*, \dots, e_n^*) \\
&= \det \varphi_1 \cdot \Delta(e_1, \dots, e_p, \varphi e_{p+1}, \dots, \varphi e_n) \cdot \Delta^*(e_1^*, \dots, e_n^*) \\
&= \det \varphi_1 \cdot \Delta^*(e_1^*, \dots, e_p^*, \varphi_2^* e_{p+1}^*, \dots, \varphi_2^* e_n^*) \cdot \Delta(e_1, \dots, e_n) \quad \text{by (2)} \\
&= \det \varphi_1 \cdot \det \varphi_2^* \cdot \Delta^*(e_1^*, \dots, e_n^*) \cdot \Delta(e_1, \dots, e_n) \\
&= \det \varphi_1 \cdot \det \varphi_2^*
\end{aligned}$$

The same result (1) holds when  $A$  has the form

$$A = \begin{pmatrix} A_1 & * \\ & A_2 \end{pmatrix}$$

Indeed, by the above,

$$\det A = \det A^* = \det A_1^* \cdot \det A_2^* = \det A_1 \cdot \det A_2$$

## § 5

*Remark.* Recall that the system (4.39) is equivalent to  $\varphi x = y$  where  $\varphi : \Gamma^n \rightarrow \Gamma^n$  is defined by  $M(\varphi) = (\alpha_k^j) = A$ ,  $x = (\xi^i)$ , and  $y = (\eta^j)$ . If  $\det A \neq 0$ , then  $\varphi$  is invertible and

$$x = \varphi^{-1} y = \frac{1}{\det A} \text{ad}(\varphi)(y)$$

It follows from the analysis of the adjoint matrix in subsection 4.13 that

$$\xi^i = \frac{1}{\det A} \sum_j \text{cof}(\alpha_i^j) \eta^j$$

Moreover, it follows from (4.38) that  $\sum_j \text{cof}(\alpha_i^j) \eta^j = \det A_i$  where  $A_i$  is the matrix obtained from  $A$  by replacing the  $i$ -th row with  $y$ .<sup>9</sup> Therefore

$$\xi^i = \frac{\det A_i}{\det A}$$

*Remark.* In subsection 4.14,  $\det B_i^j = \det S_i^j$  does *not* follow from (4.38), which only tells us that  $\det B_i^j = \det B_i^j$ . However, it follows from (4.16), or from our remarks in § 4 above.

---

<sup>9</sup>The cofactors of  $A_i$  and  $A$  along the  $i$ -th row agree since  $A_i$  and  $A$  agree on the other rows.

## § 6

**Exercise (5).** If  $\varphi_1 : E_1 \rightarrow E_1$  and  $\varphi_2 : E_2 \rightarrow E_2$  are linear, then

$$\chi_{\varphi_1 \oplus \varphi_2} = \chi_{\varphi_1} \chi_{\varphi_2}$$

where  $\chi_\varphi$  denotes the characteristic polynomial of  $\varphi$ .

*Proof.* This follows from the result in subsection 4.7 and the fact that

$$\varphi_1 \oplus \varphi_2 - \lambda \iota = (\varphi_1 - \lambda \iota_{E_1}) \oplus (\varphi_2 - \lambda \iota_{E_2}) \quad \square$$

**Exercise (6).** Let  $\varphi : E \rightarrow E$  be linear with  $E_1$  a stable subspace. If  $\varphi_1 : E_1 \rightarrow E_1$  and  $\bar{\varphi} : E/E_1 \rightarrow E/E_1$  are the induced maps, then

$$\chi_\varphi = \chi_{\varphi_1} \chi_{\bar{\varphi}}$$

*Proof.* This follows from problem 2 in § 2, the fact that  $\varphi - \lambda \iota_E$  restricted to  $E_1$  is just  $\varphi_1 - \lambda \iota_{E_1}$ , and  $\overline{\varphi - \lambda \iota_E} = \bar{\varphi} - \lambda \iota_{E/E_1}$ .  $\square$

*Remark.* Taking  $E_1 = \ker \varphi$ , we have  $\chi_{\varphi_1} = \chi_{0_{E_1}} = (-\lambda)^p$  where  $p = \dim E_1$ , so  $\chi_\varphi = (-\lambda)^p \chi_{\bar{\varphi}}$ .

**Exercise (7).** A linear map  $\varphi : E \rightarrow E$  is nilpotent if and only if  $\chi_\varphi = (-\lambda)^n$ .

*Proof.* If  $\varphi$  is nilpotent, we proceed by induction on  $k$  least such that  $\varphi^k = 0$ . If  $k = 1$ , the result is trivial. If  $k > 1$ , let  $E_1 = \ker \varphi$  and  $\bar{\varphi} : E/E_1 \rightarrow E/E_1$  the induced map. Then  $\bar{\varphi}^{k-1} = 0$  since

$$\bar{\varphi}^{k-1}(\bar{x}) = \overline{\varphi^{k-1}(x)} = \overline{\varphi^{k-1}(x)} = 0$$

as  $\varphi^{k-1}(x) \in E_1$ . By the induction hypothesis,  $\chi_{\bar{\varphi}} = (-\lambda)^{n-p}$  where  $p = \dim E_1$ , so by the previous problem,

$$\chi_\varphi = (-\lambda)^p (-\lambda)^{n-p} = (-\lambda)^n$$

Conversely, if  $\varphi \neq 0$  and  $\chi_\varphi = (-\lambda)^n$ , then the constant term  $\det \varphi = 0$ , so  $p > 0$  and by the previous problem  $(-\lambda)^n = (-\lambda)^p \chi_{\bar{\varphi}}$ , which implies  $\chi_{\bar{\varphi}} = (-\lambda)^{n-p}$ . By induction,  $\bar{\varphi}$  is nilpotent. If  $\bar{\varphi}^k = 0$ , then  $\varphi^{k+1} = 0$ , so  $\varphi$  is nilpotent.  $\square$

## § 7

*Remark.* It follows by symmetry from (4.61) that

$$\operatorname{tr} \varphi^* = \sum_i \langle \varphi^* e^{*i}, e_i \rangle = \sum_i \langle e^{*i}, \varphi e_i \rangle = \operatorname{tr} \varphi$$

where  $\varphi^*$  is dual to  $\varphi$ .

**Exercise (5).** If  $f : L(E; E) \rightarrow \Gamma$  is linear, there is  $\alpha \in L(E; E)$  unique with

$$f(\varphi) = \operatorname{tr}(\varphi \circ \alpha)$$

*Proof.* Since  $L(E; E)$  is finite-dimensional and dual to itself under the scalar product

$$\langle \varphi, \psi \rangle = \operatorname{tr}(\psi \circ \varphi)$$

it follows from Proposition I in chapter II, § 6 that there is  $\alpha$  unique with

$$f(\varphi) = \langle \alpha, \varphi \rangle = \operatorname{tr}(\varphi \circ \alpha) \quad \square$$

**Exercise (6).** If  $f : L(E; E) \rightarrow \Gamma$  is linear with

$$f(\psi \circ \varphi) = f(\varphi \circ \psi)$$

there is  $\lambda \in \Gamma$  with

$$f(\varphi) = \lambda \cdot \operatorname{tr} \varphi$$

*Proof.* By problem 5, there is  $\alpha$  with

$$f(\varphi) = \operatorname{tr}(\varphi \circ \alpha) = \operatorname{tr}(\alpha \circ \varphi)$$

We claim  $\alpha = \lambda \iota$  for some  $\lambda \in \Gamma$ , from which the result follows. By assumption,

$$\operatorname{tr}(\alpha \circ \varphi \circ \psi) = \operatorname{tr}(\alpha \circ \psi \circ \varphi) = \operatorname{tr}(\varphi \circ \alpha \circ \psi) \quad (1)$$

for all  $\varphi$  and  $\psi$ . Fix a basis of  $E$ . Let  $A = M(\alpha)$ , let  $R_l$  be the matrix with 1's in the  $l$ -th row and 0's elsewhere, and let  $C_k$  be the matrix with 1's in the  $k$ -th column and 0's elsewhere. By direct computation,

$$(C_k R_l A)_{ij} = 0 \quad (k \neq l)$$

while

$$(C_k A R_l)_{ij} = A_{kl}$$

It follows from (1) with  $R_l = M(\varphi)$  and  $C_k = M(\psi)$  that  $A$  is diagonal. Now let  $T_{kl}$  be the matrix with 1's in entries  $(k, l)$  and  $(l, k)$  and 0's elsewhere. It follows from (1) with  $T_{kl} = M(\varphi)$  and  $C_k = M(\psi)$  that

$$A_{ll} = \text{tr}(C_l A) = \text{tr}(C_k T_{kl} A) = \text{tr}(C_k A T_{kl}) = A_{kk}$$

Therefore the claim holds with  $\lambda = A_{11}$ . □

**Exercise (10).** If  $A : L(E; E) \rightarrow L(E; E)$  is linear and “functorial” in that

$$A(\varphi \circ \psi) = A(\varphi) \circ A(\psi) \quad \text{and} \quad A(\iota) = \iota$$

then  $\text{tr } A(\varphi) = \text{tr } \varphi$ .

*Proof.* The function  $\text{tr} \circ A : L(E; E) \rightarrow \Gamma$  is linear and

$$(\text{tr} \circ A)(\varphi \circ \psi) = \text{tr}(A(\varphi) \circ A(\psi)) = \text{tr}(A(\psi) \circ A(\varphi)) = (\text{tr} \circ A)(\psi \circ \varphi)$$

so  $\text{tr} \circ A = \lambda \cdot \text{tr}$  for some  $\lambda \in \Gamma$  by problem 6. But this implies

$$\lambda \cdot \text{tr } \iota = (\text{tr} \circ A)(\iota) = \text{tr } \iota$$

so  $\lambda = 1$  and  $\text{tr} \circ A = \text{tr}$ . □

**Exercise (12).** If  $\varphi_1 : E_1 \rightarrow E_1$  and  $\varphi_2 : E_2 \rightarrow E_2$  are linear, then

$$\text{tr}(\varphi_1 \oplus \varphi_2) = \text{tr } \varphi_1 + \text{tr } \varphi_2$$

*Proof.* Immediate since

$$M(\varphi_1 \oplus \varphi_2) = \begin{pmatrix} M(\varphi_1) & \\ & M(\varphi_2) \end{pmatrix} \quad \square$$

## § 8

*Remark.* In (4.68), if instead we define

$$\Delta_1(x_1, \dots, x_p) = \Delta(x_1, \dots, x_p, e_{p+1}, \dots, e_n)$$

then  $\Delta_1$  represents the original orientation in  $E_1$ . Indeed, in this case

$$\Delta_1(e_1, \dots, e_p) = \Delta(e_1, \dots, e_p, e_{p+1}, \dots, e_n) = \Delta_2(e_{p+1}, \dots, e_n) > 0$$

## Chapter V

### § 1

*Remark.* An algebra  $A$  is a *zero algebra* if  $xy = 0$  for all  $x, y \in A$ ; this is equivalent to  $A^2 = 0$ . As an example, *the zero algebra* is the algebra  $A = 0$ . A zero algebra is unital if and only if it is the zero algebra.

*Remark.* Let  $\varphi : A \rightarrow B$  be a homomorphism of algebras. If  $A_1$  is a subalgebra of  $A$  and  $B_1$  is a subalgebra of  $B$  and  $\varphi(A_1) \subseteq B_1$ , then the restriction  $\varphi_1 : A_1 \rightarrow B_1$  of  $\varphi$  to  $A_1, B_1$  is a homomorphism.

If  $A_1$  and  $B_1$  are *ideals*, then the induced linear map  $\bar{\varphi} : A/A_1 \rightarrow B/B_1$  is also a homomorphism since

$$\bar{\varphi}(\bar{x}\bar{y}) = \bar{\varphi}(\overline{xy}) = \overline{\varphi(xy)} = \overline{\varphi(x)\varphi(y)} = \overline{\varphi(x)}\overline{\varphi(y)} = \bar{\varphi}(\bar{x})\bar{\varphi}(\bar{y})$$

In the problems below,  $E$  is a finite-dimensional vector space.

**Exercise (12).** The mapping

$$\Phi : A(E; E) \rightarrow A(E^*; E^*)^{\text{opp}}$$

defined by  $\varphi \mapsto \varphi^*$  is an algebra isomorphism.

*Proof.*  $\Phi$  is a linear isomorphism by problem 9 of chapter II, § 6, and preserves products since  $(\varphi \circ \psi)^* = \psi^* \circ \varphi^*$ .  $\square$

**Exercise (16).** Every algebra automorphism  $\Phi : A(E; E) \rightarrow A(E; E)$  is an *inner* automorphism; that is, there exists  $\alpha \in GL(E)$  such that  $\Phi(\varphi) = \alpha\varphi\alpha^{-1}$  for all  $\varphi \in A(E; E)$ .

*Proof.* First, observe that every basis  $(e_i)$  of  $E$  induces a basis  $(\varphi_{ij})$  of  $A(E; E)$  defined by  $\varphi_{ij}(e_k) = \delta_{jk}e_i$ . This basis satisfies

$$\varphi_{ij}\varphi_{lk} = \delta_{jl}\varphi_{ik} \quad \text{and} \quad \sum_i \varphi_{ii} = \iota \tag{1}$$

Conversely, every basis satisfying these properties is induced by a basis of  $E$  in this manner (see problem 14). Moreover, any two of these bases are conjugate to each other via the change of basis transformation between their inducing bases of  $E$  (see problem 15).

Now fix  $(e_i)$  and  $(\varphi_{ij})$  as above. Since  $\Phi$  is an automorphism,  $(\Phi(\varphi_{ij}))$  is also a basis of  $A(E; E)$  which satisfies (1), so there is  $\alpha \in GL(E)$  with  $\Phi(\varphi_{ij}) = \alpha\varphi_{ij}\alpha^{-1}$  for all  $i, j$ . It follows that  $\Phi(\varphi) = \alpha\varphi\alpha^{-1}$  for all  $\varphi \in A(E; E)$ .  $\square$

*Remark.* The result is true for any nonzero endomorphism  $\Phi$ , since  $A(E; E)$  is simple (see subsection 5.12).

### § 3

*Remark.* We see the following examples of change of coefficient field of a vector space:

- Taking the underlying real space of a complex space (5.16): for example changing from  $\mathbb{C}$  over  $\mathbb{C}$  to  $\mathbb{C}$  over  $\mathbb{R}$ . In this case the dimension is doubled. Moreover, the underlying real space can be decomposed into “real” and “imaginary” parts of equal dimension (11.7).
- Complexifying a real space (2.16): for example changing from  $\mathbb{R}$  over  $\mathbb{R}$  to  $\mathbb{R}^2$  over  $\mathbb{C}$ . In this case the dimension is preserved.
- Inducing complex structure on a real space (8.21): for example changing from  $\mathbb{R}^2$  over  $\mathbb{R}$  to  $\mathbb{R}^2$  over  $\mathbb{C}$ . In this case the dimension is halved.

## Chapter VI

### § 1

*Remark.* The space of polynomials in one variable is positively graded by the degrees of monomials. More generally, the space of polynomials in  $p$  variables is  $p$ -graded by the multidegrees of monomials.

*Remark.* If  $E = \sum_{k \in G} E_k$  is a  $G$ -graded space and  $F = \sum_{k \in G} F \cap E_k$  is a  $G$ -graded subspace, then  $E/F = \sum_{k \in G} E_k / (F \cap E_k)$  is the  $G$ -graded factor space.<sup>10</sup>

*Remark.* The zero map between two  $G$ -graded vector spaces is homogeneous of every degree. A nonzero homogeneous map has a unique degree.

*Remark.* Let  $E$  and  $F$  be  $G$ -graded vector spaces. If  $\varphi : E \rightarrow F$  is linear and homogeneous of degree  $k$  and  $\varphi x$  is homogeneous of degree  $l$ , then we may assume without loss of generality that  $x$  is homogeneous of degree  $l - k$ . Indeed, writing  $x = \sum_m x_m$  with  $\deg x_m = m$ , we have  $\varphi x = \sum_m \varphi x_m$  with  $\deg(\varphi x_m) = m + k$ . Since  $\deg(\varphi x) = l$ , we must have  $\varphi x_m = 0$  for  $m \neq l - k$ , so  $\varphi x = \varphi x_{l-k}$ .

*Remark.* If  $E$  is a finite-dimensional  $G$ -graded vector space and  $\varphi : E \rightarrow E$  is linear and homogeneous with  $\deg \varphi \neq 0$ , then  $\text{tr } \varphi = 0$ .

---

<sup>10</sup>See problem 2 in chapter II, § 4.

*Proof.* Write  $E = E_{k_1} \oplus \cdots \oplus E_{k_n}$  with  $k_i \in G$  and  $d_i = \dim E_{k_i} < \infty$ . Let  $(e_{ij})$  be a basis of  $E$  such that for each  $1 \leq i \leq n$ ,  $(e_{ij})$  is a basis of  $E_{k_i}$  for  $1 \leq j \leq d_i$ . Let  $\Delta$  be a determinant function in  $E$  with  $\Delta(e_{ij}) = 1$ . Then

$$\operatorname{tr} \varphi = \sum_{i,j} \Delta(e_{11}, \dots, e_{1d_1}, \dots, \varphi(e_{ij}), \dots, e_{nd_n})$$

By assumption,  $\varphi(e_{ij}) \in E_{k_l}$  for some  $l \neq i$ , so each term in this sum is zero, and hence  $\operatorname{tr} \varphi = 0$ .  $\square$

As an example, formal differentiation in the space of polynomials of degree at most  $n$  (graded by the degrees of monomials) is homogeneous of degree  $-1$ , so has zero trace. This is also obvious from its matrix representation with respect to the standard basis.

**Exercise (6).** Let  $E, E^*$  and  $F, F^*$  be pairs of dual  $G$ -graded vector spaces and let  $\varphi : E \rightarrow F$  and  $\varphi^* : E^* \leftarrow F^*$  be dual linear maps. If  $\varphi$  is homogeneous of degree  $k$ , then  $\varphi^*$  is homogeneous of degree  $-k$ .

*Proof.* We have direct sum decompositions

$$E = \sum_{m \in G} E_m \quad E^* = \sum_{m \in G} E^{*m}$$

and

$$F = \sum_{n \in G} F_n \quad F^* = \sum_{n \in G} F^{*n}$$

where the pairs  $E_m, E^{*m}$  and  $F_n, F^{*n}$  are dual for all  $m, n$  under the restrictions of the scalar products between  $E, E^*$  and  $F, F^*$ , respectively (see subsection 6.5). We also have  $\varphi E_m \subseteq F_{m+k}$  for all  $m$ . We must prove  $\varphi^* F^{*n} \subseteq E^{*n-k}$  for all  $n$ .

Let  $y^* \in F^{*n}$  and  $x \in E$ . Write  $x = \sum_m x_m$  where  $x_m \in E_m$ . Then

$$\langle \varphi^* y^*, x \rangle = \langle y^*, \varphi x \rangle = \sum_m \langle y^*, \varphi x_m \rangle = \langle y^*, \varphi x_{n-k} \rangle = \langle \varphi^* y^*, x_{n-k} \rangle$$

which implies

$$\langle \varphi^* y^*, x - \pi_{n-k} x \rangle = 0 \tag{1}$$

where  $\pi_{n-k} : E \rightarrow E_{n-k}$  is the canonical projection. Now write  $\varphi^* y^* = \sum_m x^{*m}$  where  $x^{*m} \in E^{*m}$ . We claim  $x^{*m} = 0$  for all  $m \neq n-k$ . Indeed, for  $m \neq n-k$  and  $x \in E_m$  we have  $\pi_{n-k} x = 0$ , so by (1)

$$\langle x^{*m}, x \rangle = \sum_p \langle x^{*p}, x \rangle = \langle \varphi^* y^*, x \rangle = 0$$

Therefore  $x^{*m} = 0$ . It follows that  $\varphi^* y^* = x^{*n-k} \in E^{*n-k}$ , as desired.  $\square$



**Exercise (8).** Let  $E, E^*$  be a pair of almost finite dual  $G$ -graded vector spaces. If  $F$  is a  $G$ -graded subspace of  $E$ , then  $F^\perp$  is a  $G$ -graded subspace of  $E^*$  and  $F^{\perp\perp} = F$ .

*Proof.* We have direct sums  $E = \sum_{m \in G} E_m$  and  $E^* = \sum_{m \in G} E^{*m}$  where the pairs  $E_m, E^{*m}$  are dual under the restrictions of the scalar product between  $E, E^*$  and  $\dim E_m = \dim E^{*m} < \infty$  for all  $m$ . By assumption,  $F = \sum_{m \in G} F \cap E_m$ . We must prove

$$F^\perp = \sum_{m \in G} F^\perp \cap E^{*m} \quad (1)$$

Let  $x^* \in F^\perp$  and write  $x^* = \sum_m x^{*m}$  where  $x^{*m} \in E^{*m}$ . We claim  $x^{*n} \in F^\perp$  for all  $n$ . Indeed, if  $x \in F$ , write  $x = \sum_m x_m$  where  $x_m \in F \cap E_m$ . Then

$$\langle x^{*n}, x \rangle = \sum_m \langle x^{*n}, x_m \rangle = \langle x^{*n}, x_n \rangle = \sum_m \langle x^{*m}, x_n \rangle = \langle x^*, x_n \rangle = 0$$

This establishes (1). By symmetry, we have

$$F^{\perp\perp} = \sum_{m \in G} F^{\perp\perp} \cap E_m \quad (2)$$

We claim  $F^{\perp\perp} \cap E_n \subseteq F \cap E_n$  for all  $n$ . To prove this, we first show

$$F^{\perp\perp} \cap E_n \subseteq (F \cap E_n)^{\perp_n \perp_n} \quad (3)$$

where  $\perp_n$  is taken relative to the scalar product between  $E_n, E^{*n}$ . Indeed, let  $x \in F^{\perp\perp} \cap E_n$  and  $x^* \in (F \cap E_n)^{\perp_n \perp_n} \subseteq E^{*n}$ . If  $y \in F$ , write  $y = \sum_m y_m$  where  $y_m \in F \cap E_m$ . Then

$$\langle x^*, y \rangle = \sum_m \langle x^*, y_m \rangle = \langle x^*, y_n \rangle = 0$$

This implies  $x^* \in F^\perp$ , which implies  $\langle x^*, x \rangle = 0$ , which in turn implies (3). Now  $(F \cap E_n)^{\perp_n \perp_n} = F \cap E_n$  since  $\dim E_n < \infty$ , which establishes the claim. Finally, it follows from (2) that  $F^{\perp\perp} = F$ .  $\square$

## § 2

**Exercise (1).** Let  $A$  be a  $G$ -graded algebra. If  $x \in A$  is an invertible element homogeneous of degree  $k$ , then  $x^{-1}$  is homogeneous of degree  $-k$ . If  $A$  is nonzero and positively graded, then  $k = 0$ .

*Proof.* Write  $A = \sum_{m \in G} A_m$  and  $x^{-1} = \sum_m x_m$  with  $x_m \in A_m$ . Then

$$e = xx^{-1} = \sum_m xx_m$$

Since  $\deg e = 0$  and  $\deg(xx_m) = m + k$ , it follows that  $xx_m = 0$  for all  $m \neq -k$ . Therefore  $e = xx_{-k}$  and  $x^{-1} = x_{-k}$ , so  $x^{-1}$  is homogeneous of degree  $-k$ .

If  $A \neq 0$ , then  $x \neq 0$  and  $x^{-1} \neq 0$ , so  $A_k \neq 0$  and  $A_{-k} \neq 0$ . If  $A$  is positively graded, this forces  $k = 0$ .  $\square$

**Exercise (4).** Let  $E$  be a  $G$ -graded vector space. Then the subspace  $A_G(E; E)$  of  $A(E; E)$  generated by homogeneous linear transformations of  $E$  forms a  $G$ -graded subalgebra of  $A(E; E)$ .

*Proof.* First observe that  $A_G(E; E)$  is naturally graded as a vector space by the degrees of homogeneous transformations (see problem 3). If  $\varphi, \psi \in A_G(E; E)$  are homogeneous with  $\deg \varphi = m$  and  $\deg \psi = n$ , then it is obvious that  $\varphi\psi$  is homogeneous with  $\deg(\varphi\psi) = m + n$ . It follows from this that  $A_G(E; E)$  is a  $G$ -graded subalgebra.  $\square$

**Exercise (7).** Let  $E, E^*$  be a pair of almost finite dual  $G$ -graded vector spaces. Then the mapping

$$\Phi: A_G(E; E) \rightarrow A_G(E^*; E^*)^{\text{opp}}$$

defined by  $\varphi \mapsto \varphi^*$  is an algebra isomorphism.

*Proof.*  $\Phi$  is well defined by problems 6 and 10 of § 1, and is an isomorphism by problem 12 of chapter V, § 1.  $\square$

## Chapter VII

**In this chapter, all vector spaces are real.**

### § 1

*Remark.* The Riesz representation theorem shows that for a finite-dimensional inner product space  $E$ , there is a *natural* isomorphism between  $E$  and its dual space  $L(E)$  given by  $x \mapsto (x, -)$ . This is unlike for a finite-dimensional vector space, where the isomorphism is in general non-natural, and means we may naturally *identify* a vector with its dual vector.

*Remark.* If  $E$  is a finite-dimensional inner product space and  $E_1$  is a subspace of  $E$ , then  $E$  induces an inner product in  $E/E_1$  by

$$(\bar{x}, \bar{y}) = (x, y)$$

where  $x, y$  are the unique representatives of  $\bar{x}, \bar{y}$  in  $E_1^\perp$  (see problem 5).

*Proof.* The only thing to check is bilinearity, which follows from the fact that  $E_1^\perp$  is a subspace of  $E$ .  $\square$

**Exercise (5).** If  $E$  is a finite-dimensional inner product space and  $E_1$  is a subspace of  $E$ , then every element of  $E/E_1$  has exactly one representative in  $E_1^\perp$ .

*Proof.* By duality,  $\dim E = \dim E_1 + \dim E_1^\perp$ , and by definiteness of the inner product,  $E_1 \cap E_1^\perp = 0$ , so  $E = E_1 \oplus E_1^\perp$ . For  $x + E_1 \in E/E_1$ , let  $y = \pi(x) - x$  where  $\pi$  is the canonical projection onto  $E_1^\perp$ . Then  $y \in E_1^\perp$  so  $x + y \in x + E_1$  and  $x + y \in E_1^\perp$ , as desired. If  $z \in E_1$  and  $x + z \in E_1^\perp$ , then

$$y - z = (y + x) - (x + z) \in E_1 \cap E_1^\perp = 0$$

so  $y = z$ , establishing uniqueness.  $\square$

## § 2

*Remark.* Let  $E$  be an inner product space of finite dimension  $n$ . We provide an inductive proof of the existence of an orthonormal basis in  $E$ .

For  $n = 0, 1$ , the result is trivial. For  $n > 1$ , fix a unit vector  $e_1 \in E$  and let  $E_1 = \langle e_1 \rangle$ . By induction, there is an orthonormal basis  $\bar{e}_2, \dots, \bar{e}_n$  in the induced inner product space  $E/E_1$  (see above). Letting  $e_2, \dots, e_n$  be the representatives in  $E_1^\perp$ , it follows that  $e_1, \dots, e_n$  is an orthonormal basis in  $E$ .

*Remark.* In the Gram-Schmidt process, we can just let  $e_1 = a_1/|a_1|$  and

$$e_k = \frac{a_k - (a_k, e_1)e_1 - \dots - (a_k, e_{k-1})e_{k-1}}{|a_k - (a_k, e_1)e_1 - \dots - (a_k, e_{k-1})e_{k-1}|} \quad (k = 2, \dots, n)$$

At each step, we compute the difference between the current vector  $a_k$  and its orthogonal projection onto the subspace generated by the previous vectors, then normalize (compare (7.19)).

*Remark.* If  $E$  is a finite-dimensional inner product space and  $\varphi : E \rightarrow E$  is linear, then the dual transformation  $\varphi^* : E \rightarrow E$  satisfies

$$(\varphi^* x, y) = (x, \varphi y)$$

If  $\varphi$  preserves inner products, then it also preserves orthonormal bases and is invertible. In this case,  $\varphi^{-1}$  also preserves inner products, so

$$(\varphi^{-1} x, y) = (\varphi^{-1} x, \varphi^{-1} \varphi y) = (x, \varphi y)$$

and it follows that  $\varphi^{-1} = \varphi^*$ . Conversely if  $\varphi^{-1} = \varphi^*$ , then

$$(\varphi x, \varphi y) = (\varphi^* \varphi x, y) = (x, y)$$

so  $\varphi$  preserves inner products.

