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, Γ 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

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

The inclusion map $i_X : X \to C(X)$ 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 $f : X \to F$, there is a unique *linear* $\varphi : C(X) \to F$ "extending f" in the sense that $\varphi \circ i_X = f$:

$$\begin{tikzcd} X \arrow[r, "i_X"] \arrow[dr, "f"'] & C(X) \arrow[d, dashed, "\varphi"] \\ & F \end{tikzcd}$$

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

$$\begin{tikzcd} X \arrow[r, "\alpha"] \arrow[d, "i_X"'] & Y \arrow[d, "i_Y"] \\ C(X) \arrow[r, dashed, "\alpha_*"'] & C(Y) \end{tikzcd}$$

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

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

$$\begin{tikzcd} E \arrow[r, "i_E"] \arrow[dr, "\iota_E"'] & C(E) \arrow[d, dashed, "\pi_E"] \\ & E \end{tikzcd}$$

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

$$\begin{tikzcd} E \arrow[r, "\varphi"] & F \\ C(E) \arrow[u, "\pi_E"] \arrow[r, "\varphi_*"'] & C(F) \arrow[u, "\pi_F"'] \end{tikzcd}$$

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

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

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:

$$\begin{tikzcd} X \arrow[r, "\alpha"] \arrow[d, "i_X"'] & Y \arrow[r, "\beta"] \arrow[d, "i_Y"] & Z \arrow[d, "i_Z"] \\ C(X) \arrow[r, "\alpha_*"'] & C(Y) \arrow[r, "\beta_*"'] & C(Z) \end{tikzcd}$$

(iii) By (i), applied to ι_E.

(iv) If φ is linear, then $\varphi \circ \pi_E : C(E) \to F$ is linear and extends φ in the sense that $\varphi \circ \pi_E \circ i_E = \varphi \circ \iota_E = \varphi$. However, $\pi_F \circ \varphi_* : C(E) \to F$ is also linear and extends φ 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 φ is linear since $\pi_F \circ \varphi_*$ is linear and π_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 $X \to F$, for vector spaces F, and the arrows are *linear* (i.e. structure-preserving) maps $F \to G$ between the vector spaces which respect the mappings of X:

$$\begin{tikzcd} X \arrow