Such a  $\varphi$  is called an *orthogonal* transformation. Note  $\varphi$  is orthogonal if and only if  $M(\varphi)$  is an orthogonal matrix relative to an orthonormal basis.

### § 3

*Remark.* A determinant function  $\Delta$  in an inner product space  $E$  is normed if and only if  $|\Delta(e_1, \dots, e_n)| = 1$  for any orthonormal basis  $e_1, \dots, e_n$ . If  $E$  is oriented and  $\Delta$  is the normed determinant function representing the orientation, then  $\Delta(e_1, \dots, e_n) = 1$  if  $e_1, \dots, e_n$  is positive. Geometrically, this is just *the* oriented volume function in  $E$ . By comparing (4.26) and (7.23), we see that *the normed determinant functions are precisely the self-dual determinant functions*.

If  $\Delta_0 \neq 0$  is a determinant function in  $E$  and  $\Delta_0^*$  is its dual, then  $\Delta_0^* = \alpha \Delta_0$  for some real number  $\alpha$  by uniqueness of  $\Delta_0$ , so

$$\alpha \Delta_0(x_1, \dots, x_n) \Delta_0(y_1, \dots, y_n) = \det(x_i, y_j) \quad (x_i, y_i \in E)$$

Substituting  $x_i = y_i = e_i$  shows that  $\alpha > 0$ . It follows that  $\Delta_1 = \pm \sqrt{\alpha} \cdot \Delta_0$  is a normed determinant function.

*Remark.* It follows from (7.37) that if  $x$  and  $y$  are linearly independent, then

$$|x \times y| = \Delta(x, y, z)$$

where  $z = (x \times y) / |x \times y|$ . Since  $z$  is a unit vector orthogonal to  $x$  and  $y$ , it follows (subsection 7.15) that  $\Delta(x, y, z)$  is just the area of the parallelogram determined by  $x$  and  $y$ . In other words,

$$|x \times y| = |x||y| \sin \theta$$

where  $0 \leq \theta \leq \pi$  is the angle between  $x$  and  $y$ . This last equation obviously still holds if  $x$  and  $y$  are linearly dependent but nonzero.

**Exercise (12).** For vectors  $x_1, \dots, x_p$ ,

$$G(x_1, \dots, x_p) \leq |x_1|^2 \cdots |x_p|^2 \quad (1)$$

Additionally

$$\left| \begin{array}{ccc} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{array} \right|^2 \leq \sum_{k=1}^n |a_{1k}|^2 \cdots \sum_{k=1}^n |a_{nk}|^2 \quad (2)$$

*Proof.* By induction with the “base times height” rule for volume (subsection 7.15),  $V(u_1, \dots, u_p) \leq 1$  for linearly independent *unit* vectors  $u_1, \dots, u_p$ .

Now if  $x_1, \dots, x_p$  are linearly dependent, then  $G(x_1, \dots, x_p) = 0$  and (1) holds trivially. Otherwise, setting  $u_i = x_i/|x_i|$  we have

$$\sqrt{G(x_1, \dots, x_p)} = V(x_1, \dots, x_p) = |x_1| \cdots |x_p| \cdot V(u_1, \dots, u_p) \leq |x_1| \cdots |x_p|$$

Squaring both sides yields (1). Since the determinant of a matrix is a (normed) determinant function of the rows, (2) follows.  $\square$

*Remark.* These results both simply express an upper bound for the volume of a parallelepiped in terms of the lengths of the edges.

## § 4

*Remark.* The covariant components of a vector are just its coordinates in the dual space, which are also just the entries of its matrix (as a linear function). They are called “covariant” because they vary in the same way as basis vectors in the original space under a change of basis (see subsections 3.13–14). This is unlike the regular components of the vector, which vary inversely to the basis vectors and may be called the *contravariant components*.

To see why the covariant components of a vector are its inner products with the basis vectors, observe from (7.42) and (7.49) that

$$(x, e_\nu) = \langle \tau x, e_\nu \rangle = \sum_{\lambda} \xi_{\lambda} \langle e^{*\lambda}, e_\nu \rangle = \xi_{\nu}$$

Dually, the contravariant components of a vector are its inner products with the dual basis vectors:

$$(\tau x, e^{*\nu}) = \xi^{\nu}$$

*Remark.* If  $e_v$  is an orthonormal basis of  $E$ , then the dual basis  $e^{*v} = \tau e_v$  is an orthonormal basis of  $E^*$ .

## § 5

*Remark.* In subsection 7.22, observe that  $Q$  is just the set of unit vectors under the norm defined in example 3 in subsection 7.20. Since norm functions are continuous under the natural topology,  $Q$  is closed under this topology, and  $Q$  is also clearly bounded under this topology. It follows that  $Q$  is compact under this topology since  $E$  is finite-dimensional.

## § 6

*Remark.* If  $x = \lambda e + x_1$  and  $y = \mu e + y_1$  are quaternions ( $\lambda, \mu \in \mathbb{R}$  and  $x_1, y_1 \in E_1$ ), then by the definition of quaternion multiplication

$$xy = (\lambda\mu - (x_1, y_1))e + \lambda y_1 + \mu x_1 + x_1 \times y_1$$

*Remark.* In subsection 7.24, in the proof of Lemma I, observe that the result holds trivially if  $y = \pm x$  by taking  $\lambda = \mp 1$ . If  $y \neq \pm x$ , then  $e \neq 0$ . Suppose

$$\alpha x + \beta y + \gamma e = 0 \quad (\alpha, \beta, \gamma \in \mathbb{R})$$

If  $\alpha \neq 0$ , then  $x = \beta_1 y + \gamma_1 e$  for  $\beta_1, \gamma_1 \in \mathbb{R}$ , so

$$-e = x^2 = (\beta_1 y + \gamma_1 e)^2 = 2\beta_1 \gamma_1 y + (\gamma_1^2 - \beta_1^2)e$$

which implies

$$2\beta_1 \gamma_1 y = (\beta_1^2 - \gamma_1^2 - 1)e$$

If  $\beta_1 = 0$ , it follows that  $\gamma_1^2 = -1$ , which is impossible. If  $\gamma_1 = 0$ , it follows that  $\beta_1 = \pm 1$ , so  $x = \pm y$  contrary to assumption. Therefore  $\beta_1 \gamma_1 \neq 0$ , so  $y = \delta e$  for  $\delta \in \mathbb{R}$ . But then  $-e = y^2 = \delta^2 e$ , so  $\delta^2 = -1$ , which is impossible. It follows that  $\alpha = 0$ . Similarly  $\beta = 0$ . Finally,  $\gamma = 0$ . This result shows that the vectors  $x, y, e$  are linearly independent.

Now  $x + y$  and  $x - y$  are roots of polynomials of degree 1 or 2, but the linear independence of  $x, y, e$  implies that these polynomials must have degree 2, so (7.60) and (7.61) follow.

## Chapter VIII

**In this chapter, all vector spaces are real and finite-dimensional.**

*Remark.* In this chapter, it is useful to think intuitively of transformations like complex numbers, which induce transformations of the complex plane through multiplication. Under this analogy, adjoints correspond to complex conjugates; self-adjoint transformations, to real numbers; positive transformations, to non-negative real numbers; skew transformations, to purely imaginary numbers; and isometries, to complex numbers on the unit circle.

### § 1

*Remark.* In subsection 8.2, if the bases  $(x_\nu)$  and  $(y_\mu)$  are orthonormal, then they are self-dual,<sup>11</sup> so  $M(\tilde{\varphi}, y_\mu, x_\nu) = M(\varphi, x_\nu, y_\mu)^*$  by (3.4). It follows that

$$\tilde{\alpha}_\mu^\rho = M(\tilde{\varphi})_\mu^\rho = M(\varphi)_\rho^\mu = \alpha_\rho^\mu$$

*Remark.* In subsection 8.4, the natural isomorphism  $B(E, E) \cong L(E; E)$  follows from Proposition III in chapter II, § 6. There is also a natural isomorphism

$$B(E, E) \cong L(E; L(E))$$

given by

$$\Phi \mapsto (x \mapsto (y \mapsto \Phi(x, y))) \quad (1)$$

Recall from subsection 7.18 the natural isomorphism  $\tau : E \rightarrow L(E)$  given by  $x \mapsto (x, -)$ , which maps each vector to its dual vector. For a linear transformation  $\varphi : E \rightarrow E$  and its corresponding bilinear function  $\Phi$  defined by  $\Phi(x, y) = (\varphi x, y)$ , the linear map  $\Phi'$  corresponding to  $\Phi$  under the isomorphism (1) is  $\tau\varphi$ :

$$\begin{array}{ccc} x & \xrightarrow{\varphi} & \varphi x \\ & \searrow \Phi' & \downarrow \tau \\ & & (\varphi x, -) \end{array}$$

---

<sup>11</sup>See problem 2 in chapter VII, § 2.

In these ways,  $\Phi$  is naturally dual to  $\varphi$ . Properties of  $\Phi$  naturally correspond to those of  $\varphi$ ; for example,  $\Phi$  is symmetric if and only if  $\varphi$  is symmetric (self-adjoint), and  $\Phi$  is skew symmetric if and only if  $\varphi$  is skew symmetric. Since

$$\Phi(x, y) = (\varphi x, y) = (x, \tilde{\varphi} y)$$

we see that  $\varphi$  is “left-dual” to  $\Phi$ , while  $\tilde{\varphi}$  is “right-dual” to  $\Phi$ .

*Remark.* In subsection 8.5, observe for a direct sum  $E = \sum E_i$  and  $\varphi = \sum \varphi_i$  with  $\varphi_i : E_i \rightarrow E_i$ , if  $E_i$  is stable under  $\tilde{\varphi}$  then the restriction of  $\tilde{\varphi}$  to  $E_i$  is the adjoint of  $\varphi_i$ . In other words, if  $\tilde{\varphi}_i$  denotes the restriction, then

$$\tilde{\varphi}_i = \tilde{\varphi}_i$$

Indeed, for  $x, y \in E_i$ ,

$$(\tilde{\varphi}_i x, y) = (\tilde{\varphi} x, y) = (x, \varphi y) = (x, \varphi_i y)$$

It follows that  $\tilde{\varphi} = \sum \tilde{\varphi}_i$ . If additionally the  $E_i$  are orthogonal, we see that  $\varphi$  is normal if and only if each  $\varphi_i$  is normal.

## § 2

*Remark.* Geometrically, a self-adjoint transformation just independently scales the axes in some system of orthogonal axes for the space.

*Remark.* For any transformation  $\varphi$ , the transformations

$$\varphi + \tilde{\varphi} \quad \varphi \tilde{\varphi} \quad \tilde{\varphi} \varphi$$

are self-adjoint.

*Remark.* If  $E = E_1 \oplus E_2$  with  $E_1 \perp E_2$ , and  $\varphi_1 : E_1 \rightarrow E_1$  and  $\varphi_2 : E_2 \rightarrow E_2$  are self-adjoint, then  $\varphi = \varphi_1 \oplus \varphi_2$  is self-adjoint.

*Proof.* For  $x = x_1 + x_2$  and  $y = y_1 + y_2$  with  $x_i, y_i \in E_i$ ,

$$\begin{aligned} (\varphi x, y) &= (\varphi_1 x_1 + \varphi_2 x_2, y_1 + y_2) \\ &= (\varphi_1 x_1, y_1) + (\varphi_2 x_2, y_2) \\ &= (x_1, \varphi_1 y_1) + (x_2, \varphi_2 y_2) \\ &= (x_1 + x_2, \varphi_1 y_1 + \varphi_2 y_2) \\ &= (x, \varphi y) \end{aligned}$$

□



*Remark.* We provide an alternative proof of the existence of eigenvectors for self-adjoint transformations (subsection 8.6).

Let  $\varphi: E \rightarrow E$  be a transformation. First observe that for  $u \neq 0$ ,  $\varphi u = \lambda u$  for some  $\lambda$  if and only if  $\varphi u \in u^{\perp\perp}$ , which is true if and only if  $(\varphi u, v) = 0$  for all  $v \in u^\perp$ . This motivates us to study the bilinear function

$$\Phi(x, y) = (\varphi x, y)$$

We may assume without loss of generality that  $|u| = |v| = 1$ . Consider a unit vector  $w$  coincident with  $u$  and moving in the direction of  $v$ :

$$w(t) = (\cos t)u + (\sin t)v \quad 0 \leq t \leq \pi/2$$

Note  $|w(t)| = 1$  and  $w(0) = u$ . Also

$$w'(t) = (-\sin t)u + (\cos t)v$$

so  $w'(0) = v$ . This means  $\Phi(u, v) = \Phi(w(0), w'(0))$ . If  $\varphi$  is self-adjoint, then  $\Phi$  is symmetric and

$$2 \cdot \Phi(u, v) = \Phi(w'(0), w(0)) + \Phi(w(0), w'(0)) = (\Psi \circ w)'(0)$$

where  $\Psi(x) = \Phi(x, x)$  is the quadratic function associated with  $\Phi$ . Now we have  $(\Psi \circ w)'(0) = 0$  if  $t = 0$  minimizes (or maximizes)  $\Psi \circ w$ , which is true if  $w(0) = u$  minimizes (or maximizes)  $\Psi$ . But we know such a  $u$  exists since the unit sphere is compact in  $E$ , and the argument above then shows that  $u$  is an eigenvector of  $\varphi$ . Note that

$$\Psi(u) = \Phi(u, u) = (\varphi u, u) = (\lambda u, u) = \lambda(u, u) = \lambda$$

is the eigenvalue, which is the minimum (resp. maximum) eigenvalue of  $\varphi$ . In summary: the quadratic function associated with a self-adjoint transformation attains its extreme values on the unit sphere at eigenvectors, where its values are the extreme eigenvalues.

**Exercise (5).** Every positive transformation  $\varphi$  has a unique positive square root (that is, a positive transformation  $\psi$  such that  $\psi^2 = \varphi$ ).

*Proof.* There is an orthonormal basis  $e_1, \dots, e_n$  of eigenvectors of  $\varphi$ , so that  $\varphi e_i = \lambda_i e_i$  for  $\lambda_i \in \mathbb{R}$ . Now  $\lambda_i = (e_i, \varphi e_i) \geq 0$ , so setting  $\psi e_i = \sqrt{\lambda_i} e_i$  it follows that  $\psi$  is positive with  $\psi^2 = \varphi$ .

If  $\psi_1$  is positive with  $\psi_1^2 = \varphi$ , then the eigenvalues of  $\psi_1$  must be  $\sqrt{\lambda_i}$ . Also, if  $\psi_1 x = \sqrt{\lambda_i} x$ , then  $\varphi x = \lambda_i x$ , so  $E_{\psi_1}(\sqrt{\lambda_i}) \subseteq E_\varphi(\lambda_i)$ . By (8.21), it follows that  $E_{\psi_1}(\sqrt{\lambda_i}) = E_\varphi(\lambda_i)$ , so  $\psi_1 e_i = \sqrt{\lambda_i} e_i$  and  $\psi_1 = \psi$ .  $\square$

### § 3

*Remark.* In subsection 2.19, we see that a linear transformation  $\pi$  is a projection operator (that is,  $\pi^2 = \pi$ ) if and only if  $\pi = 0_{\ker \pi} \oplus \iota_{\operatorname{Im} \pi}$ . In subsection 8.11, we see that  $\pi$  is additionally an *orthogonal* projection if and only if  $\ker \pi \perp \operatorname{Im} \pi$ . In summary, for a projection operator  $\pi$ , the following are equivalent:

- $\ker \pi \perp \operatorname{Im} \pi$ .
- $\pi$  is self-adjoint.
- $\pi$  is normal.

### § 4

*Remark.* Geometrically, a skew transformation, apart from possibly killing off part of the space, induces scaled 90-degree rotations<sup>12</sup> on orthogonal stable planes in the space.

*Remark.* For any transformation  $\varphi$ , the transformation  $\varphi - \tilde{\varphi}$  is skew. Also

$$\varphi = \frac{1}{2}(\varphi + \tilde{\varphi}) + \frac{1}{2}(\varphi - \tilde{\varphi})$$

uniquely represents  $\varphi$  as a sum of self-adjoint and skew transformations.

*Remark.* If  $\varphi$  is a skew transformation of  $E$  and  $F$  is a stable subspace of  $E$ , then  $F^\perp$  is also stable.

*Proof.* If  $x \in F$  and  $y \in F^\perp$ , then  $(x, \varphi y) = -(\varphi x, y) = 0$ . □

*Remark.* In subsection 8.16, the proof of the normal form (8.35) is incorrect because it is not true in general that the  $a_n$  defined form an orthonormal basis of the space. For example in  $\mathbb{R}^4$ , if we define the transformation  $\psi$  by

$$e_1 \mapsto e_2 \quad e_2 \mapsto -e_1 \quad e_3 \mapsto e_4 \quad e_4 \mapsto -e_3$$

where  $e_i$  is the  $i$ -th standard basis vector, then  $\psi$  is skew and  $\varphi = \psi^2 = -\iota$  is diagonalized by the standard basis. If we follow the proof for this example, we get  $a_1 = e_1$ ,  $a_2 = \psi e_1 = e_2$ ,  $a_3 = e_2$ , and  $a_4 = \psi e_2 = -e_1$ , so the  $a_n$  do not form a basis of  $\mathbb{R}^4$ .

---

<sup>12</sup>No one calls these “scrotations”.

To prove the result, let  $\lambda_1, \dots, \lambda_r$  be the distinct eigenvalues of  $\varphi = \psi^2$  and write the orthogonal decomposition (8.21)

$$E = E_1 \oplus \dots \oplus E_r$$

where  $E_i$  is the eigenspace of  $\varphi$  corresponding to  $\lambda_i$ . Observe that  $E_i$  is stable under  $\psi$ , for if  $x \in E_i$  then

$$\varphi(\psi x) = \psi(\varphi x) = \psi(\lambda_i x) = \lambda_i(\psi x)$$

so  $\psi x \in E_i$ . As in the book, the  $\lambda_i$  are negative or zero. If  $\lambda_i < 0$ , construct a basis for  $F = E_i$  in the following way: let  $a_1$  be an arbitrary unit vector in  $F$  and  $a_2 = \kappa_i^{-1} \psi a_1$  where  $\kappa_i = \sqrt{-\lambda_i}$ . It is immediate that  $a_2$  is a unit vector in  $F$  orthogonal to  $a_1$  and  $H = \langle a_1, a_2 \rangle$  is stable under  $\psi$ . Let  $H^\perp$  be the orthogonal complement of  $H$  in  $F$ . If  $H^\perp = 0$ , take the basis  $a_1, a_2$ . Otherwise, adjoin to  $a_1, a_2$  the result of applying this procedure recursively to  $F = H^\perp$  (recall that  $H^\perp$  is stable under  $\psi$  by the remark above). If  $\lambda_i = 0$ , assume that  $i = r$  and choose any orthonormal basis of  $E_r$ , noting that  $\psi$  is zero on  $E_r$  since

$$|\psi x|^2 = (\psi x, \psi x) = -(x, \varphi x) = 0$$

for  $x \in E_r$ . Combine the resulting bases of the  $E_i$  to obtain an orthonormal basis of  $E$  with respect to which the matrix of  $\psi$  has the form (8.35).<sup>13</sup>

**Exercise (1).** If  $\varphi$  is a skew transformation of a plane, then

$$(\varphi x, \varphi y) = \det \varphi \cdot (x, y)$$

*Proof.* If  $\varphi = 0$ , then the result is trivial; if  $\varphi \neq 0$ , then  $\varphi$  is an automorphism since it has even rank. Now  $\Delta(x, y) = (\varphi x, y)$  is a determinant function, so

$$(\varphi^2 x, \varphi y) = \Delta(\varphi x, \varphi y) = \det \varphi \cdot \Delta(x, y) = \det \varphi \cdot (\varphi x, y)$$

Substituting  $\varphi^{-1} x$  for  $x$  yields the result. □

**Exercise (2 - Skew transformations of 3-space).** Let  $E$  be an oriented Euclidean 3-space.

(i) For  $a \in E$ ,  $\varphi_a(x) = a \times x$  is a skew transformation of  $E$ .

---

<sup>13</sup>See <https://math.stackexchange.com/q/3402347>.

- (ii) For  $a, b \in E$ ,  $\varphi_{a \times b} = \varphi_a \varphi_b - \varphi_b \varphi_a$ .
- (iii) If  $\varphi$  is a skew transformation of  $E$ , then  $\varphi = \varphi_a$  for a unique  $a \in E$ .
- (iv) The vector  $a$  in (iii) is given by

$$a = \alpha_{23}e_1 + \alpha_{31}e_2 + \alpha_{12}e_3$$

where  $e_1, e_2, e_3$  is a positive orthonormal basis of  $E$  and  $(\alpha_{ij}) = M(\varphi; e_i)$ .

- (v) If  $a \neq 0$ , then  $\ker \varphi_a = \langle a \rangle$  and  $a^\perp$  is stable under  $\varphi_a$ .
- (vi) If  $e_1, e_2$  are normal such that  $e_1, e_2, a$  is a positive orthogonal basis of  $E$ , then

$$M(\varphi_a; e_1, e_2, a) = \begin{pmatrix} 0 & |a| & 0 \\ -|a| & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

*Proof.*

- (i) It is a transformation since the cross product is bilinear, and it is skew since  $(x, a \times x) = 0$  for all  $x \in E$ .
- (ii) By the vector triple product formula (7.41),

$$(a \times b) \times x = a \times (b \times x) - b \times (a \times x)$$

- (iii) For uniqueness, note that if  $\varphi_a = \varphi_b$ , then  $a \times x = b \times x$  for all  $x \in E$ , so  $(a - b) \times x = 0$  for all  $x \in E$ , so  $a - b = 0$  and  $a = b$ .

For existence, note that the bilinear function  $(\varphi x, y)$  is skew symmetric, and hence a determinant function in any plane in  $E$ . Define

$$\Phi(x, y, z) = (\varphi y, z)x + (\varphi z, x)y + (\varphi x, y)z$$

Then  $\Phi$  is clearly trilinear and skew symmetric. By the universal property of determinants if  $\Delta$  is the normed determinant function representing the orientation in  $E$  and  $e_1, e_2, e_3$  is an orthonormal basis with  $\Delta(e_1, e_2, e_3) = 1$ , then  $\Phi = \Delta a$  where  $a = \Phi(e_1, e_2, e_3)$ .<sup>14</sup> By direct computation it is easily verified that  $a \times e_i = \varphi e_i$ , so  $\varphi = \varphi_a$ .

---

<sup>14</sup>See Proposition III in subsection 4.3.

(iv) By the proof of (iii), noting that  $a = \Phi(e_1, e_2, e_3)$  and  $\alpha_{ij} = (\varphi e_i, e_j)$ .

(v) By the fact that  $a \times x = 0$  if and only if  $a$  and  $x$  are linearly dependent.

(vi) By direct computation. □

*Remark.* The definition of  $\Phi$  in (iii) is motivated by the fundamental relationship between  $n$ -dimensional and  $(n - 1)$ -dimensional determinant functions discussed in § 1 of chapter IV above. If  $\varphi$  is skew, then its dual bilinear function  $\Delta_1(x, y) = (\varphi x, y)$  is a determinant function in any plane in  $E$ . If  $\varphi = \varphi_a$ , then

$$\Delta_1(x, y) = (a \times x, y) = \Delta(a, x, y)$$

where  $\Delta$  is the normed determinant function representing the orientation in  $E$ . In other words,  $\Delta_1$  is just the 2-dimensional area function induced by the 3-dimensional volume function  $\Delta$  and the vector  $a$ . Therefore if we *define* a 3-dimensional volume map  $\Phi$  in terms of  $\Delta_1$ , we should expect that  $\Phi = \Delta a$ .

This result, which shows that any skew linear *transformation* of the space can be represented by a unique vector under the *cross* product, is analogous to the Riesz representation theorem, which shows that any linear *function* of the space can be represented by a unique vector under the *inner* product.

**Exercise (3).** If  $\varphi \neq 0$  and  $\psi$  are skew transformations of an oriented Euclidean 3-space with  $\ker \varphi \subseteq \ker \psi$ , then  $\psi = \lambda \varphi$  for some  $\lambda \in \mathbb{R}$ .

*Proof.* By the previous problem, we can write  $\varphi(x) = a \times x$  and  $\psi(x) = b \times x$  with  $a \neq 0$ . By assumption,  $a$  and  $b$  are orthogonal to the same plane, so  $b = \lambda a$  for some  $\lambda \in \mathbb{R}$  and hence  $\psi = \lambda \varphi$ . □

**Exercise (4).**

$$(a_1 \times a_2) \times a_3 = a_2(a_1, a_3) - a_1(a_2, a_3)$$

*Proof.* Without loss of generality, we may assume that  $a_1, a_2$  are orthonormal. In particular,  $a_1 \times a_2 \neq 0$ . Define

$$\varphi(x) = (a_1 \times a_2) \times x \quad \text{and} \quad \psi(x) = a_2(a_1, x) - a_1(a_2, x)$$

Then  $\varphi \neq 0$  is skew, and  $\psi$  is skew since  $(x, \psi x) = 0$  for all  $x$ . The kernel of  $\varphi$  is the line determined by  $a_1 \times a_2$ , which is killed by  $\psi$ . By the previous problem,  $\psi = \lambda \varphi$  for some  $\lambda \in \mathbb{R}$ . Substituting  $x = a_1 + a_2$  into this equation, it follows that  $a_2 - a_1 = \lambda(a_2 - a_1)$ , so  $\lambda = 1$  and  $\psi = \varphi$ . □

**Exercise (5).** A linear transformation  $\varphi$  satisfies  $\tilde{\varphi} = \lambda\varphi$  for  $\lambda \in \mathbb{R}$  if and only if  $\varphi$  is self-adjoint or skew.

*Proof.* If  $\varphi \neq 0$  and  $\tilde{\varphi} = \lambda\varphi$  for some  $\lambda \in \mathbb{R}$ , then there exist vectors  $x, y$  such that  $(\varphi x, y) \neq 0$  and

$$(\varphi x, y) = \lambda(x, \varphi y) = \lambda^2(\varphi x, y)$$

so  $\lambda = \pm 1$  and  $\varphi$  is self-adjoint or skew, respectively. The rest is obvious.  $\square$

**Exercise (6).** If  $\Phi$  is a skew symmetric bilinear function in an oriented Euclidean 3-space, then there is a unique vector  $a$  such that

$$\Phi(x, y) = (a, x \times y)$$

*Proof.* By problem 2, the skew transformation  $\varphi$  dual to  $\Phi$  can be written in the form  $\varphi(x) = a \times x$  for some vector  $a$ . Then

$$\Phi(x, y) = (\varphi x, y) = (a \times x, y) = (a, x \times y)$$

Uniqueness is obvious.  $\square$

*Remark.* This result shows that any function measuring 2-dimensional area in planes in 3-space actually measures 3-dimensional volume relative to some fixed vector in the space.

## § 5

*Remark.* Geometrically, a rotation (and more generally an isometry) preserves length and angle. A proper rotation additionally preserves orientation, whereas an improper rotation reverses it.

**Exercise (2).** A linear transformation  $\varphi$  is regular and preserves orthogonality (that is,  $(\varphi x, \varphi y) = 0$  whenever  $(x, y) = 0$ ) if and only if  $\varphi = \lambda\tau$  where  $\lambda \neq 0$  and  $\tau$  is a rotation.

*Proof.* For the forward direction, let  $e_1, \dots, e_n$  be an orthonormal basis. Then  $\varphi e_1, \dots, \varphi e_n$  is an orthogonal basis. Also

$$|\varphi e_i|^2 - |\varphi e_j|^2 = (\varphi e_i - \varphi e_j, \varphi e_i + \varphi e_j) = (\varphi(e_i - e_j), \varphi(e_i + e_j)) = 0$$

since

$$(e_i - e_j, e_i + e_j) = 1 - 1 = 0$$

for all  $i, j$ . Let  $\lambda = |\varphi e_i| > 0$ . Then clearly  $\tau = \lambda^{-1}\varphi$  is a rotation.

The reverse direction is trivial.  $\square$

*Remark.* This proof is motivated by the geometrical fact that a rectangle is a square if and only if its diagonals are orthogonal.

**Exercise (5).** If  $\varphi : E \rightarrow E$  is a mapping such that  $\varphi(0) = 0$  and

$$|\varphi x - \varphi y| = |x - y|$$

for all  $x, y \in E$ , then  $\varphi$  is linear.

*Proof.* Substituting  $y = 0$ , we have  $|\varphi x| = |x|$  for all  $x$ , so

$$(\varphi x - \varphi y, \varphi x - \varphi y) = |x|^2 - 2(\varphi x, \varphi y) + |y|^2$$

On the other hand,

$$(\varphi x - \varphi y, \varphi x - \varphi y) = |\varphi x - \varphi y|^2 = |x - y|^2 = |x|^2 - 2(x, y) + |y|^2$$

Therefore

$$(\varphi x, \varphi y) = (x, y)$$

It now follows that

$$|\varphi(x + y) - \varphi x - \varphi y|^2 = (\varphi(x + y) - \varphi x - \varphi y, \varphi(x + y) - \varphi x - \varphi y) = 0$$

so  $\varphi(x + y) = \varphi x + \varphi y$ . Similarly  $\varphi(\lambda x) = \lambda \varphi x$ . □

## § 6

*Remark.* In subsection 8.21,  $j$  is called the canonical “complex structure” on  $E$  because it induces a complex vector space structure on the underlying set of  $E$  in which scalar multiplication is defined by

$$(\alpha + \beta i) \cdot x = \alpha x + \beta jx \quad \alpha, \beta \in \mathbb{R}$$

*Remark.* In subsection 8.21, to derive (8.40) from (8.39) and (8.41), first observe that  $j\varphi = \varphi j$ . Indeed, for  $z \neq 0$ ,  $(jz, z) = 0$ , so  $(\varphi jz, \varphi z) = 0$ . On the other hand,  $(j\varphi z, \varphi z) = 0$ . Since  $E$  is a plane, it follows that  $j\varphi z = \lambda \varphi jz$  for some  $\lambda \in \mathbb{R}$ . But

$$\begin{aligned} \lambda |z|^2 &= \lambda |\varphi jz|^2 \\ &= \lambda (\varphi jz, \varphi jz) \\ &= (j\varphi z, \varphi jz) \\ &= \Delta(\varphi z, \varphi jz) \\ &= \Delta(z, jz) \\ &= (jz, jz) \\ &= (z, z) = |z|^2 \end{aligned}$$

so  $\lambda = 1$  and  $j\varphi z = \varphi jz$ .

From this and (8.39), it follows that

$$\varphi^{-1}x = x \cdot \cos(-\Theta) + jx \cdot \sin(-\Theta) = x \cdot \cos \Theta - jx \cdot \sin \Theta$$

where  $x$  and  $\Theta$  are as in (8.39). In (8.41),

$$\Delta(x, \varphi y) = (jx, \varphi y) = (\varphi^{-1}jx, y) = (j\varphi^{-1}x, y) = \Delta(\varphi^{-1}x, y)$$

so

$$\Delta(\varphi x, y) + \Delta(x, \varphi y) = \Delta(\varphi x + \varphi^{-1}x, y) = 2 \cos \Theta \cdot \Delta(x, y)$$

and it follows that  $\cos \Theta = \frac{1}{2} \operatorname{tr} \varphi$ . Similar reasoning shows that  $\sin \Theta = -\frac{1}{2} \operatorname{tr}(j\varphi)$ , contrary to what the book says.

*Remark.* In subsection 8.21, to make sense of the “definition” in (8.43), fix  $x \neq 0$  and let  $0 \leq \Theta \leq \pi$  be the angle between  $x$  and  $\varphi x$ . Fix an orientation in  $E$  so that  $\Theta$  is the *oriented* angle between  $x$  and  $\varphi x$ . Then (8.40) holds, so  $\cos \Theta = \frac{1}{2} \operatorname{tr} \varphi$ . Clearly  $\Theta$  is independent of  $x$  and depends only on  $\varphi$ , so can be written as  $\Theta(\varphi)$  and satisfies (8.43).

*Remark.* In subsection 8.22, why would we expect the rotation vector  $u$  in (8.45) to lie on the rotation axis? Since  $\psi$  is skew,  $\Psi(x, y) = (\psi x, y)$  measures oriented area in planes in  $E$ , and since  $\psi$  kills off  $E_1$ ,  $\Psi(x, y) = 0$  if  $x \in E_1$  or  $y \in E_1$ . But we know from previous results that  $\Psi$  actually measures oriented *volume* relative to  $u$ , so it follows that  $u$  must lie in  $E_1$ .

Note that  $\psi$  stabilizes  $F = E_1^\perp$ , and if  $\psi_1$  denotes the restriction, then

$$\psi_1 = \frac{1}{2}(\varphi_1 - \widetilde{\varphi_1}) = \frac{1}{2}(\varphi_1 - \varphi_1^{-1})$$

where  $\varphi_1$  denotes the restriction of  $\varphi$ . If  $F$  is oriented by  $E$  and by  $u \neq 0$ , and  $j$  is the induced canonical complex structure on  $F$ , then

$$\psi_1 = j \cdot \sin \Theta$$

where  $0 < \Theta < \pi$  is the oriented angle of rotation. If  $x \in F$  is any unit vector, then

$$|u| = |u \times x| = |\psi_1 x| = |jx| \sin \Theta = \sin \Theta$$

as expected.



*Remark.* In subsection 8.24, for the proof of the first part of Proposition I, let  $q_i = p_i - \lambda_i e$  where  $\lambda_i = (p_i, e)$  for  $i = 1, 2$ . Since  $q_1$  and  $q_2$  each generate the same axis of rotation, either  $q_1 = 0 = q_2$  in which case  $p_1 = \pm e$  and  $p_2 = \pm e$  and the result holds, or else  $q_1 \neq 0$  and  $q_2 = \alpha q_1$  for some  $\alpha \neq 0$ . If  $\alpha > 0$ , then  $q_1$  and  $q_2$  induce the same orientation in their orthogonal plane (in  $E_1$ ), so the oriented angle  $\Theta$  of rotation in that plane is the same. It follows that

$$\lambda_1 = \cos \frac{\Theta}{2} = \lambda_2 \quad \text{and} \quad |q_1| = \sin \frac{\Theta}{2} = |q_2|$$

The second equation implies that  $\alpha = 1$ , so  $q_1 = q_2$  and  $p_1 = p_2$ . On the other hand if  $\alpha < 0$ , then  $-p_2$  is a unit quaternion inducing the same rotation as  $p_1$  and its pure part  $-q_2$  satisfies  $-q_2 = (-\alpha)q_1$  with  $-\alpha > 0$ , so by the previous case  $p_1 = -p_2$ .

**Exercise (16).** If  $p \neq \pm e$  is a unit quaternion, then the rotation vector induced by  $p$  is

$$u = 2\lambda(p - \lambda e) \quad \lambda = (p, e)$$

*Proof.* Let  $q = p - \lambda e$ . By assumption,  $q \neq 0$  and  $u = \alpha q$  since  $q$  and  $u$  both lie on the axis of rotation. If  $\alpha > 0$ , then  $q$  and  $u$  both induce the same orientation on the orthogonal plane and

$$\alpha|q| = |u| = \sin \Theta = 2\lambda|q|$$

where  $\Theta$  is the oriented angle of rotation. Therefore  $\alpha = 2\lambda$  and the result holds. If  $\alpha < 0$ , then  $-p$  induces the same rotation vector and satisfies the hypotheses of the previous case since  $u = (-\alpha)(-q)$  with  $-\alpha > 0$ , so  $-\alpha = 2(-\lambda)$  and  $\alpha = 2\lambda$  and the result holds. If  $\alpha = 0$ , then  $u = 0$  and  $\Theta = \pi$ , so  $\lambda = \cos(\pi/2) = 0$  and the result holds.  $\square$

**Exercise (17).** Let  $p \neq \pm e$  be a unit quaternion. If  $F$  denotes the plane generated by  $e$  and  $p$ , then the rotations  $\varphi x = px$  and  $\psi x = xp$  agree on  $F$  and stabilize  $F$  and  $F^\perp$ .

*Proof.* The rotations agree on  $e$  and  $p$ , hence on  $F$ . By definition of quaternion multiplication,  $p^2 \in F$ , so they also stabilize  $F$  and  $F^\perp$ .  $\square$

## Chapter XI

*Remark.* In parts of this chapter, it is implicitly assumed that unitary spaces have dimension  $n \geq 1$ .

*Remark.* A map  $\varphi : E \rightarrow F$  of complex vector spaces  $E, F$  is *conjugate-linear* if

$$\varphi(\lambda x + \mu y) = \bar{\lambda}\varphi x + \bar{\mu}\varphi y$$

for all  $x, y \in E$  and  $\lambda, \mu \in \mathbb{C}$ , where  $\bar{\lambda}, \bar{\mu}$  denote the complex conjugates of  $\lambda, \mu$  respectively.

## § 1

*Remark.* A sesquilinear function  $\Phi(x, y)$  is linear in  $x$  and conjugate-linear in  $y$ .

## § 2

*Remark.* The identity map  $\kappa : F \rightarrow \bar{F}$  is conjugate-linear. A map  $\varphi : E \rightarrow F$  is conjugate-linear if and only if  $\kappa\varphi : E \rightarrow \bar{F}$  is linear. This yields a natural bijective correspondence between conjugate-linear maps  $E \rightarrow F$  and linear maps  $E \rightarrow \bar{F}$ .

*Remark.* If  $z \mapsto \bar{z}$  is a conjugation in  $E$ , then for  $\lambda = \alpha + i\beta$  with  $\alpha, \beta \in \mathbb{R}$ ,

$$\overline{\lambda z} = \overline{\alpha z + i\beta z} = \alpha\bar{z} - i\beta\bar{z} = \bar{\lambda}\bar{z}$$

Therefore a conjugation is just a conjugate-linear involution. By the previous remark, a conjugation can also be viewed as a linear map  $E \rightarrow \bar{E}$ .

## § 3

*Remark.* If  $E$  is a unitary space, then there exists a conjugation in  $E$ . In fact, if  $z_1, \dots, z_n$  is a basis of  $E$  and  $F$  is the real span of  $z_1, \dots, z_n$ , then  $F$  is a real form of  $E$  and the map  $z \mapsto \bar{z}$  defined by  $x + iy \mapsto x - iy$  for  $x, y \in F$  is a conjugation in the *vector space*  $E$ . It is also a conjugation in the *unitary space*  $E$  since for  $z_1 = x_1 + iy_1$  and  $z_2 = x_2 + iy_2$  with  $x_j, y_j \in F$ ,

$$(z_1, z_2) = (x_1, x_2) + (y_1, y_2) + i[(y_1, x_2) - (x_1, y_2)]$$

and

$$(\bar{z}_1, \bar{z}_2) = (x_1, x_2) + (y_1, y_2) - i[(y_1, x_2) - (x_1, y_2)]$$

so  $(\bar{z}_1, \bar{z}_2) = \overline{(z_1, z_2)}$  as required.<sup>15</sup>

---

<sup>15</sup>See problem 2(i) in § 2.

*Remark.* If  $E, F$  are unitary spaces and  $\bar{E}, \bar{F}$  the corresponding conjugate spaces, then  $E$  is dual to  $\bar{E}$  under the scalar product

$$\langle x, x^* \rangle = (x, \kappa_E^{-1} x^*)$$

where  $\kappa_E : E \rightarrow \bar{E}$  is the identity map; similarly  $F$  is dual to  $\bar{F}$  under the scalar product

$$\langle y, y^* \rangle = (y, \kappa_F^{-1} y^*)$$

where  $\kappa_F : F \rightarrow \bar{F}$  is the identity map.

If  $\varphi : E \rightarrow F$  is a linear map, then the dual map  $\varphi^* : \bar{E} \leftarrow \bar{F}$  satisfies

$$(\varphi x, \kappa_F^{-1} y^*) = \langle \varphi x, y^* \rangle = \langle x, \varphi^* y^* \rangle = (x, \kappa_E^{-1} \varphi^* y^*)$$

Taking  $y^* = \kappa_F y$  yields

$$(\varphi x, y) = (x, \kappa_E^{-1} \varphi^* \kappa_F y)$$

Therefore  $\tilde{\varphi} = \kappa_E^{-1} \varphi^* \kappa_F$ :

$$\begin{array}{ccc} E & \xleftarrow{\tilde{\varphi}} & F \\ \kappa_E \downarrow & & \downarrow \kappa_F \\ \bar{E} & \xleftarrow{\varphi^*} & \bar{F} \end{array}$$

This construction of the adjoint avoids using conjugations in  $E, F$  by using the conjugate spaces  $\bar{E}, \bar{F}$  instead.

If  $E = F$ ,  $\Delta \neq 0$  is a determinant function in  $E$ , and  $\Delta^*$  is the corresponding determinant function in  $\bar{E}$  (11.6), then

$$\begin{aligned} \det \tilde{\varphi} \cdot \Delta(x_1, \dots, x_n) &= \Delta(\tilde{\varphi} x_1, \dots, \tilde{\varphi} x_n) \\ &= \Delta(\kappa_E^{-1} \varphi^* \kappa_F x_1, \dots, \kappa_E^{-1} \varphi^* \kappa_F x_n) \\ &= \overline{\Delta^*(\varphi^* \kappa_F x_1, \dots, \varphi^* \kappa_F x_n)} \\ &= \overline{\det \varphi^* \cdot \Delta^*(\kappa_F x_1, \dots, \kappa_F x_n)} \\ &= \overline{\det \varphi} \cdot \Delta(x_1, \dots, x_n) \end{aligned}$$

where  $\kappa = \kappa_E$ , so  $\det \tilde{\varphi} = \overline{\det \varphi}$ . This derivation of (11.21) similarly avoids using conjugations in  $E$ .

*Remark.* The mapping  $\varphi \mapsto \tilde{\varphi}$  is conjugate-linear.

*Remark.* If  $z \mapsto \bar{z}$  is a conjugation in  $E$  and  $\Delta$  is a determinant function in  $E$ , then

$$\bar{\Delta}(z_1, \dots, z_n) = \overline{\Delta(\bar{z}_1, \dots, \bar{z}_n)}$$

is also a determinant function in  $E$ , called the *conjugate determinant function*. The proof of (11.21) in the book uses one of these.

**Exercise (1).** If  $E$  is a unitary space and  $\varphi : E \rightarrow E$  is linear, then

$$\Phi(x, y) = (\varphi x, y)$$

is sesquilinear. Conversely, every sesquilinear function in  $E$  can be uniquely represented in this way. The adjoint  $\tilde{\varphi}$  represents the adjoint  $\tilde{\Phi}$ .

*Proof.* The forward direction is trivial. If  $\Phi$  is sesquilinear, then for any fixed vector  $x$  the function

$$y \mapsto \overline{\Phi(x, y)}$$

is linear, so by the Riesz theorem (11.5) there is a unique vector  $\varphi x$  such that

$$\overline{\Phi(x, y)} = (y, \varphi x) = \overline{(\varphi x, y)}$$

and therefore

$$\Phi(x, y) = (\varphi x, y)$$

Clearly  $\varphi$  is linear and is uniquely determined by  $\Phi$ . Also

$$(\tilde{\varphi} x, y) = \overline{(y, \tilde{\varphi} x)} = \overline{(\varphi y, x)} = \overline{\Phi(y, x)} = \tilde{\Phi}(x, y) \quad \square$$

*Remark.* This result shows that the linear transformation  $\varphi$  is naturally left-dual to the sesquilinear function  $\Phi$ . Observe that  $\Phi$  is Hermitian if and only if  $\varphi$  is Hermitian (self-adjoint) and  $\Phi$  is skew-Hermitian if and only if  $\varphi$  is skew-Hermitian.

## Chapter XII

### § 1

*Remark.* The pair  $(\Gamma[t], t)$  is a universal (initial) object in the category of pointed algebras  $(A, a)$ , where  $A$  is an associative unital  $\Gamma$ -algebra with unit  $e \in A$  and

distinguished element  $a \in A$ . In this category, an arrow  $\varphi : (A, a) \rightarrow (B, b)$  is a unital  $\Gamma$ -algebra homomorphism  $\varphi : A \rightarrow B$  with  $\varphi a = b$ . This result follows from the existence and uniqueness of the evaluation homomorphism  $\Gamma[t] \rightarrow A$ .

Identifying the scalars  $\alpha \in \Gamma$  with the elements  $\alpha \cdot e \in A$ , this result shows that  $\Gamma[t]$  is universal as an algebra obtained by adjoining a single element to  $\Gamma$ , in the sense that every such algebra can be obtained from  $\Gamma[t]$  in a unique way through evaluation.

**Exercise (2).** Define the linear mapping  $\int : \Gamma[t] \rightarrow \Gamma[t]$  by

$$\int t^p = \frac{t^{p+1}}{p+1} \quad (p \geq 0)$$

Then  $\int$  is homogeneous (of degree 1) with  $d \circ \int = \iota$ , and is unique with these properties.

*Proof.* If  $\int$  satisfies the properties, then we must have

$$\int t^p = \frac{t^{p+1}}{p+1} + \alpha_p \quad (\alpha_p \in \Gamma)$$

since  $d \circ \int = \iota$ . But  $\alpha_p = 0$  for all  $p$  since  $\int$  is homogeneous. Finally,  $\int$  is uniquely determined on  $\Gamma[t]$  since  $\int$  is linear and the  $t^p$  ( $p \geq 0$ ) form a basis of  $\Gamma[t]$ .  $\square$

**Exercise (3).** If  $\int : \Gamma[t] \rightarrow \Gamma[t]$  is the integration operator, then  $\int \circ d = \iota - \varrho$ , where  $\varrho : \Gamma[t] \rightarrow \Gamma$  is the scalar projection  $\varrho f = f(0)$ . It follows that

$$\int f g' = f g - \varrho(f g) - \int g f'$$

*Proof.* The first part is easily verified on the basis  $t^p$  ( $p \geq 0$ ). The second part then follows since

$$f g - \varrho(f g) = \int (f g)' = \int (f' g + f g') = \int f' g + \int f g' \quad \square$$

### § 3

*Remark.* If  $A = 0$ , then the minimum polynomial of  $a = 0$  is  $\mu = 1$ .

*Remark.* In the proof of Proposition I, to see why the elements  $1, \bar{t}, \dots, \bar{t}^{r-1}$  form a basis of  $\Gamma[t]/I_\mu$ , first observe that every element of  $\Gamma[t]/I_\mu$  is of the form  $f(\bar{t})$  for some  $f \in \Gamma[t]$ . Write  $f = g\mu + h$  where  $h = 0$  or  $\deg h < r$ . Then

$$f(\bar{t}) = g(\bar{t})\mu(\bar{t}) + h(\bar{t}) = h(\bar{t})$$

since  $\mu(\bar{t}) = 0$ , and clearly  $h(\bar{t})$  is in the span of  $1, \bar{t}, \dots, \bar{t}^{r-1}$ .

On the other hand, suppose

$$\alpha_0 + \alpha_1 \bar{t} + \dots + \alpha_{r-1} \bar{t}^{r-1} = 0$$

Define  $f = \alpha_0 + \dots + \alpha_{r-1} t^{r-1}$ . Then  $f = 0$  or  $\deg f < r$ . But  $f \in \ker \pi = I_\mu$ , so  $\mu \mid f$ , which forces  $f = 0$  and  $\alpha_0 = \dots = \alpha_{r-1} = 0$ . It follows that the elements  $1, \bar{t}, \dots, \bar{t}^{r-1}$  are linearly independent.

## Chapter XIII

*Remark.* In parts of this chapter, it is implicitly assumed that the vector space  $E$  has dimension  $n \geq 1$ .

### § 1

In the following remarks,  $f \in \Gamma[t]$ .

*Remark.* Why is it natural to apply polynomials to a linear transformation? If we think of a linear transformation  $\varphi : E \rightarrow E$  as being adjoined to the scalar transformations in  $A(E; E)$ , then we know that the structure of the resulting subalgebra is obtained from that of  $\Gamma[t]$  in a unique way through evaluation  $f \mapsto f(\varphi)$ , by the universal property of  $\Gamma[t]$ .<sup>16</sup> This fact makes it natural to apply polynomials to  $\varphi$  in order to study the structure of  $\varphi$ .

*Remark.* The results of this section are summarized in the following diagram, which shows that  $K$  is a homomorphism from the lattice of polynomials in  $\Gamma[t]$  under divisibility to the lattice of subspaces of  $E$  under inclusion:<sup>17</sup>

<sup>16</sup>See the remark in chapter XII, § 1 above.

<sup>17</sup>Recall from chapter XII that  $f \vee g$  denotes the greatest common divisor of  $f$  and  $g$ , while  $f \wedge g$  denotes the least common multiple of  $f$  and  $g$ —notation inspired by the lattice of *ideals*. See also problem 11.



*Remark.* If  $\varphi : E \rightarrow E$  is linear and  $E_1$  is stable under  $\varphi$ , then  $E_1$  is also stable under  $f(\varphi)$  and

$$f(\varphi)_{E_1} = f(\varphi_{E_1})$$

Dually,

$$\overline{f(\varphi)} = f(\overline{\varphi})$$

*Proof.* Stabilization of  $E_1$  is preserved by addition, scalar multiplication, and composition, so is preserved by  $f$ . It follows that  $f(\varphi)_{E_1}$  and  $f(\varphi_{E_1})$  are both transformations of  $E_1$  with the same action, so are equal. Finally,

$$\overline{f(\varphi)}(\overline{x}) = \overline{f(\varphi)(x)} = f(\overline{\varphi})(\overline{x}) \quad \square$$

*Remark.* It follows that if  $K_1(f) = \ker f(\varphi_{E_1})$ , then

$$K_1(f) = K(f) \cap E_1$$

Dually, if  $\overline{K}(f) = \ker f(\overline{\varphi})$ , then

$$\overline{K}(f) \supseteq \overline{K(f)}$$

The reverse inclusion does not hold in general. For a counterexample, consider  $f = t$  with  $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  defined by  $\varphi(x, y) = (y, 0)$  and  $E_1 = \mathbb{R} \times 0$ .

*Remark.* If  $\varphi_1 : E_1 \rightarrow E_1$  and  $\varphi_2 : E_2 \rightarrow E_2$  are linear, then

$$f(\varphi_1 \oplus \varphi_2) = f(\varphi_1) \oplus f(\varphi_2)$$

*Proof.* Write  $\varphi = \varphi_1 \oplus \varphi_2$ , so  $\varphi_{E_1} = \varphi_1$  and  $\varphi_{E_2} = \varphi_2$ . Then

$$f(\varphi) = f(\varphi)_{E_1} \oplus f(\varphi)_{E_2} = f(\varphi_1) \oplus f(\varphi_2) \quad \square$$

*Remark.* It follows that

$$r(f(\varphi)) = r(f(\varphi_1)) + r(f(\varphi_2))$$

where  $r(\psi)$  denotes the rank of  $\psi$ .

*Remark.* If  $E, E^*$  are dual spaces and  $\varphi : E \rightarrow E$  and  $\varphi^* : E^* \leftarrow E^*$  are dual linear transformations, then

$$f(\varphi)^* = f(\varphi^*)$$

*Proof.* By the results of subsection 2.25.  $\square$

*Remark.* It follows that  $f(\varphi) = 0$  if and only if  $f(\varphi^*) = 0$ , so  $\mu_{\varphi^*} = \mu_{\varphi}$ .

*Remark.* In particular, if  $E$  is an inner product space or a unitary space, then

$$\widetilde{f(\varphi)} = \overline{f(\tilde{\varphi})}$$

where  $\overline{f}$  is obtained from  $f$  by conjugating the coefficients.<sup>18</sup>

*Proof.* Let  $\overline{E}$  be the conjugate space and  $\kappa : E \rightarrow \overline{E}$  the conjugate-linear identity map. Then

$$\widetilde{f(\varphi)} = \kappa^{-1} f(\varphi)^* \kappa = \kappa^{-1} f(\varphi^*) \kappa = \overline{f}(\kappa^{-1} \varphi^* \kappa) = \overline{f}(\tilde{\varphi}) \quad \square$$

*Remark.* It follows that  $f(\varphi) = 0$  if and only if  $\overline{f}(\tilde{\varphi}) = 0$ , so  $\mu_{\tilde{\varphi}} = \overline{\mu_{\varphi}}$ .

*Remark.* If  $\varphi : E \rightarrow E$  and  $\psi : F \rightarrow F$  are linear with

$$\psi = \alpha \circ \varphi \circ \alpha^{-1}$$

where  $\alpha : E \cong F$ , then

$$f(\psi) = \alpha \circ f(\varphi) \circ \alpha^{-1}$$

*Proof.* Conjugation by  $\alpha$  is an algebra isomorphism  $A(E; E) \cong A(F; F)$ .  $\square$

---

<sup>18</sup>In the case of an inner product space,  $\overline{f} = f$ .



*Remark.* It follows that  $f(\psi) = 0$  if and only if  $f(\varphi) = 0$ , so  $\mu_\psi = \mu_\varphi$ .

*Remark.* In particular if  $E = E_1 \oplus E_2$  and  $\varphi : E \rightarrow E$  is given by  $\varphi = \varphi_1 \oplus \varphi_2$  where  $\varphi_1 : E_1 \rightarrow E_1$  and  $\varphi_2 : E_2 \rightarrow E_2$ , then  $\mu_{\varphi_1} = \mu_{\bar{\varphi}}$  where  $\bar{\varphi} : E/E_2 \rightarrow E/E_2$  is the induced transformation. Dually,  $\mu_{\varphi_2} = \mu_{\bar{\varphi}}$  where  $\bar{\varphi} : E/E_1 \rightarrow E/E_1$  is induced.<sup>19</sup>

**Exercise (2).** Consider the linear transformations  $\varphi, \psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  defined by

$$\varphi(x, y) = (y, 0) \quad \text{and} \quad \psi(x, y) = (0, y)$$

Then  $\varphi\psi = \varphi$  while  $\psi\varphi = 0$ , so

$$\mu_{\varphi\psi} = t^2 \neq t = \mu_{\psi\varphi}$$

**Exercise (3).** If  $\varphi_1 : E_1 \rightarrow E_1$  and  $\varphi_2 : E_2 \rightarrow E_2$  are linear, then

$$\mu_{\varphi_1 \oplus \varphi_2} = \mu_{\varphi_1} \wedge \mu_{\varphi_2}$$

*Proof.* Write  $\varphi = \varphi_1 \oplus \varphi_2$ . By a remark above,  $f(\varphi) = f(\varphi_1) \oplus f(\varphi_2)$ , so  $f(\varphi) = 0$  if and only if  $f(\varphi_1) = 0$  and  $f(\varphi_2) = 0$ . It follows that  $\mu_\varphi$  is the least common multiple of  $\mu_{\varphi_1}$  and  $\mu_{\varphi_2}$ .  $\square$

**Exercise (5).** If  $\varphi : E \rightarrow E$  stabilizes  $E_1$ , and  $\varphi_1 : E_1 \rightarrow E_1$  and  $\bar{\varphi} : E/E_1 \rightarrow E/E_1$  are the induced transformations, then

$$\bar{\mu} \wedge \mu_1 \mid \mu \mid \bar{\mu}\mu_1$$

where  $\mu = \mu_\varphi$ ,  $\mu_1 = \mu_{\varphi_1}$ , and  $\bar{\mu} = \mu_{\bar{\varphi}}$ .

*Proof.* By a remark above,  $\mu(\varphi_1) = 0$  and  $\mu(\bar{\varphi}) = 0$ , so  $\mu_1 \mid \mu$  and  $\bar{\mu} \mid \mu$ . On the other hand,

$$\bar{\mu}\mu_1(\varphi)(x) = \mu_1(\varphi)(\bar{\mu}(\varphi)(x))$$

But  $\bar{\mu}(\varphi)(x)$  is in  $E_1$  since

$$\overline{\bar{\mu}(\varphi)(x)} = \overline{\bar{\mu}(\varphi)(\bar{x})} = \bar{\mu}(\bar{\varphi})(\bar{x}) = 0$$

so

$$\mu_1(\varphi)(\bar{\mu}(\varphi)(x)) = \mu_1(\varphi_1)(\bar{\mu}(\varphi)(x)) = 0$$

It follows that  $\bar{\mu}\mu_1(\varphi) = 0$ , so  $\mu \mid \bar{\mu}\mu_1$ .  $\square$

<sup>19</sup>See the conjugacy diagrams in chapter II, § 4 above.

**Exercise (6).** The minimum polynomial  $\mu$  of  $\varphi : E \rightarrow E$  can be constructed in the following way: select  $x_1 \in E$  and let  $k_1$  be least such that the vectors

$$x_1, \varphi x_1, \dots, \varphi^{k_1} x_1 \quad (1)$$

are linearly dependent, with

$$\lambda_0 x_1 + \lambda_1 \varphi x_1 + \dots + \lambda_{k_1} \varphi^{k_1} x_1 = 0$$

where the  $\lambda_i$  are not all zero, so in particular  $\lambda_{k_1} \neq 0$ . Define the polynomial

$$f_1 = \lambda_0 + \lambda_1 t + \dots + \lambda_{k_1} t^{k_1}$$

If the vectors (1) do not span  $E$ , select  $x_2 \in E$  not in their span and construct the corresponding polynomial  $f_2$ . Repeat this procedure until  $E$  is exhausted. Then  $\mu$  is the (monic) least common multiple of the  $f_i$ .

*Proof.* Let  $E_i$  be the span of the vectors

$$x_i, \varphi x_i, \dots, \varphi^{k_i-1} x_i \quad (2)$$

By construction, (2) is a basis of  $E_i$  with  $\varphi^{k_i} x_i \in E_i$ , and we have the direct sum  $E = \sum_i E_i$ . The minimum polynomial of the restriction  $\varphi_i : E_i \rightarrow E_i$  is just  $\lambda_{k_i}^{-1} f_i$ , so the result follows from problem 3.  $\square$

*Remark.* The subspaces  $E_i$  are cyclic, so this procedure just produces a cyclic decomposition of  $E$ , from which the minimum polynomial is easily computed.<sup>20</sup>

**Exercise (8).** If  $E$  is an inner product space and  $\varphi : E \rightarrow E$  is a rotation, then the coefficients  $\alpha_i$  of the minimum polynomial  $\mu$  of  $\varphi$  satisfy

$$\alpha_i = \varepsilon \alpha_{k-i} \quad k = \deg \mu$$

where  $\varepsilon = \pm 1$ .

*Proof.* Since  $\tilde{\varphi} = \varphi^{-1}$ , we have  $\mu_{\varphi^{-1}} = \mu$  by a remark above, so

$$\alpha_0 + \alpha_1 \varphi^{-1} + \dots + \alpha_k \varphi^{-k} = \mu(\varphi^{-1}) = 0 \quad (1)$$

Multiplying both sides of (1) by  $\varphi^k$  yields

$$\alpha_0 \varphi^k + \alpha_1 \varphi^{k-1} + \dots + \alpha_k = \mathfrak{u}(\varphi) = 0 \quad (2)$$

where  $\mathfrak{u} = \sum_i \alpha_{k-i} t^i$ . It follows from (2) that  $\mu \mid \mathfrak{u}$ , and in fact  $\alpha_0 \mu = \mathfrak{u}$ . Since  $\alpha_0^2 = \alpha_k = 1$ , the result holds with  $\varepsilon = \alpha_0$ .  $\square$

<sup>20</sup>See Proposition IV and its corollary in § 3.

**Exercise (9).** The minimum polynomial  $\mu$  of a self-adjoint transformation of a unitary space has real coefficients.

*Proof.* By a remark above,  $\bar{\mu} = \mu$ . □

**Exercise (12).** If  $\varphi : E \rightarrow E$  is regular, then  $\varphi^{-1} = f(\varphi)$  for some  $f \in \Gamma[t]$ .

*Proof.* Let  $\mu$  be the minimum polynomial of  $\varphi$  and  $\alpha = \mu(0)$ . Since  $\varphi$  is regular, 0 is not an eigenvalue of  $\varphi$ , so  $\alpha \neq 0$ . Write

$$\alpha^{-1}\mu = 1 - tf \quad f \in \Gamma[t] \quad (1)$$

Evaluating both sides of (1) at  $\varphi$  yields  $\varphi f(\varphi) = \iota$  and therefore  $\varphi^{-1} = f(\varphi)$ . □

## § 2

*Remark.* If the minimum polynomial of  $\varphi : E \rightarrow E$  has the prime decomposition

$$\mu = f_1^{k_1} \cdots f_r^{k_r}$$

and  $E$  has the generalized eigenspace decomposition

$$E = E_1 \oplus \cdots \oplus E_r$$

define

$$g_i = f_1^{k_1} \cdots \widehat{f_i^{k_i}} \cdots f_r^{k_r}$$

and

$$F_i = E_1 \oplus \cdots \oplus \widehat{E_i} \oplus \cdots \oplus E_r$$

Then  $f_i^{k_i}(\varphi)$  kills  $E_i$  and induces an isomorphism on  $F_i$ , while  $g_i(\varphi)$  kills  $F_i$  and induces an isomorphism on  $E_i$ .

More generally if  $K \subseteq \{1, \dots, r\}$ , define

$$f_K = \prod_{i \in K} f_i^{k_i}$$

and

$$E_K = \sum_{i \in K} E_i \quad E_{\widehat{K}} = \sum_{i \notin K} E_i$$

Then  $f_K(\varphi)$  kills  $E_K$  and induces an isomorphism on  $E_{\widehat{K}}$ .

*Remark.* In Proposition I, if  $\pi_i$  and  $\pi_i^*$  are the projection operators associated with the generalized eigenspace decompositions in  $E$  and  $E^*$ , then

$$(\pi_i)^* = \pi_i^*$$

*Proof.* From (13.24), it follows that

$$\langle \pi_i^* x^*, x \rangle = \langle \pi_i^* x^*, \pi_i x \rangle = \langle x^*, \pi_i x \rangle \quad \square$$

*Remark.* If  $g_i \in \Gamma[t]$  with  $g_i(\varphi) = \pi_i$  (subsection 13.5), then

$$g_i(\varphi^*) = g_i(\varphi)^* = (\pi_i)^* = \pi_i^*$$

Alternatively, the fact that  $g_i(\varphi^*) = \pi_i^*$  follows from the fact that  $g_i$  depends only on the common minimum polynomial of  $\varphi^*$  and  $\varphi$ , and the argument above can be run in reverse to obtain (13.24).<sup>21</sup>

**Exercise (1).** The minimum polynomial  $\mu$  of  $\varphi : E \rightarrow E$  is completely reducible if and only if every nonzero subspace of  $E$  stable under  $\varphi$  contains an eigenvector of  $\varphi$ .

*Proof.* If  $F \neq 0$  is stable, then by (13.22)

$$F = \sum_i F \cap E_i$$

where the  $E_i$  are the generalized eigenspaces of  $\varphi$ . Fix  $i$  with  $F \cap E_i \neq 0$ . If  $\mu$  is completely reducible, then  $E_i = K((t - \lambda)^k)$  for some  $\lambda \in \Gamma$  and  $k \geq 1$ . Choose  $x \in F \cap E_i$  with  $x \neq 0$  and let  $1 \leq l \leq k$  be least with  $(\varphi - \lambda I)^l x = 0$ . Then it follows that  $(\varphi - \lambda I)^{l-1} x \in F$  is an eigenvector.

Conversely, suppose  $f$  is monic irreducible with  $\deg f \geq 1$  and  $f \mid \mu$ . Let  $F = K(f)$ . Then  $F \neq 0$  is stable and so contains an eigenvector by assumption. If  $\lambda$  is the associated eigenvalue, then  $(t - \lambda) \mid \mu_f = f$ , so  $f = t - \lambda$ . It follows that  $\mu$  is completely reducible.  $\square$

*Remark.* This is true in particular if  $\Gamma$  is algebraically closed (for example  $\Gamma = \mathbb{C}$ ).

**Exercise (2).** If the minimum polynomial  $\mu$  of  $\varphi : E \rightarrow E$  is completely reducible, then there is a basis of  $E$  with respect to which the matrix of  $\varphi$  is triangular:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ * & & \lambda_n \end{pmatrix}$$

---

<sup>21</sup>See problems 7 and 8.

*Proof.* By induction on  $\dim E$ .

Since  $E \neq 0$  is stable, there is an eigenvector  $x_1$  with  $\varphi x_1 = \lambda_1 x_1$ . Let  $E_1 = \langle x_1 \rangle$  and consider  $\bar{\varphi} : E/E_1 \rightarrow E/E_1$  with minimum polynomial  $\bar{\mu}$ . Since  $\bar{\mu} \mid \mu$ ,  $\bar{\mu}$  is completely reducible, and by induction there is a basis  $\bar{x}_2, \dots, \bar{x}_n$  of  $E/E_1$  with  $\bar{\varphi}(\bar{x}_k) \in \langle \bar{x}_2, \dots, \bar{x}_k \rangle$  for  $2 \leq k \leq n$ . It follows that  $x_1, \dots, x_n$  is a basis of  $E$  with  $\varphi x_k \in \langle x_1, \dots, x_k \rangle$  for  $1 \leq k \leq n$ .  $\square$

*Remark.* The diagonal entries of this matrix are just the eigenvalues of  $\varphi$ .

**Exercise (9).** Let  $F$  be a subspace of  $E$  stable under  $\varphi : E \rightarrow E$  and consider the induced mappings  $\varphi_F : F \rightarrow F$  and  $\bar{\varphi} : E/F \rightarrow E/F$ . Let  $E = \sum_i E_i$  be the generalized eigenspace decomposition of  $E$  under  $\varphi$ .

(a) The generalized eigenspace decomposition of  $F$  under  $\varphi_F$  is

$$F = \sum_i F \cap E_i$$

(b) The generalized eigenspace decomposition of  $E/F$  under  $\bar{\varphi}$  is

$$E/F = \sum_i E_i/F$$

Moreover,  $E_i/F \cong E_i/(F \cap E_i)$ .

(c) If  $\pi_i$ ,  $\pi_i^F$ , and  $\bar{\pi}_i$  are the projection operators in  $E$ ,  $F$ , and  $E/F$  associated with the decompositions, then the diagram

$$\begin{array}{ccccc} F & \xrightarrow{j} & E & \xrightarrow{\rho} & E/F \\ \pi_i^F \downarrow & & \pi_i \downarrow & & \downarrow \bar{\pi}_i \\ F & \xrightarrow{j} & E & \xrightarrow{\rho} & E/F \end{array}$$

commutes, where  $j$  and  $\rho$  are the canonical injection and projection. Moreover,

$$\pi_i^F = (\pi_i)_F \quad \text{and} \quad \bar{\pi}_i = \overline{\pi_i}$$

so  $\pi_i^F$  and  $\bar{\pi}_i$  are unique making the diagram commute.

(d) If  $g_i \in \Gamma[t]$  with  $g_i(\varphi) = \pi_i$ , then

$$g_i(\varphi_F) = \pi_i^F \quad \text{and} \quad g_i(\overline{\varphi}) = \overline{\pi}_i$$

*Proof.* For (a), we already know from (13.22) that

$$F = \sum_i F \cap E_i \tag{1}$$

Let

$$F = \sum_i F_i \tag{2}$$

be the generalized eigenspace decomposition of  $F$  under  $\varphi_F$ . Since  $\mu_{\varphi_F} \mid \mu_\varphi$ , we may assume that  $F_i = K_F(f_i^{j_i})$  and  $E_i = K(f_i^{k_i})$  for some  $f_i$  with  $j_i \leq k_i$ . Then

$$F_i = K_F(f_i^{j_i}) = F \cap K(f_i^{j_i}) \subseteq F \cap K(f_i^{k_i}) = F \cap E_i \tag{3}$$

It follows from (1), (2), and (3) that  $F_i = F \cap E_i$ , establishing (a).

For (b), it follows from (1) that

$$E/F = \sum_i E_i/F \tag{4}$$

is a direct sum decomposition of  $E/F$ , and it is immediate that  $E_i/F$  is stable under  $\overline{\varphi}$  since  $E_i$  is stable under  $\varphi$ . Let

$$E/F = \sum_i (E/F)_i \tag{5}$$

be the generalized eigenspace decomposition of  $E/F$  under  $\overline{\varphi}$ . Since  $\mu_{\overline{\varphi}} \mid \mu_\varphi$ , we may assume that  $(E/F)_i = \overline{K}(f_i^{m_i})$  with  $m_i \leq k_i$ . Now

$$E_i/F = K(f_i^{k_i})/F \subseteq \overline{K}(f_i^{k_i})$$

so if  $\mu_i$  denotes the minimum polynomial of  $\overline{\varphi}_{E_i/F}$ , then  $\mu_i \mid f_i^{k_i}$ . On the other hand,  $\mu_i \mid \mu_{\overline{\varphi}}$ , so we must have  $\mu_i \mid f_i^{m_i}$  and therefore

$$E_i/F \subseteq \overline{K}(f_i^{m_i}) = (E/F)_i \tag{6}$$

It now follows from (4), (5), and (6) that  $(E/F)_i = E_i/F$ . Finally, the restriction to  $E_i$  of the canonical projection induces the isomorphism  $E_i/F \cong E_i/F$ .

For (c), commutativity of the left square in the diagram follows from (a) and commutativity of the right square follows from (b). The rest of (c) follows from the diagram.

Finally, (d) follows immediately from (c).  $\square$

**Exercise (10).** If the minimum polynomial  $\mu$  of  $\varphi : E \rightarrow E$  is completely reducible, then  $\mu \mid \chi$  where  $\chi$  is the characteristic polynomial of  $\varphi$ .

*Proof.* First consider the case  $\mu = (t - \lambda)^k$ . By problem 2, there is a basis of  $E$  with respect to which the matrix of  $\varphi$  has the form

$$\begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ * & & \lambda \end{pmatrix}$$

It follows immediately that  $(\varphi - \lambda I)^n = 0$  where  $n = \dim E$ , so  $\mu \mid (t - \lambda)^n$  and therefore  $k \leq n$ . Moreover,  $\chi = (-1)^n(t - \lambda)^n$ , so  $\mu \mid \chi$ .

In the general case, take the prime decomposition

$$\mu = (t - \lambda_1)^{k_1} \cdots (t - \lambda_r)^{k_r}$$

and let  $E = \sum_i E_i$  be the corresponding generalized eigenspace decomposition. Let  $\varphi_i : E_i \rightarrow E_i$  be the restriction with minimum polynomial  $\mu_i = (t - \lambda_i)^{k_i}$  and characteristic polynomial  $\chi_i$ . Then  $\mu_i \mid \chi_i$  by the previous case, so

$$\mu = \mu_1 \cdots \mu_r \mid \chi_1 \cdots \chi_r = \chi \quad \square$$

*Remark.* This is just a special case of the Cayley-Hamilton theorem.<sup>22</sup>

### § 3

*Remark.* In Proposition I, the vector  $a$  is intuitively “hard to kill”, and therefore generates a cyclic subspace that is as large as possible.<sup>23</sup>

*Remark.* A vector space  $E$  is cyclic with generator  $a \in E$  under  $\varphi : E \rightarrow E$  if and only if  $E$  is the smallest subspace of  $E$  containing  $a$  and stable under  $\varphi$ .

*Remark.* To prove Proposition III, just recall from the proof of Proposition II that  $\sigma_a$  induces an isomorphism  $\overline{\sigma}_a : \Gamma[t]/I_\mu \cong E$  defined by  $\overline{\sigma}_a(\bar{f}) = \sigma_a(f)$ . Since  $1, \bar{t}, \dots, \bar{t}^{m-1}$  is a basis of  $\Gamma[t]/I_\mu$  and  $\sigma_a(t^i) = \varphi^i(a)$ , it follows that

$$a, \varphi(a), \dots, \varphi^{m-1}(a)$$

is a basis of  $E$ .

---

<sup>22</sup>See Theorem II in § 5.

<sup>23</sup>See Proposition V.

*Remark.* In the proof of the corollary to Proposition IV, to see that  $\mu(\varphi) = 0$ , first observe that

$$\mu(\varphi)(a_0) = \varphi^n(a_0) - \sum_{v=0}^{n-1} \alpha_v \varphi^v(a_0) = \varphi(a_{n-1}) - \sum_{v=0}^{n-1} \alpha_v a_v = 0$$

Now if  $\mu(\varphi)(a_v) = 0$  for some  $0 \leq v \leq n-2$ , then

$$\mu(\varphi)(a_{v+1}) = \mu(\varphi)(\varphi(a_v)) = \varphi(\mu(\varphi)(a_v)) = \varphi(0) = 0$$

By induction,  $\mu(\varphi)(a_v) = 0$  for all  $0 \leq v \leq n-1$ , so  $\mu(\varphi) = 0$ .

*Remark.* If  $E$  and  $E^*$  are dual spaces and  $\varphi : E \rightarrow E$  and  $\varphi^* : E^* \leftarrow E^*$  are dual transformations, then  $E$  is cyclic with respect to  $\varphi$  if and only if  $E^*$  is cyclic with respect to  $\varphi^*$ .

*Proof.* By Theorem I, if  $E$  is cyclic then

$$\dim E^* = \dim E = \deg \mu_\varphi = \deg \mu_{\varphi^*}$$

so  $E^*$  is cyclic. The converse follows by symmetry.  $\square$

## § 4

*Remark.* If  $E = \sum_i E_i$  is the generalized eigenspace decomposition of  $E$  and  $E_i = \sum_j E_{ij}$  is a cyclic decomposition of  $E_i$ , then  $E = \sum_{i,j} E_{ij}$  is an irreducible decomposition of  $E$ . Conversely, an irreducible decomposition of  $E$  yields an irreducible (and hence cyclic) decomposition of each  $E_i$ .<sup>24</sup>

*Remark.* In subsection 13.14, observe that

$$\begin{aligned} \sum_{v=0}^{p-1} \alpha_v f^{i-1}(\varphi) \varphi^v e + f(\varphi)^{i-1} \varphi^p e &= f(\varphi)^{i-1} \sum_{v=0}^p \alpha_v \varphi^v e \\ &= f(\varphi)^{i-1} f(\varphi) e \\ &= f(\varphi)^i e \end{aligned}$$

*Remark.* In subsection 13.17, observe from the formula for  $r(\psi^j)$  that we can first determine  $N_k$  from  $r(\psi^{k-1})$ , then determine  $N_{k-1}$  from  $r(\psi^{k-2})$  and  $N_k$ , then determine successively  $N_{k-2}, \dots, N_1$ .

<sup>24</sup>See the remarks at the end of subsection 13.17.



*Remark.* Proposition II shows that two linear transformations are conjugate if and only if they have the same Jordan canonical matrix. The matrix is therefore a particularly simple representative for the conjugacy class.

The proof of case II can be simplified by noting that  $\dim E_i = \dim F_i$  follows from the fact that  $r(f_i(\varphi_i)^j) = r(f_i(\psi_i)^j)$  for all  $j \geq 1$ .

*Remark.* In the proof of Corollary II to Proposition II, it follows from the results of subsection 13.20 that the common characteristic polynomial of the induced restrictions of  $\varphi$  and  $\psi$  to  $E_i$  and  $F_i$  is just  $f_i^{m_i}$ .<sup>25</sup>

**Exercise (5).** If  $\varphi : E \rightarrow E$  stabilizes  $E_1$ , and  $\varphi_1 : E_1 \rightarrow E_1$  and  $\bar{\varphi} : E/E_1 \rightarrow E/E_1$  are the induced transformations, then  $E$  is cyclic if and only if (i)  $E_1$  is cyclic, (ii)  $E/E_1$  is cyclic, and (iii)  $\mu = \mu_1 \bar{\mu}$ , where  $\mu = \mu_\varphi$ ,  $\mu_1 = \mu_{\varphi_1}$ , and  $\bar{\mu} = \mu_{\bar{\varphi}}$ .

*Proof.* If  $E$  is cyclic, then  $\sigma_a : \Gamma[t] \rightarrow E$  is surjective for some  $a \in E$ . It follows that  $\pi\sigma_a : \Gamma[t] \rightarrow E/E_1$  is also surjective, where  $\pi : E \rightarrow E/E_1$  is the canonical projection. But

$$(\pi\sigma_a)(f) = \pi(f(\varphi)(a)) = f(\bar{\varphi})(\bar{a}) = \sigma_{\bar{a}}(f)$$

so  $\sigma_{\bar{a}} = \pi\sigma_a$  is surjective, and therefore  $E/E_1$  is cyclic. Now

$$\deg \mu_1 \leq \dim E_1 = \dim E - \dim E/E_1 = \deg \mu - \deg \bar{\mu} \leq \deg \mu_1$$

since  $\mu \mid \mu_1 \bar{\mu}$ . It follows that  $\dim E_1 = \deg \mu_1$  so  $E_1$  is cyclic, and  $\deg \mu = \deg \mu_1 \bar{\mu}$  so  $\mu = \mu_1 \bar{\mu}$ .

Conversely, if  $E_1$  and  $E/E_1$  are cyclic and  $\mu = \mu_1 \bar{\mu}$ , then

$$\dim E = \dim E_1 + \dim E/E_1 = \deg \mu_1 + \deg \bar{\mu} = \deg \mu_1 \bar{\mu} = \deg \mu$$

so  $E$  is cyclic. □

*Remark.* If  $\varphi = \varphi_1 \oplus \varphi_2$  with  $\mu_2 = \mu_{\varphi_2}$ , then  $\bar{\mu} = \mu_2$ , so if  $E$  is cyclic it follows that  $\mu = \mu_1 \mu_2$ . Since also  $\mu = \mu_1 \wedge \mu_2$ , this implies that  $\mu_1$  and  $\mu_2$  are relatively prime.

**Exercise (6).** Let  $E = F_1 \oplus \cdots \oplus F_s$  and  $\varphi = \varphi_1 \oplus \cdots \oplus \varphi_s$  where  $\varphi_j : F_j \rightarrow F_j$  with  $\mu = \mu_\varphi$  and  $\mu_j = \mu_{\varphi_j}$ .

(a)  $E$  is cyclic if and only if each  $F_j$  is cyclic and the  $\mu_j$  are relatively prime.

(b) If  $E$  is cyclic, then each  $F_j$  is a sum of generalized eigenspaces for  $\varphi$ .

---

<sup>25</sup>See the remarks in § 5 below.

(c) If  $E$  is cyclic and

$$a = a_1 + \cdots + a_s \quad a_j \in F_j$$

is any vector in  $E$ , then  $a$  generates  $E$  if and only if each  $a_j$  generates  $F_j$ .

*Proof.* For (a), if  $E$  is cyclic then by problem 5 each  $F_j$  is cyclic and (by induction on  $s$ )

$$\mu_1 \wedge \cdots \wedge \mu_s = \mu = \mu_1 \cdots \mu_s$$

so the  $\mu_j$  are relatively prime. Conversely, if each  $F_j$  is cyclic and the  $\mu_j$  are relatively prime, then

$$\begin{aligned} \dim E &= \dim F_1 + \cdots + \dim F_s \\ &= \deg \mu_1 + \cdots + \deg \mu_s \\ &= \deg \mu_1 \cdots \mu_s \\ &= \deg \mu_1 \wedge \cdots \wedge \mu_s \\ &= \deg \mu \end{aligned}$$

so  $E$  is cyclic.

For (b), since  $\mu = \mu_1 \cdots \mu_s$  and the  $\mu_j$  are relatively prime, the generalized eigenspaces of  $\varphi_j$  in  $F_j$  are generalized eigenspaces of  $\varphi$  in  $E$ . It follows that the projection operator associated with  $F_j$  is a polynomial  $g_j(\varphi)$  in  $\varphi$ .

For (c), if  $a$  generates  $E$  and  $x \in F_j$ , then there is  $f \in \Gamma[t]$  with

$$x = f(\varphi)a = f(\varphi_1)a_1 + \cdots + f(\varphi_s)a_s = f(\varphi_j)a_j$$

so  $a_j$  generates  $F_j$ . Conversely, if each  $a_j$  generates  $F_j$  and

$$x = x_1 + \cdots + x_s \quad x_j \in F_j$$

is a vector in  $E$ , there are  $f_j \in \Gamma[t]$  with  $x_j = f_j(\varphi_j)a_j$ . Define  $f = \sum f_j g_j$ . Then

$$f(\varphi)a = \sum f_j(\varphi)g_j(\varphi)a = \sum f_j(\varphi)a_j = \sum x_j = x$$

so  $a$  generates  $E$ . □

**Exercise (7).** If  $\varphi : E \rightarrow E$  stabilizes  $F$ , and  $\varphi_F : F \rightarrow F$  and  $\bar{\varphi} : E/F \rightarrow E/F$  are the induced transformations, then  $E$  is irreducible if and only if (i)  $E/F$  is irreducible, (ii)  $F$  is irreducible, and (iii)  $\mu_{\varphi_F} = f^k$ ,  $\mu_{\bar{\varphi}} = f^l$ , and  $\mu_\varphi = f^{k+l}$  where  $f$  is irreducible.

*Proof.* By Theorem I and problem 5. □

**Exercise (8).** Let  $E$  be irreducible with respect to  $\varphi$ , with  $\mu_\varphi = f^k$  ( $f$  irreducible).

(a) The  $k$  subspaces  $K(f), \dots, K(f^k)$  are the only nontrivial stable subspaces of  $E$ .

(b)  $\text{Im } f(\varphi)^m = K(f^{k-m})$  for  $0 \leq m \leq k$ .

*Proof.* We may assume that  $f$  is monic with  $\deg f = p$ . For (a), we know that

$$0 = K(1) \subset K(f) \subset \dots \subset K(f^k) = E$$

are stable subspaces of  $E$ . If  $F \neq 0$  is a stable subspace of  $E$ , then  $\mu_F = f^m$  for some  $1 \leq m \leq k$  and hence  $F \subseteq K(f^m)$ . Since  $F$  and  $K(f^m)$  are both cyclic,

$$\dim F = pm = \dim K(f^m)$$

so  $F = K(f^m)$ .

For (b),  $\text{Im } f(\varphi)^m$  is stable and

$$\begin{aligned} \dim \text{Im } f(\varphi)^m &= \dim E - \dim \ker f(\varphi)^m \\ &= pk - pm \\ &= p(k - m) \\ &= \dim \ker f(\varphi)^{k-m} \end{aligned}$$

so we must have  $\text{Im } f(\varphi)^m = K(f^{k-m})$  by (a). □

## § 5

*Remark.* The proof of Lemma I is essentially the same as the proof that the product of cyclic groups of orders  $m$  and  $n$  is cyclic if and only if  $m$  and  $n$  are relatively prime.

In the proof, if  $Y_i$  is the cyclic subspace generated by  $y_i$ , then  $Y_i = Y \cap F_i$ . The contradiction is that  $Y = Y_1 \oplus Y_2$  is a cyclic subspace but the orders of  $y_1$  and  $y_2$  under  $f_1(\varphi)$  are not relatively prime.<sup>26</sup>

*Remark.* The minimum polynomial  $\mu$  and the characteristic polynomial  $\chi$  of a linear transformation  $\varphi : E \rightarrow E$  have the same prime factors.

---

<sup>26</sup>See also problem 6 in § 4 above.

*Proof.* Let  $E = \sum_i F_i$  be a cyclic decomposition of  $E$  and  $\varphi_i : F_i \rightarrow F_i$  the induced restriction of  $\varphi$  with minimum polynomial  $\mu_i$  and characteristic polynomial  $\chi_i$ . Then

$$\mu = \bigwedge \mu_i$$

and

$$\chi = \prod \chi_i = \pm \prod \mu_i$$

by the cyclic case of the Cayley-Hamilton theorem (Lemma II). It follows that the prime factors of  $\mu$  and  $\chi$  are the prime factors of the  $\mu_i$ .<sup>27</sup>  $\square$

*Remark.* If  $\varphi : E \rightarrow E$ , let

$$\chi = f_1^{m_1} \cdots f_r^{m_r} \quad \text{and} \quad \mu = f_1^{k_1} \cdots f_r^{k_r} \quad (k_i \leq m_i)$$

be the prime decompositions of the characteristic and minimum polynomials, by the Cayley-Hamilton theorem. If

$$E = E_1 \oplus \cdots \oplus E_r$$

is the generalized eigenspace decomposition of  $E$  and  $\varphi_i : E_i \rightarrow E_i$  the induced restriction of  $\varphi$  with minimum polynomial  $\mu_i$  and characteristic polynomial  $\chi_i$ , then

$$\chi_i = f_i^{m_i} \quad \text{and} \quad \mu_i = f_i^{k_i}$$

*Proof.* We already know that  $\mu_i = f_i^{k_i}$ , and it follows from the previous remark that  $f_i$  does not divide  $\chi_j$  if  $i \neq j$ . On the other hand  $\chi = \chi_1 \cdots \chi_r$ , so we must have  $\chi_i = f_i^{m_i}$ .  $\square$

*Remark.* This result shows that the algebraic multiplicity of an eigenvalue  $\lambda$  (the number of times it occurs as a root of the characteristic polynomial) is the dimension of the generalized eigenspace corresponding to  $\lambda$ , which is greater than or equal to the geometric multiplicity (the dimension of the eigenspace).

*Remark.* In subsection 13.20, we see that

$$\mu_\varphi(t) = \mu_{\varphi - \lambda I}(t - \lambda)$$

Substituting  $t + \lambda$  for  $t$ , we obtain

$$\mu_{\varphi - \lambda I}(t) = \mu_\varphi(t + \lambda)$$

In particular, it follows that  $\mu_\varphi(\lambda)$  is the constant term of  $\mu_{\varphi - \lambda I}$ , and  $\mu_{\varphi - \lambda I}(-\lambda)$  is the constant term of  $\mu_\varphi$ .

---

<sup>27</sup>It also follows from this proof that  $\mu \mid \chi$ .

*Remark.* In subsections 13.21 and 13.22, we have

$$C^2(\varphi) = \Gamma(\varphi) \subseteq C(\varphi) = C(\Gamma(\varphi)) = C^3(\varphi)$$

**Exercise (1).** If  $\varphi : E \rightarrow E$ , let  $\mu = f_1^{k_1} \cdots f_r^{k_r}$  be the prime decomposition of the minimum polynomial with  $p_i = \deg f_i$  and  $E = \sum_i E_i$  the generalized eigenspace decomposition. Denote by  $N_{ij}$  the number of irreducible subspaces of  $E_i$  of dimension  $p_i j$  for  $1 \leq j \leq k_i$  and set

$$l_i = \sum_{j=1}^{k_i} j N_{ij}$$

Then the characteristic polynomial of  $\varphi$  is  $\chi = f_1^{l_1} \cdots f_r^{l_r}$ .

*Proof.* By a remark above, we know that the characteristic polynomial of the induced restriction  $\varphi_i : E_i \rightarrow E_i$  is  $\chi_i = f_i^{m_i}$  where

$$p_i m_i = \deg \chi_i = \dim E_i = \sum_{j=1}^{k_i} p_i j N_{ij} = p_i l_i$$

so  $m_i = l_i$  and

$$\chi = \chi_1 \cdots \chi_r = f_1^{l_1} \cdots f_r^{l_r} \quad \square$$

**Exercise (2).**  $E$  is cyclic with respect to  $\varphi$  if and only if  $\chi_\varphi = \pm \mu_\varphi$ .

*Proof.* The forward direction follows from Lemma II, and the reverse from

$$\dim E = \deg \chi_\varphi = \deg \mu_\varphi \quad \square$$

**Exercise (3).** If  $\varphi : E \rightarrow E$ , and  $E = \sum_j F_j$  is a decomposition of  $E$  as a direct sum of stable subspaces  $F_j$ , then each  $F_j$  is a sum of generalized eigenspaces for  $\varphi$  if and only if each  $F_j$  is stable under  $C(\varphi)$ .

*Proof.* If each  $F_j$  is a sum of generalized eigenspaces, then by Theorem I the projection operators  $\rho_j$  are in  $\Gamma(\varphi)$ . If  $\psi \in C(\varphi)$ , then  $\psi$  commutes with  $\rho_j$ , so

$$\psi F_j = \psi \rho_j F_j = \rho_j \psi F_j \subseteq F_j$$

Conversely, if each  $F_j$  is stable under  $C(\varphi)$  and  $\psi \in C(\varphi)$ , then for  $x = \sum_j x_j$  with  $x_j \in F_j$ , we have

$$\rho_j \psi x = \psi x_j = \psi \rho_j x$$

so  $\rho_j$  commutes with  $\psi$ . Since  $\psi$  was arbitrary,  $\rho_j \in C^2(\varphi)$ . But  $C^2(\varphi) = \Gamma(\varphi)$  by Theorem III, so  $\rho_j \in \Gamma(\varphi)$  and each  $F_j$  is a sum of generalized eigenspaces by Theorem I.  $\square$

**Exercise (4).** Let  $\varphi : E \rightarrow E$  and let  $P$  be the set of projection operators in  $C(\varphi)$ .

(a)  $P = \{0, \iota\}$  if and only if  $E$  is irreducible.

(b)  $P \subseteq C^2(\varphi)$  if and only if  $E$  is cyclic.

*Proof.* By Proposition I,  $P$  consists of the projection operators for stable direct sum decompositions of  $E$ . For (a), the condition is therefore equivalent to the existence of only trivial stable decompositions. For (b), since  $C^2(\varphi) = \Gamma(\varphi)$  by Theorem III, the condition is equivalent to the existence of only generalized eigenspace decompositions by Theorem I, which is equivalent to  $E$  being cyclic by problem 6 in § 4 above.  $\square$

**Exercise (5).** Let  $\varphi : E \rightarrow E$ .

(a)  $C^3(\varphi) = C(\varphi)$

(b)  $C^2(\varphi) = C(\varphi)$  if and only if  $E$  is cyclic.

*Proof.* By a remark above and problem 4.  $\square$

## § 6

*Remark.* If  $\varphi : E \rightarrow E$  is nilpotent of degree  $k$  and  $E$  is cyclic with respect to  $\varphi$ , then  $a$  generates  $E$  if and only if  $k$  is least with  $\varphi^k a = 0$ , which is true if and only if  $\varphi^{k-1} a \neq 0$ .

*Proof.* If  $a$  generates  $E$ , then the vectors

$$a, \varphi a, \dots, \varphi^{k-1} a \quad (1)$$

form a basis of  $E$  and are therefore nonzero, while  $\varphi^k a = 0$ . Conversely, if  $k$  is least with  $\varphi^k a = 0$  and the vectors (1) are linearly dependent, there are scalars  $\alpha_0, \dots, \alpha_{k-1}$  not all zero with

$$\alpha_0 a + \alpha_1 \varphi a + \dots + \alpha_{k-1} \varphi^{k-1} a = 0 \quad (2)$$

Let  $j$  be least with  $\alpha_j \neq 0$ . Applying  $\varphi^{k-j-1}$  to (2) yields

$$\alpha_j \varphi^{k-1} a = 0$$

which is a contradiction.  $\square$

*Remark.* In the proof of Theorem I, Proposition I is used as a lemma. On the other hand, Proposition I follows from Theorem I by the remarks at the end of subsection 13.23.

**Exercise (1).** If  $\varphi : E \rightarrow E$  is nilpotent and  $N_\lambda$  is the number of subspaces of dimension  $\lambda$  in an irreducible decomposition of  $E$ , then

$$\dim \ker \varphi = \sum_{\lambda} N_{\lambda}$$

*Proof.* From subsection 13.17,

$$r(\varphi) = \sum_{\lambda} (\lambda - 1) N_{\lambda} = \dim E - \sum_{\lambda} N_{\lambda}$$

so

$$\sum_{\lambda} N_{\lambda} = \dim E - \dim \operatorname{Im} \varphi = \dim \ker \varphi \quad \square$$

**Exercise (3).** Let  $\varphi : E \rightarrow E$  and  $\varphi^* : E^* \leftarrow E^*$  be dual. If  $\varphi$  is nilpotent of degree  $k$  and  $E$  is cyclic with respect to  $\varphi$ , then  $\varphi^*$  is nilpotent of degree  $k$  and  $E^*$  is cyclic with respect to  $\varphi^*$ . If  $a$  generates  $E$ , then  $a^*$  generates  $E^*$  if and only if

$$\langle a^*, \varphi^{k-1} a \rangle \neq 0$$

*Proof.* We know  $\mu_{\varphi^*} = \mu_{\varphi} = t^k$ , so  $\varphi^*$  is nilpotent of degree  $k$ , and

$$\dim E^* = \dim E = k$$

so  $E^*$  is cyclic with respect to  $\varphi^*$ .<sup>28</sup>

By a remark above,  $a^*$  generates  $E^*$  if and only if  $(\varphi^*)^{k-1} a^* \neq 0$ . Since

$$\langle (\varphi^*)^{k-1} a^*, \varphi^j a \rangle = \langle a^*, \varphi^{k+j-1} a \rangle = 0 \quad (1)$$

for all  $j \geq 1$  and the  $(\varphi^j a)_{j \geq 0}$  form a basis of  $E$ , it follows that  $(\varphi^*)^{k-1} a^* = 0$  if and only if (1) holds for  $j = 0$ .  $\square$

**Exercise (4).** A linear transformation is diagonalizable if and only if it is semi-simple and its minimum polynomial is completely reducible, or equivalently if its minimum polynomial is the product of distinct monic linear factors.

---

<sup>28</sup>See the remark in § 4 above.

*Proof.* A linear transformation is diagonalizable if and only if its generalized eigenspace decomposition is an eigenspace decomposition, which is true if and only if the minimum polynomial has the indicated form. The rest follows from Theorem I.  $\square$

**Exercise (5).** If  $\varphi, \psi : E \rightarrow E$  are commuting diagonalizable transformations, they are simultaneously diagonalizable (that is, there exists a basis of  $E$  which diagonalizes both  $\varphi$  and  $\psi$ ).

*Proof.* Let

$$\sum_i E_i = E = \sum_j F_j$$

be the generalized eigenspace decompositions for  $\varphi$  and  $\psi$  respectively, which are eigenspace decompositions by diagonalizability of  $\varphi$  and  $\psi$ . It follows from commutativity of  $\varphi$  and  $\psi$  and the results of subsection 13.21 that there is an induced decomposition

$$E = \sum_{i,j} E_i \cap F_j$$

A basis of  $E$  obtained by joining bases of the subspaces  $E_i \cap F_j$  diagonalizes both  $\varphi$  and  $\psi$ .  $\square$

*Remark.* It follows from this problem and the previous problem that if  $\varphi$  and  $\psi$  are commuting semisimple transformations over the field of complex numbers (which is algebraically closed), then  $\varphi + \psi$  and  $\varphi\psi$  are semisimple.

**Exercise (6).** Let  $E$  be a complex vector space and  $\varphi : E \rightarrow E$ . Let  $E = \sum_i E_i$  be the generalized eigenspace decomposition of  $E$  and  $\pi_i : E \rightarrow E$  the corresponding projection operators. Assume that the minimum polynomial of the induced transformation  $\varphi_i : E_i \rightarrow E_i$  is  $\mu_i = (t - \lambda_i)^{k_i}$ . Then the semisimple part of  $\varphi$  is

$$\varphi_S = \sum_i \lambda_i \pi_i$$

*Proof.* Let  $\iota_i : E_i \rightarrow E_i$  be the identity map and write

$$\varphi_i = (\varphi_i - \lambda_i \iota_i) + \lambda_i \iota_i$$

Note that  $\varphi_i - \lambda_i \iota_i$  is nilpotent,  $\lambda_i \iota_i$  is semisimple, and the two transformations commute. Since

$$\varphi = \sum_i \varphi_i = \sum_i (\varphi_i - \lambda_i \iota_i) + \sum_i \lambda_i \iota_i$$



and direct sums preserve nilpotency, semisimplicity, and commutativity, the result now follows from the uniqueness of the Jordan-Chevalley decomposition (Theorem II).  $\square$

**Exercise (8).** Consider the transformations  $\varphi, \psi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  defined by

$$\varphi(x, y) = (y, 0) \quad \text{and} \quad \psi(x, y) = (y, x)$$

The only 1-dimensional subspace stable under  $\varphi$  is the line  $y = 0$ , while the only 1-dimensional subspaces stable under  $\psi$  are the lines  $y = \pm x$ , so the set  $\{\varphi, \psi\}$  is semisimple but  $\varphi$  is not. Note that  $\varphi$  and  $\psi$  do not commute since, for example

$$\varphi\psi(1, 0) = (1, 0) \neq (0, 0) = \psi\varphi(1, 0)$$

## § 7

**Exercise (2).** If  $E$  is an inner product space, then  $\varphi : E \rightarrow E$  is normal if and only if  $\tilde{\varphi} = f(\varphi)$  for some  $f \in \mathbb{R}[t]$ .

*Proof.* If  $\tilde{\varphi} = f(\varphi)$ , then  $\tilde{\varphi}\varphi = \varphi\tilde{\varphi}$ , so  $\varphi$  is normal. Conversely, if  $\varphi$  is normal let

$$E = E_1 \oplus \cdots \oplus E_r \tag{1}$$

be the orthogonal generalized eigenspace decomposition of  $E$  and  $\varphi_i : E_i \rightarrow E_i$  the induced homothetic restriction (Theorem I). Write  $\varphi_i = \lambda_i \tau_i$  where  $\lambda_i \in \mathbb{R}$  and  $\tau_i$  is a rotation. We know  $\tau_i^{-1} = f_i(\tau_i)$  for some  $f_i \in \mathbb{R}[t]$ .<sup>29</sup> By subsection 8.5, we have the direct sum

$$\tilde{\varphi} = \sum \tilde{\varphi}_i = \sum \lambda_i \tilde{\tau}_i = \sum \lambda_i \tau_i^{-1} = \sum \lambda_i f_i(\tau_i) \tag{2}$$

Define  $g_i \in \mathbb{R}[t]$  by

$$g_i(t) = \begin{cases} 0 & \text{if } \lambda_i = 0 \\ \lambda_i f_i(\lambda_i^{-1} t) & \text{if } \lambda_i \neq 0 \end{cases}$$

Let  $\pi_i = h_i(\varphi)$  be the  $i$ -th projection operator for (1), where  $h_i \in \mathbb{R}[t]$ . It follows from (2) that

$$\tilde{\varphi} = \sum g_i(\varphi) h_i(\varphi)$$

Setting  $f = \sum g_i h_i$ , we have  $\tilde{\varphi} = f(\varphi)$ .  $\square$

---

<sup>29</sup>See problem 12 in § 1.

# Multilinear Algebra

## Chapter 1

### § 1

*Remark.* If  $\lambda^1 \lambda^4 - \lambda^2 \lambda^3 = 0$ , we want  $\xi^1, \xi^2, \eta^1, \eta^2$  with

$$\lambda^1 = \xi^1 \eta^1 \quad \lambda^2 = \xi^1 \eta^2 \quad \lambda^3 = \xi^2 \eta^1 \quad \lambda^4 = \xi^2 \eta^2$$

If  $\lambda^1 = 0$ , then  $\lambda^2 \lambda^3 = \lambda^1 \lambda^4 = 0$ , so  $\lambda^2 = 0$  or  $\lambda^3 = 0$ .

- If  $\lambda^2 = 0$ , we take  $\xi^1 = 0$ ,  $\xi^2 = 1$ ,  $\eta^1 = \lambda^3$ , and  $\eta^2 = \lambda^4$ .
- If  $\lambda^3 = 0$ , we take  $\xi^1 = \lambda^2$ ,  $\xi^2 = \lambda^4$ ,  $\eta^1 = 0$ , and  $\eta^2 = 1$ .

If  $\lambda^1 \neq 0$  and  $\lambda^2 = 0$ , then  $\lambda^4 = 0$  and we take  $\xi^1 = \lambda^1$ ,  $\xi^2 = \lambda^3$ ,  $\eta^1 = 1$ , and  $\eta^2 = 0$ .

If  $\lambda^1 \neq 0$  and  $\lambda^2 \neq 0$ , we take  $\xi^1 = 1$ ,  $\xi^2 = \lambda^3 / \lambda^1 = \lambda^4 / \lambda^2$ ,  $\eta^1 = \lambda^1$ , and  $\eta^2 = \lambda^2$ .

### § 2

*Remark.* In the construction of the first induced bilinear map  $\tilde{\varphi}$ , note that for each  $y \in F$  the linear map  $\varphi(-, y) : E \rightarrow G$  sends  $E_1$  into  $G_1$  and hence induces a linear map  $\bar{\varphi}(-, y) : E/E_1 \rightarrow G/G_1$  by  $\bar{\varphi}(\rho x, y) = \pi \varphi(x, y)$ :

$$\begin{array}{ccc} E & \xrightarrow{\varphi(-, y)} & G \\ \rho \downarrow & & \downarrow \pi \\ E/E_1 & \xrightarrow{\bar{\varphi}(-, y)} & G/G_1 \end{array}$$

Since  $\bar{\varphi}(-, y)$  depends linearly on  $y$ ,  $\bar{\varphi}$  is bilinear and we define  $\tilde{\varphi} = \bar{\varphi}$ .

In the construction of the second induced bilinear map  $\tilde{\varphi}$ , note that the linear map  $y \mapsto \bar{\varphi}(-, y)$  kills  $F_1$ , so it factors through  $\sigma$ :

$$\begin{array}{ccc} F & \xrightarrow{y \mapsto \bar{\varphi}(-, y)} & L(E/E_1; G/G_1) \\ \sigma \downarrow & \nearrow \sigma y \mapsto \bar{\varphi}(-, y) & \\ F/F_1 & & \end{array}$$

This allows us to define the bilinear map  $\tilde{\varphi}(\rho x, \sigma y) = \overline{\varphi}(\rho x, y) = \pi\varphi(x, y)$ .

### § 3

*Remark.* The claim in problem 5(b) is false because it implies, for example, that any two inner products in  $\mathbb{R}^2$  agree on orthogonality, which is false. The claim holds if and only if  $\psi$  preserves linear relations satisfied by  $\varphi$ .

**Exercise (7).** If  $E, E^*$  are finite-dimensional dual spaces and  $\Phi: E^* \times E \rightarrow \Gamma$  is a bilinear function such that

$$\Phi(\tau^{*-1}x^*, \tau x) = \Phi(x^*, x)$$

for every pair of dual automorphisms, then there is  $\lambda \in \Gamma$  such that

$$\Phi(x^*, x) = \lambda \langle x^*, x \rangle$$

*Proof.* Let  $\varphi: E \rightarrow E$  be the linear transformation with  $\Phi(x^*, x) = \langle x^*, \varphi x \rangle$ . If  $\tau$  is an automorphism of  $E$ , then since  $(\tau^{-1})^* = (\tau^*)^{-1}$ ,

$$\langle x^*, \tau^{-1}\varphi\tau(x) \rangle = \langle \tau^{*-1}x^*, \varphi\tau(x) \rangle = \Phi(\tau^{*-1}x^*, \tau x) = \Phi(x^*, x) = \langle x^*, \varphi x \rangle$$

It follows that  $\tau^{-1}\varphi\tau = \varphi$ . Since  $\tau$  was arbitrary,  $\varphi = \lambda \text{id}$  for some  $\lambda \in \Gamma$ , from which the result follows.  $\square$

### § 4

*Remark.* The tensor product  $E \otimes F$  is a universal (initial) object in the category of “vector spaces with bilinear maps of  $E \times F$  into them”. In this category, the objects are bilinear maps  $E \times F \rightarrow G$ , and the arrows are linear maps  $G \rightarrow H$  which respect the bilinear maps:

$$\begin{array}{ccc} E \times F & \xrightarrow{\quad} & H \\ \downarrow & \nearrow & \\ G & & \end{array}$$

Every object  $E \times F \rightarrow G$  in this category can be obtained from the tensor product  $\otimes: E \times F \rightarrow E \otimes F$  in a unique way. This is why  $\otimes$  is said to satisfy the “universal property”. This is only possible because the elements of  $E \otimes F$  satisfy only those relations required to make  $E \otimes F$  into a vector space and to make  $\otimes$  bilinear. By category theoretic abstract nonsense,  $E \otimes F$  is unique up to isomorphism.

## § 5

*Remark.* Lemma 1.5.1 generalizes the consequence of linear independence to tensor products other than scalar multiplication.

We provide an alternative proof. For a linear function  $g \in L(F)$ , consider the bilinear map  $E \times F \rightarrow E$  defined by  $(x, y) \mapsto g(y)x$ . By the universal property of the tensor product, there is a linear map  $h : E \otimes F \rightarrow E$  with  $h(x \otimes y) = g(y)x$ . Now

$$0 = h\left(\sum a_i \otimes b_i\right) = \sum h(a_i \otimes b_i) = \sum g(b_i)a_i$$

By linear independence of the  $a_i$ , it follows that  $g(b_i) = 0$  for all  $i$ . Since  $g$  was arbitrary, it follows that  $b_i = 0$  for all  $i$ .

Note  $h = \iota \otimes g : E \otimes F \rightarrow E \otimes \Gamma$  (see § 16).

*Remark.* Lemma 1.5.2 generalizes the existence and uniqueness of a representation relative to a basis.

## § 8

*Remark.* In the proof of Proposition 1.8.1, note that if  $f$  is injective, it has a left inverse  $g : G \rightarrow E \otimes F$  with  $g \circ f = \iota$ . If  $\psi : E \times F \rightarrow K$  is bilinear, there is a linear map  $h : E \otimes F \rightarrow K$  with

$$\psi = h \circ \otimes = h \circ \iota \circ \otimes = h \circ g \circ f \circ \otimes = h \circ g \circ \varphi$$

So  $\psi$  factors through  $\varphi$ . Since  $\psi$  was arbitrary,  $\varphi$  satisfies  $\otimes_2$ .

**Exercise (3).** If  $S$  and  $T$  are sets, then  $C(S \times T) \cong C(S) \otimes C(T)$ .

*Proof.* By the universal property of the free space, there is a unique linear map  $\varphi : C(S \times T) \rightarrow C(S) \otimes C(T)$  with  $\varphi(s, t) = s \otimes t$  for  $s \in S$  and  $t \in T$ :

$$\begin{array}{ccc} S \times T & \xrightarrow{i \times i} & C(S) \times C(T) \\ \downarrow i & & \downarrow \otimes \\ C(S \times T) & \xrightarrow{\varphi} & C(S) \otimes C(T) \end{array}$$

Observe that  $\varphi$  is injective since

$$0 = \varphi\left(\sum_i \lambda_i(s_i, t_i)\right) = \sum_i \lambda_i s_i \otimes t_i$$

implies  $\lambda_i = 0$  by linear independence of the distinct  $s_i$  in  $C(S)$  and  $t_i$  in  $C(T)$  (1.5.1). Also  $\varphi$  is surjective since

$$\left(\sum_i \lambda_i s_i\right) \otimes \left(\sum_j \mu_j t_j\right) = \sum_{i,j} \lambda_i \mu_j s_i \otimes t_j = \sum_{i,j} \lambda_i \mu_j \varphi(s_i, t_j)$$

and the elements on the left generate  $C(S) \otimes C(T)$ .  $\square$

**Exercise (4).** If  $a \otimes b \neq 0$ , then  $a \otimes b = a' \otimes b'$  if and only if  $a' = \lambda a$  and  $b' = \lambda^{-1} b$  for some  $\lambda \neq 0$ .

*Proof.* For the forward direction, note all the vectors are nonzero by bilinearity of  $\otimes$ . Now

$$a \otimes b + a' \otimes (-b') = 0$$

so we must have  $a' = \lambda a$  for some  $\lambda \neq 0$  (1.5.1). It follows that

$$a \otimes (b - \lambda b') = 0$$

so  $b - \lambda b' = 0$  and  $b' = \lambda^{-1} b$ .

The reverse direction follows from bilinearity of  $\otimes$ .  $\square$

## § 10

*Remark.* To obtain the tensor product of  $E/E_1$  and  $F/F_1$ , just take the tensor product of  $E$  and  $F$  and then kill off all the products by elements of  $E_1$  and by elements of  $F_1$ .

## § 11

*Remark.* The tensor product operation is bilinear on *spaces*:

$$\left(\bigoplus_{\alpha} E_{\alpha}\right) \otimes \left(\bigoplus_{\beta} F_{\beta}\right) = \bigoplus_{\alpha, \beta} E_{\alpha} \otimes F_{\beta}$$

In particular,  $E \otimes 0 = 0 = 0 \otimes E$ .

## § 12

*Remark.* In the proof of (1.2), the idea is that  $E \cong \bigoplus_{\alpha} E_{\alpha}$  and  $F \cong \bigoplus_{\beta} F_{\beta}$ , so

$$E \otimes F \cong \left(\bigoplus_{\alpha} E_{\alpha}\right) \otimes \left(\bigoplus_{\beta} F_{\beta}\right) = \bigoplus_{\alpha, \beta} E_{\alpha} \otimes F_{\beta}$$

by the result for external direct sums in the previous subsection. Observe that  $h = f \otimes g$  (see § 16).

## § 15

*Remark.* By (1.7), the intersection of tensor products is the tensor product of the intersections. Observe that (1.4) and (1.5) are special cases of (1.7).

In the proof of (1.4),  $u_\beta = v_\beta$  follows immediately from Lemma 1.5.2. In the proof of (1.5), if  $z \in (E_1 \otimes F) \cap (E \otimes F_1)$ , then in particular  $z = x + y$  with  $x \in E_1 \otimes F_1$  and  $y \in E_1 \otimes F'$ . Now  $y = z - x \in E \otimes F_1$ , so

$$y \in (E \otimes F_1) \cap (E \otimes F') = E \otimes (F_1 \cap F') = E \otimes 0 = 0$$

by (1.4). Hence  $y = 0$  and  $z = x \in E_1 \otimes F_1$ . Note that this argument makes (1.6) superfluous.

*Remark.* If  $\sum_{i=1}^r x_i \otimes y_i = \sum_{j=1}^s x'_j \otimes y'_j$  and the  $x_i$  are linearly independent, then the  $y_i$  are in the span of the  $y'_j$ .

*Proof.* By induction on  $s$ . If the vectors in  $\{x_i\} \cup \{x'_j\}$  are linearly independent, then  $y_i = 0$  for all  $i$  (1.5.1), so the result holds trivially. Otherwise, since the  $x_i$  are linearly independent, we may assume (relabeling if necessary) that

$$x_s = \sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^{s-1} \mu_j x'_j \quad (\lambda_i, \mu_j \in \Gamma)$$

By bilinearity of the tensor product, it follows that

$$\sum_{i=1}^r x_i \otimes (y_i - \lambda_i y'_s) = \sum_{j=1}^{s-1} x'_j \otimes (y'_j + \mu_j y'_s)$$

By induction, the  $y_i - \lambda_i y'_s$  are in the span of the  $y'_j + \mu_j y'_s$ , so the  $y_i$  are in the span of the  $y'_j$ .  $\square$

*Remark.* If  $\sum_{i=1}^r x_i \otimes y_i = \sum_{j=1}^s x'_j \otimes y'_j$  and the  $x_i$  and  $y_i$  are respectively linearly independent, then  $r \leq s$ .

*Proof.* By the previous remark and the elementary fact that the size of a linearly independent set is at most the size of a spanning set in a subspace.  $\square$

**Exercise (1).** If  $\sum_{i=1}^r x_i \otimes y_i = \sum_{j=1}^s x'_j \otimes y'_j$  and the  $x_i$ ,  $y_i$ ,  $x'_j$ , and  $y'_j$  are each respectively linearly independent, then  $r = s$ .

*Proof.* By the previous remark,  $r \leq s$  and  $s \leq r$ .  $\square$

*Remark.* It follows from this result and the proof of Lemma 1.5.3 that the tensor representations of minimal length are precisely the representations by linearly independent vectors. These representations are *not* unique, as already seen in problem 1.8.4.

**Exercise (2).** A bilinear mapping  $\varphi : E \times F \rightarrow G$  satisfies  $\otimes_2$  if and only if the vectors  $\varphi(x_\alpha, y_\beta)$  are linearly independent whenever the vectors  $x_\alpha \in E$  and  $y_\beta \in F$  are linearly independent.

*Proof.* If  $f : E \otimes F \rightarrow G$  is the induced linear map with  $\varphi = f \circ \otimes$ , then  $\varphi$  satisfies  $\otimes_2$  if and only if  $f$  is injective (1.8.1)—that is, if and only if  $f$  preserves linear independence. But  $f$  preserves linear independence if and only if  $\varphi$  does, since  $\otimes$  does (1.5.1).  $\square$

**Exercise (3).** If  $A \neq 0$  is a finite-dimensional algebra forming a tensor product under the algebra multiplication, then  $\dim A = 1$ .

*Proof.* By (1.3)  $\dim A = (\dim A)^2$ , and  $\dim A \neq 0$ , so  $\dim A = 1$ .  $\square$

**Exercise (5).** If  $E, E^*$  and  $F, F^*$  are pairs of dual spaces of finite dimension and  $\beta : E \times F \rightarrow B(E^*, F^*)$  is the bilinear map given by

$$\beta_{x,y}(x^*, y^*) = \langle x^*, x \rangle \langle y^*, y \rangle$$

then  $(B(E^*, F^*), \beta)$  is the tensor product of  $E$  and  $F$ .

*Proof.* Let  $x_1, \dots, x_n$  be a basis in  $E$  with dual basis  $x^{*1}, \dots, x^{*n}$  in  $E^*$  and let  $y_1, \dots, y_m$  be a basis in  $F$  with dual basis  $y^{*1}, \dots, y^{*m}$  in  $F^*$ . Let  $\varphi^{*kl} : E^* \times F^* \rightarrow \Gamma$  be the basis function in  $B(E^*, F^*)$  defined by

$$\varphi^{*kl}(x^{*i}, y^{*j}) = \delta_i^k \delta_j^l$$

Then

$$\beta_{x_k, y_l}(x^{*i}, y^{*j}) = \langle x^{*i}, x_k \rangle \langle y^{*j}, y_l \rangle = \delta_i^k \delta_j^l = \varphi^{*kl}(x^{*i}, y^{*j})$$

so  $\beta_{x_k, y_l} = \varphi^{*kl}$ . It follows that  $\text{Im } \beta = B(E^*, F^*)$ , so  $\beta$  satisfies  $\otimes_1$ .

If  $\varphi : E \times F \rightarrow G$  is bilinear, define  $f : B(E^*, F^*) \rightarrow G$  by  $f(\varphi^{*kl}) = \varphi(x_k, y_l)$ . Then  $\varphi(x_k, y_l) = f(\beta_{x_k, y_l})$ , so  $\varphi = f\beta$ . It follows that  $\beta$  satisfies  $\otimes_2$ .  $\square$

*Remark.* This result allows us to view an element  $\sum x_i \otimes y_i$  of the tensor product as a bilinear function: given linear scalar substitutions  $x_i \rightarrow \lambda_i$  and  $y_i \rightarrow \mu_i$  as inputs, it produces the scalar  $\sum \lambda_i \mu_i$  as output. This is analogous to viewing a vector as a linear function. Indeed, if we identify  $x$  with  $\langle -, x \rangle$  and  $y$  with  $\langle -, y \rangle$ , then we can identify  $x \otimes y$  with  $\langle -, x \rangle \langle -, y \rangle$ .

## § 16

*Remark.* The tensor product operation is a bifunctor in the category of vector spaces. It maps the objects  $E$  and  $F$  to the object  $E \otimes F$  and the arrows  $\varphi : E \rightarrow E'$  and  $\psi : F \rightarrow F'$  to the arrow  $\varphi \otimes \psi : E \otimes F \rightarrow E' \otimes F'$ .

*Remark.* Corollary III to Proposition 1.16.1 is not established in subsection 1.27, where it is also assumed that  $E'$  and  $F'$  are finite-dimensional. However, it is true. In fact, for fixed bases in  $E, E', F, F'$ , the tensor products of the induced basis maps in  $L(E; E')$  and  $L(F; F')$  are the basis maps in  $L(E \otimes F, E' \otimes F')$  induced by the tensor products of the basis vectors.<sup>30</sup>

## § 19

*Remark.* The proof of (1.12) is captured in the following commutative diagram:

$$\begin{array}{ccc}
 E \otimes F & \xrightarrow{\varphi \otimes \psi} & E' \otimes F' \\
 \pi_1 \otimes \pi_2 \downarrow & \searrow \pi & \nearrow \chi \\
 \overline{E} \otimes \overline{F} & \xrightarrow{g} & \overline{E} \otimes \overline{F}
 \end{array}$$

$\overline{\varphi} \otimes \overline{\psi}$

*Remark.* The “only if” part of the claim in problem 1(a) requires  $E \neq 0$  and  $F \neq 0$ . For example, if  $E \neq 0$  and  $F = 0$  and  $\varphi = 0$  and  $\psi = 0$ , then  $E \otimes F = 0$  and  $\varphi \otimes \psi = 0$  is injective despite the fact that  $\varphi$  is not.

**Exercise (2).** If  $\varphi : E \rightarrow E$  and  $\psi : F \rightarrow F$  are linear with  $\dim E = n$  and  $\dim F = m$ , then

$$\operatorname{tr}(\varphi \otimes \psi) = \operatorname{tr} \varphi \cdot \operatorname{tr} \psi$$

and

$$\det(\varphi \otimes \psi) = (\det \varphi)^m \cdot (\det \psi)^n$$

---

<sup>30</sup>See problem 1.19.4.



*Proof.* Let  $e_1, \dots, e_n$  be a basis of  $E$  and  $f_1, \dots, f_m$  be a basis of  $F$ . Write

$$\begin{aligned}\varphi e_i &= \sum_{k=1}^n \alpha_i^k e_k \\ \psi f_j &= \sum_{l=1}^m \beta_j^l f_l\end{aligned}$$

Then  $e_1 \otimes f_1, \dots, e_n \otimes f_m$  is a basis of  $E \otimes F$  and

$$(\varphi \otimes \psi)(e_i \otimes f_j) = \varphi e_i \otimes \psi f_j = \sum_{k,l} \alpha_i^k \beta_j^l e_k \otimes f_l$$

It follows that

$$\text{tr}(\varphi \otimes \psi) = \sum_{i,j} \alpha_i^i \beta_j^j = \left( \sum_i \alpha_i^i \right) \left( \sum_j \beta_j^j \right) = \text{tr } \varphi \cdot \text{tr } \psi$$

For the determinant, observe that

$$\varphi \otimes \psi = (\varphi \otimes \iota_F) \circ (\iota_E \otimes \psi)$$

so

$$\det(\varphi \otimes \psi) = \det(\varphi \otimes \iota_F) \cdot \det(\iota_E \otimes \psi)$$

But  $M(\varphi \otimes \iota_F) = (\alpha_i^k \delta_j^l)$  is block diagonal with  $m$  blocks each equal to  $M(\varphi)$ , so  $\det(\varphi \otimes \iota_F) = (\det \varphi)^m$ . Similarly  $\det(\iota_E \otimes \psi) = (\det \psi)^n$ .  $\square$

**Exercise (3).** If  $\alpha, \beta : E \rightarrow E$  are linear with  $\dim E = n$ , and  $\Phi : L(E; E) \rightarrow L(E; E)$  is defined by

$$\Phi \sigma = \alpha \circ \sigma \circ \beta$$

then

$$\text{tr } \Phi = \text{tr } \alpha \cdot \text{tr } \beta$$

and

$$\det \Phi = \det(\alpha \circ \beta)^n$$

*Proof.* Let  $e_1, \dots, e_n$  be a basis of  $E$  and let  $\psi_{ij} : E \rightarrow E$  be the induced basis transformation defined by  $\psi_{ij} e_k = \delta_k^i e_j$ . The isomorphism  $\Psi : E \otimes E \rightarrow L(E; E)$  defined by  $\Psi(e_i \otimes e_j) = \psi_{ij}$  induces an isomorphism

$$\hat{\Psi} : L(E \otimes E; E \otimes E) \cong L(L(E; E); L(E; E))$$

by  $\widehat{\Psi}(\sigma \otimes \tau) = \Psi \circ (\sigma \otimes \tau) \circ \Psi^{-1}$ :

$$\begin{array}{ccc} E \otimes E & \xrightarrow{\Psi} & L(E; E) \\ \sigma \otimes \tau \downarrow & & \downarrow \widehat{\Psi}(\sigma \otimes \tau) \\ E \otimes E & \xrightarrow{\Psi} & L(E; E) \end{array}$$

It is easy to verify that

$$\widehat{\Psi}(\beta^* \otimes \alpha) = \Phi$$

where  $\beta^*$  is the transpose of  $\beta$ . Therefore by the previous exercise,

$$\text{tr } \Phi = \text{tr}(\beta^* \otimes \alpha) = \text{tr}(\beta^*) \cdot \text{tr } \alpha = \text{tr } \alpha \cdot \text{tr } \beta$$

and

$$\det \Phi = \det(\beta^* \otimes \alpha) = (\det \beta^*)^n (\det \alpha)^n = \det(\alpha \circ \beta)^n \quad \square$$

*Remark.* It is more natural to do this problem using the composition algebra (see § 26). Indeed, writing  $\alpha = a^* \otimes b$ ,  $\beta = c^* \otimes d$ , and  $\sigma = x^* \otimes y$ , we have

$$\begin{aligned} \alpha \circ \sigma \circ \beta &= (a^* \otimes b) \circ (x^* \otimes y) \circ (c^* \otimes d) \\ &= (a^* \otimes b) \circ (\langle x^*, d \rangle (c^* \otimes y)) \\ &= \langle x^*, d \rangle \langle a^*, y \rangle (c^* \otimes b) \\ &= (\langle x^*, d \rangle c^*) \otimes (\langle a^*, y \rangle b) \\ &= \beta^* x^* \otimes \alpha y \\ &= (\beta^* \otimes \alpha) \sigma \end{aligned}$$

so  $\Phi = \beta^* \otimes \alpha$ . Here  $\beta^* = d \otimes c^*$ .<sup>31</sup> Alternately, we could use the isomorphism  $\Omega$  from Proposition 1.29.1.

## § 20

*Remark.* A *linear* map is universal if and only if it is an isomorphism. In fact, it satisfies  $\otimes_1$  if and only if it is surjective, and it satisfies  $\otimes_2$  if and only if it is injective (has a left inverse). In this sense, a tensor product map is a multilinear analogue of an isomorphism.

---

<sup>31</sup>See problem 1.29.1.

*Remark.* To see that there is a unique isomorphism

$$f : (E_1 \otimes \cdots \otimes E_p) \otimes (E_{p+1} \otimes \cdots \otimes E_{p+q}) \rightarrow E_1 \otimes \cdots \otimes E_{p+q}$$

with

$$f((x_1 \otimes \cdots \otimes x_p) \otimes (x_{p+1} \otimes \cdots \otimes x_{p+q})) = x_1 \otimes \cdots \otimes x_{p+q} \quad (1)$$

observe that for each  $p$ -tuple  $(x_1, \dots, x_p) \in E_1 \times \cdots \times E_p$ , there is a  $q$ -linear map  $\varphi_{x_1 \dots x_p} : E_{p+1} \times \cdots \times E_{p+q} \rightarrow E_1 \otimes \cdots \otimes E_{p+q}$  given by

$$\varphi_{x_1 \dots x_p}(x_{p+1}, \dots, x_{p+q}) = x_1 \otimes \cdots \otimes x_{p+q}$$

By the universal property of  $E_{p+1} \otimes \cdots \otimes E_{p+q}$ , it follows that there is a linear map  $f_{x_1 \dots x_p} : E_{p+1} \otimes \cdots \otimes E_{p+q} \rightarrow E_1 \otimes \cdots \otimes E_{p+q}$  with

$$f_{x_1 \dots x_p}(x_{p+1} \otimes \cdots \otimes x_{p+q}) = x_1 \otimes \cdots \otimes x_{p+q}$$

Now the mapping  $(x_1, \dots, x_p) \mapsto f_{x_1 \dots x_p}$  is  $p$ -linear, so by the universal property of  $E_1 \otimes \cdots \otimes E_p$ , there is a linear map

$$\widehat{f} : E_1 \otimes \cdots \otimes E_p \rightarrow L(E_{p+1} \otimes \cdots \otimes E_{p+q}; E_1 \otimes \cdots \otimes E_{p+q})$$

with  $\widehat{f}(x_1 \otimes \cdots \otimes x_p) = f_{x_1 \dots x_p}$ . Define

$$\varphi : (E_1 \otimes \cdots \otimes E_p) \times (E_{p+1} \otimes \cdots \otimes E_{p+q}) \rightarrow E_1 \otimes \cdots \otimes E_{p+q}$$

by  $\varphi(x, y) = \widehat{f}(x)(y)$ . Then  $\varphi$  is bilinear, so by the universal property of  $(E_1 \otimes \cdots \otimes E_p) \otimes (E_{p+1} \otimes \cdots \otimes E_{p+q})$ , there is a linear map  $f$  satisfying  $f \circ \otimes = \varphi$ , which implies (1).

By the universal property of  $E_1 \otimes \cdots \otimes E_{p+q}$ , the  $(p+q)$ -linear map

$$\psi : E_1 \times \cdots \times E_{p+q} \rightarrow (E_1 \otimes \cdots \otimes E_p) \otimes (E_{p+1} \otimes \cdots \otimes E_{p+q})$$

defined by

$$\psi(x_1, \dots, x_{p+q}) = (x_1 \otimes \cdots \otimes x_p) \otimes (x_{p+1} \otimes \cdots \otimes x_{p+q})$$

induces a linear map

$$g : E_1 \otimes \cdots \otimes E_{p+q} \rightarrow (E_1 \otimes \cdots \otimes E_p) \otimes (E_{p+1} \otimes \cdots \otimes E_{p+q})$$

with

$$g(x_1 \otimes \cdots \otimes x_{p+q}) = (x_1 \otimes \cdots \otimes x_p) \otimes (x_{p+1} \otimes \cdots \otimes x_{p+q})$$

By universal properties again,  $f \circ g = \iota$  and  $g \circ f = \iota$  so  $f$  is an isomorphism, and  $f$  is uniquely determined by (1).

*Remark.* If the vectors  $a_v^i$  are linearly independent in  $E_i$  ( $i = 1, \dots, p$ ), then their tensor products  $a_{v_1}^1 \otimes \dots \otimes a_{v_p}^p$  are linearly independent in  $E_1 \otimes \dots \otimes E_p$ . Indeed, if  $p = 2$ , this follows from Lemma 1.5.1. If  $p > 2$ , we use the natural isomorphism

$$E_1 \otimes \dots \otimes E_p \cong (E_1 \otimes \dots \otimes E_{p-1}) \otimes E_p$$

If

$$\sum_{v_1, \dots, v_p} \lambda_{v_1 \dots v_p} a_{v_1}^1 \otimes \dots \otimes a_{v_p}^p = 0$$

then

$$\sum_{v_1, \dots, v_p} (\lambda_{v_1 \dots v_p} a_{v_1}^1 \otimes \dots \otimes a_{v_{p-1}}^{p-1}) \otimes a_{v_p}^p = 0$$

Since the  $a_{v_p}^p$  are linearly independent in  $E_p$ , it follows from Lemma 1.5.1 that

$$\lambda_{v_1 \dots v_p} a_{v_1}^1 \otimes \dots \otimes a_{v_{p-1}}^{p-1} = 0$$

for all  $v_1, \dots, v_p$ . But  $a_{v_1}^1 \otimes \dots \otimes a_{v_{p-1}}^{p-1} \neq 0$  since  $a_v^i \neq 0$  for all  $i, v$  (see problem 1), so  $\lambda_{v_1 \dots v_p} = 0$  for all  $v_1, \dots, v_p$ .

If the vectors  $a_v^i$  span  $E_i$  ( $i = 1, \dots, p$ ), then clearly their tensor products span  $E_1 \otimes \dots \otimes E_p$ . It follows that if the vectors  $a_v^i$  form a basis of  $E_i$  ( $i = 1, \dots, p$ ), then their tensor products form a basis of  $E_1 \otimes \dots \otimes E_p$ .

*Remark.* To see that for  $\varphi_i : E_i \rightarrow F_i$  the map  $(\varphi_1, \dots, \varphi_p) \mapsto \varphi_1 \otimes \dots \otimes \varphi_p$  induces an injection

$$L(E_1; F_1) \otimes \dots \otimes L(E_p; F_p) \rightarrow L(E_1 \otimes \dots \otimes E_p; F_1 \otimes \dots \otimes F_p)$$

proceed by induction on  $p \geq 2$ . For  $p = 2$ , this is just Proposition 1.16.1. For  $p > 2$ , if we make the appropriate identifications we have

$$\begin{aligned} L(E_1; F_1) \otimes \dots \otimes L(E_p; F_p) &= (L(E_1; F_1) \otimes \dots \otimes L(E_{p-1}; F_{p-1})) \otimes L(E_p; F_p) \\ &\subseteq L(E_1 \otimes \dots \otimes E_{p-1}; F_1 \otimes \dots \otimes F_{p-1}) \otimes L(E_p; F_p) \\ &\subseteq L((E_1 \otimes \dots \otimes E_{p-1}) \otimes E_p; (F_1 \otimes \dots \otimes F_{p-1}) \otimes F_p) \\ &= L(E_1 \otimes \dots \otimes E_p; F_1 \otimes \dots \otimes F_p) \end{aligned}$$

**Exercise (1).** In  $E_1 \otimes \dots \otimes E_p$ :

- (a)  $x_1 \otimes \dots \otimes x_p = 0$  if and only if at least one  $x_i = 0$ .

(b) If  $x_1 \otimes \cdots \otimes x_p \neq 0$ , then

$$x_1 \otimes \cdots \otimes x_p = y_1 \otimes \cdots \otimes y_p$$

if and only if  $y_i = \lambda_i x_i$  with  $\lambda_1 \cdots \lambda_p = 1$ .

*Proof.* The reverse directions follow from multilinearity of the tensor product. The forward directions follow by induction on  $p$  using the natural isomorphism

$$E_1 \otimes \cdots \otimes E_p \cong (E_1 \otimes \cdots \otimes E_{p-1}) \otimes E_p$$

(a) If  $p = 2$ , the result follows from Lemma 1.5.1. If  $p > 2$ , then  $(x_1 \otimes \cdots \otimes x_{p-1}) \otimes x_p = 0$ , so by Lemma 1.5.1 either  $x_1 \otimes \cdots \otimes x_{p-1} = 0$  and  $x_i = 0$  for some  $i = 1, \dots, p-1$  by induction, or else  $x_p = 0$ .

(b) If  $p = 2$ , the result follows from problem 1.8.4. If  $p > 2$ , then

$$(x_1 \otimes \cdots \otimes x_{p-1}) \otimes x_p = (y_1 \otimes \cdots \otimes y_{p-1}) \otimes y_p$$

so by the same problem,

$$y_1 \otimes \cdots \otimes y_{p-1} = \lambda_p^{-1} x_1 \otimes \cdots \otimes x_{p-1} \quad \text{and} \quad y_p = \lambda_p x_p$$

for some  $\lambda_p \neq 0$ . Setting  $\mu = \lambda_p^{-1}/(p-1)$ , we have

$$y_1 \otimes \cdots \otimes y_{p-1} = \mu x_1 \otimes \cdots \otimes \mu x_{p-1}$$

so  $y_i = \mu_i \mu x_i$  for  $i = 1, \dots, p-1$  with  $\mu_1 \cdots \mu_{p-1} = 1$  by induction. Setting  $\lambda_i = \mu_i \mu$  for  $i = 1, \dots, p-1$ , we have  $y_i = \lambda_i x_i$  for  $i = 1, \dots, p$  and  $\lambda_1 \cdots \lambda_p = 1$  as desired.  $\square$

## § 21

*Remark.* For  $p, q \geq 1$ , let

$$\varphi_i : \prod_{j=1}^p E_j^i \rightarrow E_{p+1}^i \quad (i = 1, \dots, q)$$

be a family of  $q$   $p$ -linear maps. Then there is a unique  $p$ -linear map

$$\varphi = \varphi_1 \otimes \cdots \otimes \varphi_q : \prod_{j=1}^p \left( \bigotimes_{i=1}^q E_j^i \right) \rightarrow \bigotimes_{i=1}^q E_{p+1}^i$$

with

$$\varphi(x_1^1 \otimes \cdots \otimes x_1^q, \dots, x_p^1 \otimes \cdots \otimes x_p^q) = \varphi_1(x_1^1, \dots, x_p^1) \otimes \cdots \otimes \varphi_q(x_1^q, \dots, x_p^q)$$

Moreover,

$$\ker_j \varphi = \sum_{i=1}^q E_j^1 \otimes \cdots \otimes \ker_j \varphi_i \otimes \cdots \otimes E_j^q \quad (j = 1, \dots, p)$$

In particular, if each  $\varphi_i$  is nondegenerate, then  $\varphi$  is nondegenerate.

## § 22

*Remark.* If  $\Phi$  is bilinear in  $E \times E'$  and  $\Psi$  is bilinear in  $F \times F'$ , then nondegeneracy of  $\Phi \otimes \Psi$  does not imply nondegeneracy of  $\Phi$  and  $\Psi$  unless all the spaces are nonzero, contrary to what the book says. For example, if  $E \neq 0$  and  $E' = 0$ ,  $F = F' = 0$ , and  $\Phi = 0$  and  $\Psi = 0$ , then  $\Phi$  is degenerate, but  $E \otimes F = 0$  and  $E' \otimes F' = 0$ , so  $\Phi \otimes \Psi = 0$  is nondegenerate.<sup>32</sup>

## § 26

*Remark.* If  $e_1, \dots, e_n$  is a basis in  $E$  and  $f_1, \dots, f_n$  is its dual basis in  $L(E)$ , then the basis transformation  $\varphi_{ij}$  in  $L(E; E)$  with  $\varphi_{ij}(e_k) = \delta_{ik} e_j$  is given by  $x \mapsto f_i(x) e_j$ . Therefore it is natural to consider the isomorphism  $T : L(E) \otimes E \rightarrow L(E; E)$  with  $T(f_i \otimes e_j) = \varphi_{ij}$ . In the algebra induced by  $T^{-1}$  in  $L(E) \otimes E$ ,

$$(f_i \otimes e_j) \circ (f_k \otimes e_l) = T^{-1}(\varphi_{ij} \circ \varphi_{kl}) = T^{-1}(\delta_{il} \varphi_{kj}) = f_i(e_l)(f_k \otimes e_j)$$

This motivates the definition of the composition algebra. An alternative way to discover the isomorphism is (see § 28)

$$L(E; E) \cong L(E \otimes \Gamma; \Gamma \otimes E) \cong L(E; \Gamma) \otimes L(\Gamma; E) \cong L(E) \otimes E$$

**Exercise (1).** For the bilinear map

$$\gamma : L(E, E'; E'') \times L(F, F'; F'') \rightarrow L(E \otimes F, E' \otimes F'; E'' \otimes F'')$$

with

$$\gamma(\varphi, \psi) : (x \otimes y, x' \otimes y') \mapsto \varphi(x, x') \otimes \psi(y, y')$$

the pair  $(\text{Im } \gamma, \gamma)$  is the tensor product of  $L(E, E'; E'')$  and  $L(F, F'; F'')$ .

<sup>32</sup>This is essentially the same error as in problem 1.19.1(a) above.

*Proof.* By transfer of Corollary II of Proposition 1.16.1, using the isomorphism between linear and bilinear maps induced by the tensor product.

In detail, consider the following commutative diagram:

$$\begin{array}{ccc}
 L(E \otimes E'; E'') \otimes L(F \otimes F'; F'') & \xrightarrow{f} & L((E \otimes E') \otimes (F \otimes F'); E'' \otimes F'') \\
 \uparrow \cong & & \downarrow \cong \\
 & & L((E \otimes F) \otimes (E' \otimes F'); E'' \otimes F'') \\
 & & \downarrow \cong \\
 L(E, E'; E'') \otimes L(F, F'; F'') & \xrightarrow{g} & L(E \otimes F, E' \otimes F'; E'' \otimes F'') \\
 \uparrow \otimes & \nearrow \gamma & \\
 L(E, E'; E'') \times L(F, F'; F'') & & 
 \end{array}$$

Note  $g$  is injective since  $f$  is injective (1.16.1), so  $\gamma$  satisfies  $\otimes_2$  (1.8.1). Since  $\gamma$  also satisfies  $\otimes_1$ , it follows that  $\gamma$  is the tensor product.  $\square$

**Exercise (2).** If  $E, E^*$  and  $F, F^*$  are pairs of dual spaces with  $E_1 \subseteq E$  and  $F_1 \subseteq F$  subspaces, then a scalar product is induced between

$$(E^* \otimes F^*) / (E_1^\perp \otimes F^* + E^* \otimes F_1^\perp) \quad \text{and} \quad E_1 \otimes F_1$$

by the scalar product between  $E^* \otimes F^*$  and  $E \otimes F$ . In particular,

$$(E_1 \otimes F_1)^\perp = E_1^\perp \otimes F^* + E^* \otimes F_1^\perp$$

*Proof.* A scalar product is induced between  $E^* / E_1^\perp$  and  $E_1$  by  $\langle \overline{x^*}, x \rangle = \langle x^*, x \rangle$ , and between  $F^* / F_1^\perp$  and  $F_1$  by  $\langle \overline{y^*}, y \rangle = \langle y^*, y \rangle$ . Therefore a scalar product is induced between  $(E^* / E_1^\perp) \otimes (F^* / F_1^\perp)$  and  $E_1 \otimes F_1$  by

$$\langle \overline{x^*} \otimes \overline{y^*}, x \otimes y \rangle = \langle x^*, x \rangle \langle y^*, y \rangle = \langle x^* \otimes y^*, x \otimes y \rangle$$

where the scalar product on the right is between  $E^* \otimes F^*$  and  $E \otimes F$ . However,

$$(E^* \otimes F^*) / (E_1^\perp \otimes F^* + E^* \otimes F_1^\perp) \cong (E^* / E_1^\perp) \otimes (F^* / F_1^\perp)$$

where  $\overline{x^* \otimes y^*} \mapsto \overline{x^*} \otimes \overline{y^*}$ , so there is a scalar product

$$\langle \overline{x^* \otimes y^*}, x \otimes y \rangle = \langle x^* \otimes y^*, x \otimes y \rangle$$

as required.  $\square$

**Exercise (3).** If  $E, E^*$  are dual spaces and  $\varphi : E \rightarrow E$  and  $\varphi^* : E^* \leftarrow E^*$  are dual transformations, then  $\varphi \otimes \varphi^*$  is self-dual.

*Proof.* Recall that  $E \otimes E^*$  is self-dual under the scalar product

$$\langle x \otimes x^*, y \otimes y^* \rangle = \langle x^*, y \rangle \langle y^*, x \rangle$$

Now

$$\begin{aligned} \langle (\varphi \otimes \varphi^*)(x \otimes x^*), y \otimes y^* \rangle &= \langle \varphi x \otimes \varphi^* x^*, y \otimes y^* \rangle \\ &= \langle \varphi^* x^*, y \rangle \langle y^*, \varphi x \rangle \\ &= \langle x^*, \varphi y \rangle \langle \varphi^* y^*, x \rangle \\ &= \langle x \otimes x^*, \varphi y \otimes \varphi^* y^* \rangle \\ &= \langle x \otimes x^*, (\varphi \otimes \varphi^*)(y \otimes y^*) \rangle \end{aligned}$$

so  $(\varphi \otimes \varphi^*)^* = \varphi \otimes \varphi^*$ .  $\square$

## § 28

**In this subsection, all vector spaces are finite-dimensional.**

*Remark.* By (1.30), if we view  $a^* \otimes a$  as a linear transformation, then its trace is just  $\langle a^*, a \rangle$ . This gives a natural (coordinate-free) definition of the trace. This also gives a slick proof that  $\text{tr}(\varphi \otimes \psi) = \text{tr} \varphi \cdot \text{tr} \psi$ . Indeed, if we write  $\varphi = a^* \otimes a$  and  $\psi = b^* \otimes b$ , then we can write  $\varphi \otimes \psi = (a^* \otimes b^*) \otimes (a \otimes b)$  since

$$\begin{aligned} (\varphi \otimes \psi)(x \otimes y) &= \varphi x \otimes \psi y \\ &= (\langle a^*, x \rangle a) \otimes (\langle b^*, y \rangle b) \\ &= \langle a^*, x \rangle \langle b^*, y \rangle (a \otimes b) \\ &= \langle a^* \otimes b^*, x \otimes y \rangle (a \otimes b) \end{aligned}$$

Therefore

$$\text{tr}(\varphi \otimes \psi) = \langle a^* \otimes b^*, a \otimes b \rangle = \langle a^*, a \rangle \langle b^*, b \rangle = \text{tr} \varphi \cdot \text{tr} \psi$$



## § 29

**In this subsection, all vector spaces are finite-dimensional.**

*Remark.* In the proof of Proposition 1.29.1, note that if  $\alpha = a^* \otimes b$  and  $\beta = c^* \otimes d$ , then

$$F(\alpha \otimes \beta)(x^* \otimes y) = \langle a^* \otimes b, x^* \otimes y \rangle (c^* \otimes d) = \langle x^*, b \rangle \langle a^*, y \rangle (c^* \otimes d)$$

Comparing with the calculation after problem 1.19.3 above, we immediately see that  $Q$ , which swaps  $b$  and  $d$ , satisfies

$$F(Q(\alpha \otimes \beta)) = \beta^* \otimes \alpha = \Omega(\alpha \otimes \beta)$$

These results show that we can view  $L(A; A)$  as a tensor product in multiple ways.

**Exercise (1).** Let  $E, E^*$  and  $F, F^*$  be pairs of dual spaces. If  $a^* \in E^*$  and  $b \in F$ , then

$$(a^* \otimes b)^* = b \otimes a^*$$

*Proof.* Recall  $a^* \otimes b : E \rightarrow F$  is defined by  $x \mapsto \langle a^*, x \rangle b$  and  $b \otimes a^* : F^* \rightarrow E^*$  is defined by  $y^* \mapsto \langle y^*, b \rangle a^*$ . For  $y^* \in F^*$  and  $x \in E$ ,

$$\langle (b \otimes a^*) y^*, x \rangle = \langle y^*, b \rangle \langle a^*, x \rangle = \langle y^*, (a^* \otimes b) x \rangle \quad \square$$

**Exercise (2).** If  $E, F \neq 0$  are Euclidean spaces with  $\varphi : E \rightarrow E$  and  $\psi : F \rightarrow F$ , then  $\varphi \otimes \psi : E \otimes F \rightarrow E \otimes F$  is a rotation if and only if  $\varphi = \lambda \tau_E$  and  $\psi = \lambda^{-1} \tau_F$  where  $\tau_E$  and  $\tau_F$  are rotations and  $\lambda \neq 0$ .

*Proof.* If  $\varphi = \lambda \tau_E$  and  $\psi = \lambda^{-1} \tau_F$ , then  $\varphi \otimes \psi = \tau_E \otimes \tau_F$ , and

$$\begin{aligned} \widetilde{\tau_E \otimes \tau_F} \circ (\tau_E \otimes \tau_F) &= (\widetilde{\tau_E} \otimes \widetilde{\tau_F}) \circ (\tau_E \otimes \tau_F) \\ &= (\widetilde{\tau_E} \circ \tau_E) \otimes (\widetilde{\tau_F} \circ \tau_F) \\ &= \iota_E \otimes \iota_F \\ &= \iota_{E \otimes F} \end{aligned}$$

so  $\widetilde{\varphi \otimes \psi} = (\varphi \otimes \psi)^{-1}$  and  $\varphi \otimes \psi$  is a rotation.

Conversely if  $\varphi \otimes \psi$  is a rotation, then

$$\widetilde{\varphi \otimes \psi} = \widetilde{\varphi \otimes \psi} = (\varphi \otimes \psi)^{-1} = \varphi^{-1} \otimes \psi^{-1}$$

It follows that  $\tilde{\varphi} = \mu\varphi^{-1}$  and  $\tilde{\psi} = \mu^{-1}\psi^{-1}$  for some  $\mu \neq 0$ .<sup>33</sup> We may assume  $\mu > 0$  (otherwise consider  $-\varphi$  and  $-\psi$ ). Set  $\lambda = \sqrt{\mu} > 0$  and define  $\tau_E = \lambda^{-1}\varphi$  and  $\tau_F = \lambda\psi$ . Then  $\widetilde{\tau_E} \circ \tau_E = \iota_E$ , so  $\widetilde{\tau_E} = \tau_E^{-1}$  and  $\tau_E$  is a rotation, and similarly for  $\tau_F$ .  $\square$

*Remark.* The proofs of the other parts of this problem in the book are similar.

**Exercise (3).** Let  $E$  be a real vector space with  $\dim E = n$  and  $\varphi, \psi : E \rightarrow E$  be regular transformations.

- (a) If  $n$  is even, then  $\varphi \otimes \psi$  preserves orientation.
- (b) If  $n$  is odd, then  $\varphi \otimes \psi$  preserves orientation if and only if  $\varphi$  and  $\psi$  either both preserve orientation or both reverse orientation.

*Proof.* This follows from  $\det(\varphi \otimes \psi) = (\det \varphi \cdot \det \psi)^n$ .<sup>34</sup>  $\square$

## Chapter 2

### § 2

*Remark.* If  $A$  and  $B$  are algebras and  $A \otimes B$  denotes the tensor product of the underlying *vector spaces*, define  $\varphi : A \times B \times A \times B \rightarrow A \otimes B$  by

$$\varphi(x_1, y_1, x_2, y_2) = x_1 x_2 \otimes y_1 y_2$$

Then  $\varphi$  is multilinear, so it induces a linear map  $f : (A \otimes B) \otimes (A \otimes B) \rightarrow A \otimes B$  satisfying

$$f((x_1 \otimes y_1) \otimes (x_2 \otimes y_2)) = x_1 x_2 \otimes y_1 y_2$$

Now the bilinear map  $f \circ \otimes$  makes  $A \otimes B$  into an algebra with

$$(x_1 \otimes y_1)(x_2 \otimes y_2) = x_1 x_2 \otimes y_1 y_2$$

This construction avoids the use of structure maps.

---

<sup>33</sup>See problem 1.8.4.

<sup>34</sup>See problem 1.19.2.

### § 3

*Remark.* If  $\varphi_1 : A_1 \rightarrow B_1$  and  $\varphi_2 : A_2 \rightarrow B_2$  are algebra homomorphisms and  $\varphi = \varphi_1 \otimes \varphi_2$  is the tensor product of the underlying *linear maps*, then

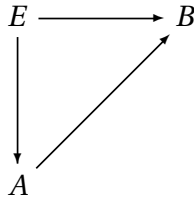
$$\begin{aligned}\varphi((x_1 \otimes x_2)(y_1 \otimes y_2)) &= \varphi(x_1 y_1 \otimes x_2 y_2) \\ &= \varphi_1(x_1 y_1) \otimes \varphi_2(x_2 y_2) \\ &= (\varphi_1 x_1 \varphi_1 y_1) \otimes (\varphi_2 x_2 \varphi_2 y_2) \\ &= (\varphi_1 x_1 \otimes \varphi_2 x_2)(\varphi_1 y_1 \otimes \varphi_2 y_2) \\ &= \varphi(x_1 \otimes x_2) \varphi(y_1 \otimes y_2)\end{aligned}$$

Since  $\varphi$  is linear, it follows that  $\varphi$  is an algebra homomorphism. This proof avoids the use of structure maps.

## Chapter 3

### § 3

*Remark.* The tensor algebra  $\otimes E$  is a universal (initial) object in the category of “associative unital algebras with linear maps of  $E$  into them”. In this category, the objects are linear maps  $E \rightarrow A$ , for associative unital algebras  $A$ , and the arrows are algebra homomorphisms  $A \rightarrow B$  which preserve the units and respect the linear maps from  $E$ :



Every object in this category can be obtained from the tensor algebra  $\otimes E$  in a unique way. This is why  $\otimes E$  is said to satisfy the “universal property”. This is only possible because the elements of  $\otimes E$  satisfy only those properties that are required to make  $\otimes E$  into an associative unital algebra containing  $E$ . By category theoretic abstract nonsense,  $\otimes E$  is unique up to isomorphism (§ 4).

### § 5

*Remark.* We have a functor from the category of vector spaces into the category of associative unital algebras, which sends vector spaces  $E$  and  $F$  to the tensor

algebras  $\otimes E$  and  $\otimes F$ , and which sends a linear map  $\varphi : E \rightarrow F$  to the algebra homomorphism  $\varphi_{\otimes} : \otimes E \rightarrow \otimes F$ .

*Remark.* By a remark from § 21 of chapter 1 above,

$$\ker \varphi_{\otimes} = \sum_{p=0}^{\infty} \sum_{i=1}^p E \otimes \cdots \otimes \ker \varphi \otimes \cdots \otimes E$$

where  $\ker \varphi$  is in position  $i$  of  $p$  on the right.

## § 6

*Remark.* There is a connection between the derivation and the trace of a linear transformation. If  $E$  is an  $n$ -dimensional vector space and  $\varphi : E \rightarrow E$  is linear, let  $\Delta : E^n \rightarrow \Gamma$  be a nonzero determinant function in  $E$ . Recall

$$\operatorname{tr} \varphi \cdot \Delta(x_1, \dots, x_n) = \sum_{i=1}^n \Delta(x_1, \dots, \varphi x_i, \dots, x_n)$$

By the universal property of the tensor product  $\otimes^n E$ , there is an induced linear function  $\Delta_{\otimes} : \otimes^n E \rightarrow \Gamma$  with  $\Delta_{\otimes}(x_1 \otimes \cdots \otimes x_n) = \Delta(x_1, \dots, x_n)$ . Now

$$\begin{aligned} \operatorname{tr} \varphi \cdot \Delta_{\otimes}(x_1 \otimes \cdots \otimes x_n) &= \sum_{i=1}^n \Delta_{\otimes}(x_1 \otimes \cdots \otimes \varphi x_i \otimes \cdots \otimes x_n) \\ &= \Delta_{\otimes}\left(\sum_{i=1}^n x_1 \otimes \cdots \otimes \varphi x_i \otimes \cdots \otimes x_n\right) \\ &= \Delta_{\otimes}(\theta_{(\overline{n})}(\varphi)(x_1 \otimes \cdots \otimes x_n)) \end{aligned}$$

It follows that

$$\Delta_{\otimes} \circ (\operatorname{tr} \varphi \cdot \iota) = \operatorname{tr} \varphi \cdot \Delta_{\otimes} = \Delta_{\otimes} \circ \theta_{(\overline{n})}(\varphi)$$

Since  $\theta_{(\overline{n})}$  is linear by (3.7), it follows that the trace is also linear. If  $\psi : E \rightarrow E$  is another linear transformation, then it follows from (3.8) that

$$\operatorname{tr}(\varphi\psi) - \operatorname{tr}(\psi\varphi) = \operatorname{tr}(\varphi\psi - \psi\varphi) = \operatorname{tr} \varphi \cdot \operatorname{tr} \psi - \operatorname{tr} \psi \cdot \operatorname{tr} \varphi = 0$$

so  $\operatorname{tr}(\varphi\psi) = \operatorname{tr}(\psi\varphi)$ .<sup>35</sup>

---

<sup>35</sup>This is what problem 3.7.3 was supposed to be.

## § 7

**Exercise (1).** If  $u_1 = a_1 \otimes b_1 \neq 0$  and  $u_2 = a_2 \otimes b_2$  are decomposable tensors, then  $u_1 + u_2$  is decomposable if and only if  $a_2 = \lambda a_1$  or  $b_2 = \mu b_1$  for some  $\lambda, \mu \in \Gamma$ .

*Proof.* By the remark after problem 1.15.1 above.  $\square$

## § 11

*Remark.* In the finite-dimensional case, the following are familiar types of mixed tensors in  $E^*, E$ :

Bidegree	Type
(0,0)	Scalar
(0,1)	Vector
(1,0)	Linear function
(1,1)	Linear transformation

## § 12

*Remark.* We have

$$\otimes(E^*, E) = \sum_{p,q} \otimes_q^p(E^*, E)$$

An element of this algebra is a finite sum of decomposable homogeneous mixed tensors. As an example of multiplication, if  $x^* \otimes x$  and  $y^* \otimes y$  are (1,1)-tensors, then

$$(x^* \otimes x)(y^* \otimes y) = x^* \otimes y^* \otimes x \otimes y$$

is a (2,2)-tensor.

*Remark.* The space  $\otimes(E^*, E)$  is clearly dual to itself, and to  $\otimes(E, E^*)$ .

## § 14

*Remark.* The mapping  $T : GL(E) \rightarrow GL(\otimes_q^p(E^*, E))$  given by  $\alpha \mapsto T_\alpha$  is a group homomorphism. The map  $T_\alpha$  is tensorial if  $\alpha$  is in the center of  $GL(E)$ .

*Remark.* We have  $T_\alpha^* = T_{\alpha^{-1}}$ . Indeed, since  $(\alpha^\otimes)^{-1} = (\alpha^{-1})^\otimes$ ,

$$\begin{aligned}\langle y^* \otimes y, T_\alpha(x^* \otimes x) \rangle &= \langle y^* \otimes y, (\alpha^\otimes)^{-1} x^* \otimes \alpha_\otimes x \rangle \\ &= \langle y^*, \alpha_\otimes x \rangle \langle (\alpha^\otimes)^{-1} x^*, y \rangle \\ &= \langle \alpha^\otimes y^*, x \rangle \langle x^*, (\alpha^{-1})_\otimes y \rangle \\ &= \langle \alpha^\otimes y^* \otimes (\alpha^{-1})_\otimes y, x^* \otimes x \rangle \\ &= \langle T_{\alpha^{-1}}(y^* \otimes y), x^* \otimes x \rangle\end{aligned}$$

*Remark.* Writing  $\otimes_q^p(E^*, E) \otimes \otimes_q^p(E^*, E) = \otimes_q^p(E^* \otimes E^*, E \otimes E)$  (see problem 10), we have  $T_\alpha \otimes T_\beta = T_{\alpha \otimes \beta}$ . Writing  $\otimes_q^p(E^*, E) \otimes \otimes_s^r(E^*, E) = \otimes_{q+s}^{p+r}(E^*, E)$ , we have  $T_\alpha \otimes T_\beta = T_\alpha$ .

*Remark.* If  $e_i, e^{*i}$  and  $\bar{e}_i, \bar{e}^{*i}$  are pairs of dual bases in  $E, E^*$  and  $\alpha$  is the change of basis transformation  $e_i \mapsto \bar{e}_i$  in  $E$ , then  $\alpha^{*-1}$  is the corresponding change of basis transformation  $e^{*i} \mapsto \bar{e}^{*i}$  in  $E^*$ .<sup>36</sup> Therefore in  $\otimes_q^p(E^*, E)$ ,

$$T_\alpha(e_{\mu_1 \dots \mu_q}^{v_1 \dots v_p}) = \bar{e}_{\mu_1 \dots \mu_q}^{v_1 \dots v_p}$$

In other words,  $T_\alpha$  is the induced change of basis transformation in  $\otimes_q^p(E^*, E)$ . It follows that a mapping is tensorial if and only if its matrix is invariant under this type of change of basis.

*Remark.* For  $\alpha = \lambda \iota$  with  $\lambda \neq 0$ ,  $\alpha^{*-1} = \lambda^{-1} \iota$ , so  $T_\alpha = \lambda^{q-p} \iota$  on  $\otimes_q^p(E^*, E)$ .

*Remark.* If  $E$  is finite-dimensional and we identify  $\otimes_1^1(E^*, E)$  with  $L(E; E)$  in the natural way, then for any automorphism  $\alpha$  of  $E$ ,

$$T_\alpha(a^* \otimes b) = \alpha \circ (a^* \otimes b) \circ \alpha^{-1}$$

**Exercise (1).** If  $e_v, e^{*v}$  and  $\bar{e}_v, \bar{e}^{*v}$  are pairs of dual bases in  $E, E^*$  related by

$$\bar{e}_v = \sum_\lambda \alpha_v^\lambda e_\lambda \quad \text{and} \quad \bar{e}^{*v} = \sum_\lambda \beta_\lambda^v e^{*\lambda}$$

then the components of a mixed tensor  $w \in \otimes_q^p(E^*, E)$  are related by

$$\bar{\zeta}_{v_1 \dots v_p}^{\mu_1 \dots \mu_q} = \sum_{(\kappa), (\lambda)} \alpha_{v_1}^{\lambda_1} \dots \alpha_{v_p}^{\lambda_p} \beta_{\kappa_1}^{\mu_1} \dots \beta_{\kappa_q}^{\mu_q} \zeta_{\lambda_1 \dots \lambda_p}^{\kappa_1 \dots \kappa_q}$$

<sup>36</sup>See the remark in chapter III, § 3 of [1] above.

*Proof.* In this proof, we use the Einstein summation convention to simplify the notation. Since  $e_\kappa = \beta_\kappa^\mu \bar{e}_\mu$  and  $e^{*\lambda} = \alpha_\nu^\lambda \bar{e}^{*\nu}$ , we have

$$\begin{aligned}
w &= \zeta_{\lambda_1 \dots \lambda_p}^{\kappa_1 \dots \kappa_q} e_{\kappa_1 \dots \kappa_q}^{\lambda_1 \dots \lambda_p} \\
&= \zeta_{\lambda_1 \dots \lambda_p}^{\kappa_1 \dots \kappa_q} e^{*\lambda_1} \otimes \dots \otimes e^{*\lambda_p} \otimes e_{\kappa_1} \otimes \dots \otimes e_{\kappa_q} \\
&= \zeta_{\lambda_1 \dots \lambda_p}^{\kappa_1 \dots \kappa_q} (\alpha_{\nu_1}^{\lambda_1} \bar{e}^{*\nu_1}) \otimes \dots \otimes (\alpha_{\nu_p}^{\lambda_p} \bar{e}^{*\nu_p}) \otimes (\beta_{\kappa_1}^{\mu_1} \bar{e}_{\mu_1}) \otimes \dots \otimes (\beta_{\kappa_q}^{\mu_q} \bar{e}_{\mu_q}) \\
&= \zeta_{\lambda_1 \dots \lambda_p}^{\kappa_1 \dots \kappa_q} \alpha_{\nu_1}^{\lambda_1} \dots \alpha_{\nu_p}^{\lambda_p} \beta_{\kappa_1}^{\mu_1} \dots \beta_{\kappa_q}^{\mu_q} \bar{e}_{\mu_1 \dots \mu_q}^{\nu_1 \dots \nu_p}
\end{aligned}$$

Therefore

$$\bar{\zeta}_{\nu_1 \dots \nu_p}^{\mu_1 \dots \mu_q} = \alpha_{\nu_1}^{\lambda_1} \dots \alpha_{\nu_p}^{\lambda_p} \beta_{\kappa_1}^{\mu_1} \dots \beta_{\kappa_q}^{\mu_q} \zeta_{\lambda_1 \dots \lambda_p}^{\kappa_1 \dots \kappa_q} \quad \square$$

*Remark.* This relationship serves as the basis of an alternative definition of a *tensor* as a multidimensional array of numbers which transforms according to this rule under change of basis.

**Exercise (3).** For  $u^* \in \otimes^p E^*$  and  $u \in \otimes^p E$ ,

$$\langle u^*, u \rangle = (C_1^1)^p (u^* \otimes u)$$

*Proof.* By induction on  $p$ . For  $p = 0$  the result is trivial, and for  $p = 1$  it follows from the definition of  $C_1^1$ . For  $p > 1$ , write  $u^* = u^{*1} \otimes \tilde{u}^*$  and  $u = u_1 \otimes \tilde{u}$ . Then

$$\begin{aligned}
(C_1^1)^p (u^* \otimes u) &= (C_1^1)^{p-1} (C_1^1 (u^* \otimes u)) \\
&= \langle u^{*1}, u_1 \rangle (C_1^1)^{p-1} (\tilde{u}^* \otimes \tilde{u}) \\
&= \langle u^{*1}, u_1 \rangle \langle \tilde{u}^*, \tilde{u} \rangle \\
&= \langle u^*, u \rangle
\end{aligned} \quad \square$$

**Exercise (4).** If the mappings

$$\Phi : \otimes_q^p (E^*, E) \rightarrow \otimes_s^r (E^*, E) \quad \text{and} \quad \Phi^* : \otimes_p^q (E^*, E) \leftarrow \otimes_r^s (E^*, E)$$

are dual and  $\Phi$  is tensorial, then  $\Phi^*$  is tensorial.

*Proof.* If  $\alpha$  is an automorphism of  $E$ , then by a remark above

$$\Phi^* \circ T_\alpha = \Phi^* \circ T_{\alpha^{-1}}^* = (T_{\alpha^{-1}} \circ \Phi)^* = (\Phi \circ T_{\alpha^{-1}})^* = T_{\alpha^{-1}}^* \circ \Phi^* = T_\alpha \circ \Phi^* \quad \square$$

**Exercise (5).** The sum, composite, and tensor product of tensorial mappings are tensorial.

*Proof.* If  $\Phi$  and  $\Psi$  are tensorial, then it is obvious that  $\Phi + \Psi$  and  $\Phi \circ \Psi$  are also tensorial. By a remark above,

$$\begin{aligned}
 (\Phi \otimes \Psi) \circ T_\alpha &= (\Phi \otimes \Psi) \circ (T_\alpha \otimes T_\alpha) \\
 &= (\Phi \circ T_\alpha) \otimes (\Psi \circ T_\alpha) \\
 &= (T_\alpha \circ \Phi) \otimes (T_\alpha \circ \Psi) \\
 &= (T_\alpha \otimes T_\alpha) \circ (\Phi \otimes \Psi) \\
 &= T_\alpha \circ (\Phi \otimes \Psi)
 \end{aligned}$$

so  $\Phi \otimes \Psi$  is tensorial.  $\square$

**Exercise (6).** If  $\Phi : \bigotimes_q^p(E^*, E) \rightarrow \bigotimes_s^r(E^*, E)$  is a nonzero tensorial mapping, then  $r - p = s - q$ .

*Proof.* Let  $\lambda > 1$ . By a remark above,  $\lambda^{q-p}\Phi = \lambda^{s-r}\Phi$  since  $\Phi$  is tensorial. It follows that  $\lambda^{q-p+r-s} = 1$  since  $\Phi \neq 0$ , so  $q - p + r - s = 0$  since  $\lambda > 1$ .  $\square$

**Exercise (7).** If  $E$  is finite-dimensional and for  $a \in E^* \otimes E$  the linear map  $\mu(a)$  is defined by  $z \mapsto a \otimes z$  for  $z \in E$ , then  $\mu(a)$  is tensorial if and only if  $a = \lambda t$  where  $t$  is the unit tensor.

*Proof.* By direct computation, it is easily verified that  $\mu(a)$  is tensorial if and only if  $T_\alpha(a) = a$  for all automorphisms  $\alpha$  of  $E$ . By a remark above, the latter is true if and only if  $a$ , when viewed as a transformation of  $E$ , is preserved under conjugation, which is true if and only if it is a scalar transformation. Since the unit tensor corresponds to the identity transformation, the result follows.  $\square$

**Exercise (8).** If  $E$  is finite-dimensional, then every tensorial map  $\Phi : E^* \otimes E \rightarrow \Gamma$  is of the form  $\Phi = \lambda \cdot C$  where  $C$  is the contraction operator.

*Proof.* This follows from problem 1.3.7 applied to  $\Phi \circ \otimes$ .  $\square$

## § 15

*Remark.* The induced inner products in  $\bigotimes^p E$  and  $\bigotimes E$  are just the induced scalar products from § 8 with  $E^* = E$  under the inner product.



## § 17

*Remark.* The metric tensors encode all of the information required to compute inner products, and hence metric properties (length, angle, etc.), in  $E$  and  $E^*$ .

*Remark.* We have

$$\tau_{\otimes}(g) = \sum_v \tau e_v \otimes \tau e_v = g^*$$

so  $\tau_{\otimes}^{-1}(g^*) = g$ . This yields

$$\begin{aligned} (x^*, y^*) &= (\tau^{-1}x^*, \tau^{-1}y^*) \\ &= \langle g^*, \tau^{-1}x^* \otimes \tau^{-1}y^* \rangle \\ &= \langle g^*, \tau_{\otimes}^{-1}(x^* \otimes y^*) \rangle \\ &= \langle x^* \otimes y^*, \tau_{\otimes}^{-1}g^* \rangle \\ &= \langle x^* \otimes y^*, g \rangle \end{aligned}$$

**Exercise (4).** If  $E, E^*$  is a pair of dual Euclidean spaces and  $\varphi = a^* \otimes b$  is a linear transformation of  $E$ , then  $\tilde{\varphi} = \tau b \otimes \tau^{-1}a^*$ .

*Proof.* Let  $\psi = \tau b \otimes \tau^{-1}a^*$ . For  $x \in E$ ,

$$\psi x = \langle \tau b, x \rangle \tau^{-1}a^* = (x, b) \tau^{-1}a^*$$

so for  $y \in E$ ,

$$(\psi x, y) = (x, b) (\tau^{-1}a^*, y) = (x, b) \langle a^*, y \rangle = (x, \varphi y)$$

Therefore  $\psi = \tilde{\varphi}$ . □

**Exercise (7).** If  $E, E^*$  is a pair of dual Euclidean spaces and  $x_v, x^{*v}$  a pair of dual bases, then

$$g^* = \sum_{v, \mu} (x_v, x_{\mu}) x^{*v} \otimes x^{*\mu} \quad \text{and} \quad g = \sum_{v, \mu} (x^{*v}, x^{*\mu}) x_v \otimes x_{\mu}$$

*Proof.* Since  $x_v \otimes x_{\mu}$  and  $x^{*v} \otimes x^{*\mu}$  are dual bases of  $E \otimes E$  and  $E^* \otimes E^*$ , we have

$$g^* = \sum_{v, \mu} g_{v\mu} x^{*v} \otimes x^{*\mu} \quad \text{and} \quad g = \sum_{v, \mu} g^{v\mu} x_v \otimes x_{\mu}$$

where

$$g_{v\mu} = \langle g^*, x_v \otimes x_{\mu} \rangle = (x_v, x_{\mu})$$

and

$$g^{v\mu} = \langle x^{*v} \otimes x^{*\mu}, g \rangle = (x^{*v}, x^{*\mu})$$

□

*Remark.* It follows that the components of  $g^*$  vary in the same way as the basis vectors in  $E$  under a change of basis, while the components of  $g$  vary inversely. This is why  $g^*$  is called “covariant” and  $g$  is called “contravariant”.

## § 19

*Remark.* The definition of the substitution operator  $i_\nu(h)$  does not specify the behavior on  $p$ -linear functions with  $p < \nu$ .

*Remark.* The definitions of the substitution operators  $i_A(h)$  and  $i_S(h)$  given in this subsection are inappropriate for use with the algebras of skew-symmetric and symmetric multilinear functions, respectively. See below.

## § 23

*Remark.* We have

$$T_q^p(E) \cong (\otimes^p E^*) \otimes (\otimes^q E) = \otimes_q^p(E^*, E)$$

and

$$T(E) \cong (\otimes E^*) \otimes (\otimes E) = \otimes(E^*, E)$$

It follows that  $T_q^p(E)$  is dual to  $T_p^q(E)$ , and  $T(E)$  is dual to itself and to  $T(E^*)$ .

*Remark.* The results in this subsection show that it is possible to develop tensor algebra using multilinear function spaces instead of tensor product spaces (at least in the finite-dimensional case). While it is easier to get started this way, it is messier in the long run.

# Chapter 4

## § 4

*Remark.* We see from (4.10) that the skew-symmetric part of a product depends only on the skew-symmetric parts of the factors.

**Exercise (1).** The map  $\sigma : \otimes^p E \rightarrow \otimes^p E$  is tensorial. If  $E$  is finite-dimensional, then  $\sigma$  is generated by  $\mu(t)$  and  $C_j^i$ , where  $t$  is the unit tensor.

*Proof.* If  $\alpha$  is an automorphism of  $E$  and  $u = x_1 \otimes \cdots \otimes x_p \in \otimes^p E$ , then

$$\sigma T_\alpha u = \alpha x_{\sigma^{-1}(1)} \otimes \cdots \otimes \alpha x_{\sigma^{-1}(p)} = T_\alpha \sigma u$$

so  $\sigma T_\alpha = T_\alpha \sigma$  and hence  $\sigma$  is tensorial.

For the second claim, we may assume that  $\sigma = (12 \cdots j)$  where  $1 < j$ . Write  $t = \sum_v e^{*v} \otimes e_v$  where  $e_v, e^{*v}$  is any pair of dual bases in  $E, E^*$ . Then

$$\begin{aligned} C_{j+1}^1(t \otimes u) &= \sum_v C_{j+1}^1(e^{*v} \otimes e_v \otimes x_1 \otimes \cdots \otimes x_p) \\ &= \sum_v \langle e^{*v}, x_j \rangle e_v \otimes x_1 \otimes \cdots \otimes x_{j-1} \otimes x_{j+1} \otimes \cdots \otimes x_p \\ &= x_j \otimes x_1 \otimes \cdots \otimes x_{j-1} \otimes x_{j+1} \otimes \cdots \otimes x_p \\ &= \sigma u \end{aligned}$$

so  $C_{j+1}^1 \circ \mu(t) = \sigma$ . □

## § 5

*Remark.* In the following subsections, an algebra of skew-symmetric tensors is constructed in two ways: first from the top down, by starting with the tensor algebra  $\otimes E$  and killing off tensors with equal factors in  $N(E)$  to form  $\otimes E / N(E)$ ; and then from the bottom up, by collecting skew-symmetric tensors in  $\otimes E$  to form  $X(E)$ . The two constructions are essentially equivalent by (4.17), although scalar products between dual algebras are defined slightly differently in § 8.

## § 6

*Remark.* If we make the appropriate identifications, we have

$$\pi(\otimes^p E) = \otimes^p E / (N(E) \cap \otimes^p E) = \otimes^p E / N^p(E) = X^p(E)$$

so we already know

$$\otimes E / N(E) = \sum_p \pi(\otimes^p E) = \sum_p X^p(E) = X(E)$$

*Remark.* To see how (4.16) follows from (4.13), write  $u = \pi(a)$  and  $v = \pi(b)$  with  $a \in \otimes^p E$  and  $b \in \otimes^q E$ . By (4.15), (4.17), and (4.13),

$$uv = \pi a \cdot \pi b = \pi(a \otimes b) = \rho \pi_A(a \otimes b) = (-1)^{pq} \rho \pi_A(b \otimes a) = (-1)^{pq} vu$$

## § 7

*Remark.*  $X(E)$  is not a subalgebra of  $\otimes E$ , but can be made into an algebra (see problem 1).

*Remark.* The inverse of  $\rho$  is the induced isomorphism  $\overline{\pi_X} : \otimes E / N(E) \cong X(E)$ :

$$\begin{array}{ccc} \otimes E & \xrightarrow{\pi_X} & X(E) \\ \downarrow \pi & \nearrow \rho^{-1} = \overline{\pi_X} & \\ \otimes E / N(E) & & \end{array}$$

**Exercise (1).** If we define a product in  $X(E)$  by

$$u \cdot v = \pi_A(u \otimes v) \quad u, v \in X(E)$$

then  $X(E)$  becomes an algebra and  $\rho$  an algebra isomorphism. Moreover,

$$\pi_A u \cdot \pi_A v = \pi_A(u \otimes v) \quad u, v \in \otimes E$$

*Proof.* Clearly the product makes  $X(E)$  into an algebra. By (4.17) and (4.15),

$$\rho(u \cdot v) = \rho \pi_A(u \otimes v) = \pi(u \otimes v) = \pi u \cdot \pi v = \rho u \cdot \rho v$$

for all  $u, v \in X(E)$ , so  $\rho$  is an algebra isomorphism. By (4.10),

$$\pi_A u \cdot \pi_A v = \pi_A(\pi_A u \otimes \pi_A v) = \pi_A(u \otimes v)$$

for all  $u, v \in \otimes E$ . □

*Remark.* The proof also shows that the product is uniquely determined by the requirement that  $\rho$  be an algebra isomorphism.

## Chapter 5

### § 1

**Exercise (1).** Let  $L_p(E; F)$  denote the space of  $p$ -linear mappings  $E^p \rightarrow F$ , and  $A_p(E; F)$  the subspace of skew-symmetric mappings. If  $T : L_p(E; F) \rightarrow L_p(E; F)$  is linear such that  $T\varphi = \varphi$  for all  $\varphi \in A_p(E; F)$ , and  $T(\sigma\varphi) = \varepsilon_\sigma T\varphi$  for all  $\sigma \in S_p$  and  $\varphi \in L_p(E; F)$ , then  $T = A$  (the alternator).

*Proof.* By assumption,  $TA = A$  and

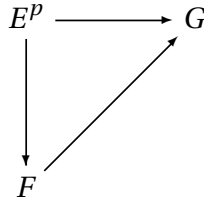
$$(TA)\varphi = T(A\varphi) = \frac{1}{p!} \sum_{\sigma} \varepsilon_{\sigma} T(\sigma\varphi) = \frac{1}{p!} \sum_{\sigma} T\varphi = T\varphi$$

so  $TA = T$ . □

*Remark.* This result shows that the alternator is the unique “skew-symmetric” transformation of  $p$ -linear mappings which fixes skew-symmetric mappings.

### § 3

*Remark.* The  $p$ -th exterior power  $\wedge^p E$  of  $E$  is a universal (initial) object in the category of “vector spaces with skew-symmetric  $p$ -linear maps of  $E$  into them”. In this category, the objects are skew-symmetric  $p$ -linear maps  $E^p \rightarrow F$ , and the arrows are linear maps  $F \rightarrow G$  which respect the  $p$ -linear maps:



Every object  $E^p \rightarrow F$  in this category can be obtained from the exterior product  $\wedge^p : E^p \rightarrow \wedge^p E$  in a unique way. This is why  $\wedge^p$  is said to satisfy the “universal property”. This is only possible because the elements of  $\wedge^p E$  satisfy only those relations required to make  $\wedge^p$  skew-symmetric and  $p$ -linear. By category theoretic abstract nonsense,  $\wedge^p E$  is unique up to isomorphism.

*Remark.* Geometrically, the  $p$ -vector

$$\wedge^p(x_1, \dots, x_p) = x_1 \wedge \cdots \wedge x_p$$

represents the oriented (signed) volume of the  $p$ -dimensional parallelepiped determined by  $x_1, \dots, x_p$ . It can also be thought of as a family of  $p$ -dimensional parallelepipeds with the same orientation and volume.

*Remark.* If  $E$  is  $n$ -dimensional and  $\Delta \neq 0$  is a determinant function in  $E$ , then since  $\wedge^n$  is skew-symmetric and  $n$ -linear in  $E$ ,

$$x_1 \wedge \cdots \wedge x_n = \Delta(x_1, \dots, x_n) \cdot e_1 \wedge \cdots \wedge e_n$$

where  $\Delta(e_1, \dots, e_n) = 1$ . Here  $e_1 \wedge \dots \wedge e_n$  represents the unit volume under  $\Delta$ . On the other hand, since  $\Delta$  is skew-symmetric and  $n$ -linear, there is a linear function  $\Delta_\wedge : \wedge^n E \rightarrow \Gamma$  with

$$\Delta_\wedge(x_1 \wedge \dots \wedge x_n) = \Delta(x_1, \dots, x_n)$$

Note  $\Delta_\wedge$  is inverse to  $e_1 \wedge \dots \wedge e_n \cdot \iota$ , so it is an isomorphism. Also

$$\Delta_\wedge(x_1 \wedge \dots \wedge x_n) = \Delta_\otimes(x_1 \otimes \dots \otimes x_n)$$

These results show that the universality of determinant functions is a special case of the universality of the exterior product.

#### § 4

*Remark.* The projection  $\pi : \otimes E \rightarrow \otimes E / N(E)$  and the projection  $\pi : \otimes E \rightarrow \wedge E$  constructed in this subsection are related by the following commutative diagram:

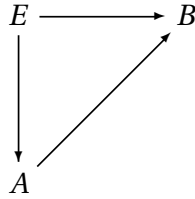
$$\begin{array}{ccc} \otimes E & \xrightarrow{\pi} & \wedge E \\ \pi \downarrow & \nearrow f & \\ \otimes E / N(E) & & \end{array}$$

*Remark.* If  $u \in \wedge^p E$  and  $v \in \wedge^q E$  with  $pq$  even, then

$$\begin{aligned} (u + v)^k &= \frac{1}{k!} \underbrace{(u + v) \wedge \dots \wedge (u + v)}_k \\ &= \frac{1}{k!} \sum_{i+j=k} \frac{k!}{i!(k-i)!} \underbrace{u \wedge \dots \wedge u}_i \wedge \underbrace{v \wedge \dots \wedge v}_j \\ &= \sum_{i+j=k} \frac{1}{i!j!} (i! u^i) \wedge (j! v^j) \\ &= \sum_{i+j=k} u^i \wedge v^j \end{aligned}$$

## § 5

*Remark.* The exterior algebra  $\wedge E$  is a universal (initial) object in the category of “associative unital algebras with linear maps of  $E$  into them whose squares are zero”. In this category, the objects are linear maps  $\varphi : E \rightarrow A$  for associative unital algebras  $A$  with  $\varphi^2 = 0$ , and the arrows are algebra homomorphisms  $A \rightarrow B$  which preserve the units and respect the linear maps:



Every object in this category can be obtained from the exterior algebra  $\wedge E$  in a unique way. This is why  $\wedge E$  is said to satisfy the “universal property”. This is only possible because the elements of  $\wedge E$  satisfy only those properties that are required to make  $\wedge E$  into an associative unital algebra containing  $E$  with skew-symmetric multiplication in  $E$ . By category theoretic abstract nonsense,  $\wedge E$  is unique up to isomorphism.

## § 7

*Remark.* If  $E$  is  $n$ -dimensional, then  $\wedge^p E = 0$  for  $p > n$ . Geometrically, this just reflects the fact that you can’t fit a  $p$ -dimensional parallelepiped into the  $n$ -dimensional space.

*Remark.* If  $E$  is a Euclidean space and  $x_1, \dots, x_p \in E$ , then

$$|x_1 \wedge \cdots \wedge x_p| = \sqrt{\det(x_i, x_j)}$$

is the volume of the  $p$ -dimensional parallelepiped determined by  $x_1, \dots, x_p$ .

**Exercise (1).** Let  $E, E^*$  be a pair of finite-dimensional dual vector spaces and  $A \subseteq L(E^*; E)$  be the subspace of linear maps  $\varphi : E^* \rightarrow E$  for which  $\varphi^* = -\varphi$ . If  $\varphi : E \times E \rightarrow A$  is the bilinear map defined by

$$\varphi_{a,b}(x^*) = \langle x^*, a \rangle b - \langle x^*, b \rangle a$$

then  $(A, \varphi)$  is a second exterior power of  $E$ .

*Proof.* Identifying  $E \otimes E$  with  $L(E^*; E)$  under the natural isomorphism

$$a \otimes b \mapsto (x^* \mapsto \langle x^*, a \rangle b)$$

$A$  is just the subspace of skew-symmetric tensors in  $E \otimes E$ , and

$$\varphi_{a,b} = a \otimes b - b \otimes a$$

is (up to a factor of 2) just the skew-symmetric part of  $a \otimes b$ . □

**Exercise** (4 - Lagrange identity).

$$\left| \begin{array}{ccc} \sum_{v=1}^n \xi_1^v \eta_v^1 & \cdots & \sum_{v=1}^n \xi_1^v \eta_v^p \\ \vdots & \ddots & \vdots \\ \sum_{v=1}^n \xi_p^v \eta_v^1 & \cdots & \sum_{v=1}^n \xi_p^v \eta_v^p \end{array} \right| = \sum_{<} \left| \begin{array}{ccc} \xi_1^{v_1} & \cdots & \xi_1^{v_p} \\ \vdots & \ddots & \vdots \\ \xi_p^{v_1} & \cdots & \xi_p^{v_p} \end{array} \right| \left| \begin{array}{ccc} \eta_{v_1}^1 & \cdots & \eta_{v_1}^p \\ \vdots & \ddots & \vdots \\ \eta_{v_p}^1 & \cdots & \eta_{v_p}^p \end{array} \right| \quad (1)$$

*Proof.* Let  $L$  denote the left-hand side of (1). Let  $E, E^*$  be a pair of dual vector spaces with dual bases  $e_1, \dots, e_n$  and  $e^{*1}, \dots, e^{*n}$ . Define

$$x_i = \sum_{v=1}^n \xi_i^v e_v \quad \text{and} \quad y^{*j} = \sum_{\mu=1}^n \eta_\mu^j e^{*\mu} \quad (i, j = 1, \dots, p)$$

Since

$$\langle y^{*j}, x_i \rangle = \sum_{v=1}^n \xi_i^v \eta_v^j$$

it follows from (5.10) and (5.13) that

$$\begin{aligned} L &= \langle y^{*1} \wedge \cdots \wedge y^{*p}, x_1 \wedge \cdots \wedge x_p \rangle \\ &= \sum_{(v), (\mu)} \xi_1^{v_1} \cdots \xi_p^{v_p} \eta_{\mu_1}^1 \cdots \eta_{\mu_p}^p \langle e^{*\mu_1} \wedge \cdots \wedge e^{*\mu_p}, e_{v_1} \wedge \cdots \wedge e_{v_p} \rangle \\ &= \sum_{(v), (\mu)} \xi_1^{v_1} \cdots \xi_p^{v_p} \eta_{\mu_1}^1 \cdots \eta_{\mu_p}^p \det(\delta_{v_i}^{\mu_j}) \end{aligned}$$

Now

$$\det(\delta_{v_i}^{\mu_j}) = \sum_{\sigma} \varepsilon_{\sigma} \delta_{v_{\sigma(1)}}^{\mu_1} \cdots \delta_{v_{\sigma(p)}}^{\mu_p}$$

so  $\det(\delta_{v_i}^{\mu_j}) = \varepsilon_{\tau}$  if the  $v_i$  are distinct and  $\mu_i = v_{\tau(i)}$  for some permutation  $\tau$ , otherwise  $\det(\delta_{v_i}^{\mu_j}) = 0$ . The  $v_i$  are distinct if and only if  $v_{\sigma(1)} < \cdots < v_{\sigma(p)}$  for



some permutation  $\sigma$ , so we have

$$\begin{aligned} L &= \sum_{v_1 < \dots < v_p} \sum_{\sigma} \xi_1^{v_{\sigma(1)}} \dots \xi_p^{v_{\sigma(p)}} \sum_{\tau} \varepsilon_{\tau} \eta_{v_{\tau\sigma(1)}}^1 \dots \eta_{v_{\tau\sigma(p)}}^p \\ &= \sum_{v_1 < \dots < v_p} \sum_{\sigma} \varepsilon_{\sigma} \xi_1^{v_{\sigma(1)}} \dots \xi_p^{v_{\sigma(p)}} \sum_{\tau} \varepsilon_{\tau\sigma} \eta_{v_{\tau\sigma(1)}}^1 \dots \eta_{v_{\tau\sigma(p)}}^p \end{aligned}$$

Since  $\tau\sigma$  runs over all permutations as  $\tau$  does, we have

$$L = \sum_{\sigma} \left( \sum_{\varepsilon_{\sigma}} \xi_1^{v_{\sigma(1)}} \dots \xi_p^{v_{\sigma(p)}} \right) \left( \sum_{\varepsilon_{\rho}} \eta_{v_{\rho(1)}}^1 \dots \eta_{v_{\rho(p)}}^p \right)$$

which is just (1). □

## § 8

*Remark.* Let  $E$  be an  $n$ -dimensional space and  $\Delta$  a determinant function in  $E$  with  $\Delta(e_1, \dots, e_n) = 1$ . If  $\varphi : E \rightarrow E$  is linear, then

$$\begin{aligned} \varphi_{\wedge}(x_1 \wedge \dots \wedge x_n) &= \varphi x_1 \wedge \dots \wedge \varphi x_n \\ &= \Delta(\varphi x_1, \dots, \varphi x_n) \cdot e_1 \wedge \dots \wedge e_n \\ &= \det \varphi \cdot \Delta(x_1, \dots, x_n) \cdot e_1 \wedge \dots \wedge e_n \\ &= \det \varphi \cdot x_1 \wedge \dots \wedge x_n \end{aligned}$$

for all  $x_1, \dots, x_n \in E$ . Therefore

$$\varphi_{\wedge}^{(n)} = \det \varphi \cdot \iota$$

where  $\varphi_{\wedge}^{(n)}$  is the restriction of  $\varphi_{\wedge}$  to  $\wedge^n E$ . Equivalently

$$\text{tr } \varphi_{\wedge}^{(n)} = \det \varphi$$

## § 10

*Remark.* Let  $E$  be an  $n$ -dimensional space and  $\Delta$  a determinant function in  $E$  with  $\Delta(e_1, \dots, e_n) = 1$ . If  $\varphi : E \rightarrow E$  is linear, then

$$\begin{aligned} \theta_{\wedge}(\varphi)(x_1 \wedge \dots \wedge x_n) &= \sum_{i=1}^n x_1 \wedge \dots \wedge \varphi x_i \wedge \dots \wedge x_n \\ &= \sum_{i=1}^n \Delta(x_1, \dots, \varphi x_i, \dots, x_n) \cdot e_1 \wedge \dots \wedge e_n \\ &= \text{tr } \varphi \cdot \Delta(x_1, \dots, x_n) \cdot e_1 \wedge \dots \wedge e_n \\ &= \text{tr } \varphi \cdot x_1 \wedge \dots \wedge x_n \end{aligned}$$

for all  $x_1, \dots, x_n \in E$ . Therefore

$$\theta_{\wedge}^{(n)}(\varphi) = \text{tr } \varphi \cdot \iota$$

where  $\theta_{\wedge}^{(n)}(\varphi)$  is the restriction of  $\theta_{\wedge}(\varphi)$  to  $\wedge^n E$ . Equivalently

$$\text{tr } \theta_{\wedge}^{(n)}(\varphi) = \text{tr } \varphi$$

This is also equivalent to the result in subsection 3.6 above, by the following commutative diagram:

$$\begin{array}{ccccc}
 & & E^n & & \\
 & & \downarrow \otimes^n & \searrow \Delta & \\
 \otimes^n E & \xrightarrow{\theta_{\otimes}(\varphi)} & \otimes^n E & \xrightarrow{\Delta_{\otimes}} & \Gamma \\
 \downarrow \pi & & \downarrow \pi & \nearrow \Delta_{\wedge} & \\
 \wedge^n E & \xrightarrow{\theta_{\wedge}^{(n)}(\varphi)} & \wedge^n E & & 
 \end{array}$$

## § 11

*Remark.* If  $\omega = \iota_{\wedge}$  and  $\varphi : E \rightarrow E$  is linear, then  $\Omega(\varphi) = \theta_{\wedge}(\varphi)$ . This shows that the construction in § 10 is just a special case of the construction in this subsection.

## § 12

**Exercise (8).** If  $\psi : E \rightarrow E$  is a linear transformation with  $\dim E = n$ , then there exist linear transformations  $\psi_i^{(n)} : \wedge^n E \rightarrow \wedge^n E$  unique such that

$$(\psi - \lambda \iota)_{\wedge}^{(n)} = \sum_{i=0}^n \psi_i^{(n)} \lambda^{n-i} \quad \lambda \in \Gamma$$

Moreover,  $\text{tr } \psi_i^{(n)} = \alpha_i$  where  $\alpha_i$  is the  $i$ -th characteristic coefficient of  $\psi$ .

*Proof.* By an argument similar to that given in subsection 4.19 of [1],

$$(\psi - \lambda \iota)_{\wedge}^{(n)} = \sum_{i=0}^n \tilde{S}_i$$

where

$$\tilde{S}_i(x_1 \wedge \cdots \wedge x_n) = \frac{(-\lambda)^{n-i}}{i!(n-i)!} \sum_{\sigma} \varepsilon_{\sigma} \varphi x_{\sigma(1)} \wedge \cdots \wedge \varphi x_{\sigma(i)} \wedge x_{\sigma(i+1)} \wedge \cdots \wedge x_{\sigma(n)}$$

This uniquely determines  $\psi_i^{(n)}$  with  $\tilde{S}_i = \psi_i^{(n)} \lambda^{n-i}$ . By (4.47) and (4.48) of [1],

$$\Delta_{\wedge} \circ \tilde{S}_i = \alpha_i \lambda^{n-i} \Delta_{\wedge} = \Delta_{\wedge} \circ (\alpha_i \lambda^{n-i} \iota)$$

where  $\Delta \neq 0$  is a determinant function in  $E$ . Since  $\Delta_{\wedge}$  is injective, this implies

$$\psi_i^{(n)} \lambda^{n-i} = \tilde{S}_i = \alpha_i \lambda^{n-i} \iota$$

so  $\psi_i^{(n)} = \alpha_i \iota$  and  $\text{tr} \psi_i^{(n)} = \alpha_i$ . □

*Remark.* We have  $\psi_0^{(n)} = (-1)^n \iota$ ,  $\psi_1^{(n)} = (-1)^{n-1} \theta_{\wedge}^{(n)}(\psi)$ , and  $\psi_n^{(n)} = \psi_{\wedge}^{(n)}$ .

## § 14

**Exercise (1).** If  $u^* \in \wedge^p E^*$  and  $i(a)u^* = 0$  for all  $a \in \wedge^k E$  for some  $k \leq p$ , then  $u^* = 0$ .

*Proof.* We claim  $\langle u^*, v \rangle = 0$  for all  $v \in \wedge^p E$ , from which it follows that  $u^* = 0$ . Since  $k \leq p$ , we may write

$$v = \sum_{\nu} a_{\nu} \wedge v_{\nu} \quad a_{\nu} \in \wedge^k E, v_{\nu} \in \wedge^{p-k} E$$

Then

$$\langle u^*, v \rangle = \sum_{\nu} \langle u^*, a_{\nu} \wedge v_{\nu} \rangle = \sum_{\nu} \langle i(a_{\nu})u^*, v_{\nu} \rangle = 0 \quad \square$$

*Remark.* This result generalizes Proposition 5.14.2.

**Exercise (2).** Let  $E, E^*$  be a pair of dual  $n$ -dimensional vector spaces and  $e_{\nu}, e^{*\nu}$  a pair of dual bases. If  $\varphi : E \rightarrow E$  is linear, then

$$\theta_{\wedge}(\varphi) = \sum_{\nu} \mu(\varphi e_{\nu}) i(e^{*\nu}) \quad (1)$$

$$\theta^{\wedge}(\varphi) = \sum_{\nu} \mu(e^{*\nu}) i(\varphi e_{\nu}) \quad (2)$$

*Proof.* By Corollary II to Proposition 5.14.1,

$$i(e^{*v})(x_1 \wedge \cdots \wedge x_p) = \sum_{\mu=1}^p (-1)^{\mu-1} \langle e^{*v}, x_\mu \rangle x_1 \wedge \cdots \wedge \widehat{x_\mu} \wedge \cdots \wedge x_p$$

It follows that

$$\mu(\varphi e_v) i(e^{*v})(x_1 \wedge \cdots \wedge x_p) = \sum_{\mu=1}^p x_1 \wedge \cdots \wedge \varphi(\langle e^{*v}, x_\mu \rangle e_v) \wedge \cdots \wedge x_p$$

and so

$$\begin{aligned} \sum_{v=1}^n \mu(\varphi e_v) i(e^{*v})(x_1 \wedge \cdots \wedge x_p) &= \sum_{\mu=1}^p x_1 \wedge \cdots \wedge \varphi\left(\sum_{v=1}^n \langle e^{*v}, x_\mu \rangle e_v\right) \wedge \cdots \wedge x_p \\ &= \sum_{\mu=1}^p x_1 \wedge \cdots \wedge \varphi x_\mu \wedge \cdots \wedge x_p \\ &= \theta_\wedge(\varphi)(x_1 \wedge \cdots \wedge x_p) \end{aligned}$$

This establishes (1). Now (2) follows from (1) by duality since

$$\theta^\wedge(\varphi) = \theta_\wedge(\varphi)^* = \sum_{v=1}^n i(e^{*v})^* \mu(\varphi e_v)^* = \sum_{v=1}^n \mu(e^{*v}) i(\varphi e_v) \quad \square$$

**Exercise (3).** Let  $E, E^*$  be a pair of dual  $n$ -dimensional vector spaces and  $e_v, e^{*v}$  a pair of dual bases.

(a) For  $x \in E$  and  $x^* \in E^*$ ,

$$i(x^*)\mu(x) + \mu(x)i(x^*) = \langle x^*, x \rangle \iota$$

(b) For  $u \in \wedge^p E$ ,

$$\sum_v \mu(e_v) i(e^{*v}) u = pu$$

(c) For  $u \in \wedge^p E$ ,

$$\sum_v i(e^{*v}) \mu(e_v) u = (n-p)u$$

*Proof.* For (a), dualize Corollary I to Proposition 5.14.1; for (b), apply problem 2 with  $\varphi = \iota$ ; for (c), apply (a) and (b) to obtain

$$\sum_{v=1}^n i(e^{*v}) \mu(e_v) u = \sum_{v=1}^n \langle e^{*v}, e_v \rangle u - \sum_{v=1}^n \mu(e_v) i(e^{*v}) u = nu - pu \quad \square$$

**Exercise (6).** If  $a \in \wedge^p E$  and  $p \leq q$ , then

$$i(a)(x_1^* \wedge \cdots \wedge x_q^*) = \sum_{v_1 < \cdots < v_p} (-1)^{\sum_{i=1}^p (v_i - i)} \langle x_{v_1}^* \wedge \cdots \wedge x_{v_p}^*, a \rangle x_{v_{p+1}}^* \wedge \cdots \wedge x_{v_q}^*$$

where  $(v_{p+1}, \dots, v_q)$  is the ordered tuple complementary to  $(v_1, \dots, v_p)$ .

*Proof.* By induction on  $p$ . For  $p = 0$  the result is trivial, and for  $p = 1$  it is just Corollary II to Proposition 5.14.1. We illustrate the induction step for  $p = 2$ , where we may assume that  $a = a_1 \wedge a_2$  and write  $u^* = x_1^* \wedge \cdots \wedge x_q^*$ :

$$\begin{aligned} i(a)u^* &= i(a_2)i(a_1)(x_1^* \wedge \cdots \wedge x_q^*) \\ &= i(a_2) \sum_{v_1} (-1)^{v_1-1} \langle x_{v_1}^*, a_1 \rangle x_1^* \wedge \cdots \wedge \widehat{x_{v_1}^*} \wedge \cdots \wedge x_q^* \\ &= \sum_{v_1} (-1)^{v_1-1} \langle x_{v_1}^*, a_1 \rangle i(a_2)(x_1^* \wedge \cdots \wedge \widehat{x_{v_1}^*} \wedge \cdots \wedge x_q^*) \\ &= \sum_{v_1 < v_2} (-1)^{(v_1-1)+(v_2-2)} C_{v_1, v_2} x_1^* \wedge \cdots \wedge \widehat{x_{v_1}^*} \wedge \cdots \wedge \widehat{x_{v_2}^*} \wedge \cdots \wedge x_q^* \end{aligned}$$

where

$$C_{v_1, v_2} = \sum_{\sigma} \varepsilon_{\sigma} \langle x_{v_{\sigma(1)}}^*, a_1 \rangle \langle x_{v_{\sigma(2)}}^*, a_2 \rangle = \det(\langle x_{v_i}^*, a_j \rangle) = \langle x_{v_1}^* \wedge x_{v_2}^*, a_1 \wedge a_2 \rangle \quad \square$$

## § 29

*Remark.* Recall that if  $E, E^*$  is a pair of finite-dimensional dual vector spaces, then there is a natural isomorphism  $\alpha : \otimes^p E^* \rightarrow T^p(E)$ .<sup>37</sup> Since

$$\langle x^{*1}, x_{\sigma(1)} \rangle \cdots \langle x^{*p}, x_{\sigma(p)} \rangle = \langle x^{*\sigma^{-1}(1)}, x_1 \rangle \cdots \langle x^{*\sigma^{-1}(p)}, x_p \rangle$$

for all  $x_i \in E$ ,  $x^{*i} \in E^*$ , and  $\sigma \in S_p$ , the following diagram commutes:

$$\begin{array}{ccc} \otimes^p E^* & \xrightarrow{\sigma} & \otimes^p E^* \\ \alpha \downarrow & & \downarrow \alpha \\ T^p(E) & \xrightarrow{\sigma} & T^p(E) \end{array}$$

<sup>37</sup>See subsection 3.20.

It follows that the alternator diagram also commutes:

$$\begin{array}{ccc}
 \otimes^p E^* & \xrightarrow{\pi^A} & \otimes^p E^* \\
 \alpha \downarrow & & \downarrow \alpha \\
 T^p(E) & \xrightarrow{A} & T^p(E)
 \end{array}$$

Therefore (5.81) follows from (4.11).

### § 32

*Remark.* In this subsection, the substitution operator  $i_A(h)$  should be redefined on  $T^p(E)$  as

$$i_A(h) = \frac{1}{p} \sum_{v=1}^p (-1)^{v-1} i_v(h)$$

Then for  $\Phi \in A^p(E)$ ,

$$\begin{aligned}
 p \cdot (i_A(h)\Phi)(x_1, \dots, x_{p-1}) &= \sum_{v=1}^p (-1)^{v-1} (i_v(h)\Phi)(x_1, \dots, x_{p-1}) \\
 &= \sum_{v=1}^p (-1)^{v-1} \Phi(x_1, \dots, x_{v-1}, h, x_{v+1}, \dots, x_{p-1}) \\
 &= \sum_{v=1}^p \Phi(h, x_1, \dots, x_{p-1}) \\
 &= p \cdot \Phi(h, x_1, \dots, x_{p-1}) \\
 &= p \cdot (i_1(h)\Phi)(x_1, \dots, x_{p-1})
 \end{aligned}$$

so

$$i_A(h)\Phi = i_1(h)\Phi$$

as desired. This redefinition is required to make  $i_A(h)$  an antiderivation in the algebra of skew-symmetric multilinear functions (5.32.1), although under the redefinition  $i_A(h)$  is no longer an antiderivation in the algebra of multilinear functions.<sup>38</sup>

---

<sup>38</sup>See subsection 3.19.

### § 33

*Remark.* The algebra isomorphism  $\beta \circ \eta : \wedge E^* \rightarrow A^\bullet(E)$  satisfies

$$f_1^* \wedge \cdots \wedge f_p^* \mapsto f_1^* \wedge \cdots \wedge f_p^* = p! A(f_1^* \bullet \cdots \bullet f_p^*) \quad f_i^* \in E^*$$

The product on the left is the exterior product and the product on the right is the Grassmann product, where  $f_i^*$  has been identified with  $\langle f_i^*, - \rangle$ .

### § 34

*Remark.* We see that the algebra isomorphisms  $\wedge E^* \cong A^\bullet(E)$  and  $\wedge E \cong A_\bullet(E)$  preserve scalar products.

## References

- [1] Greub, W. *Linear Algebra*, 4th ed. Springer, 1975.
- [2] Greub, W. *Multilinear Algebra*, 2nd ed. Springer, 1978.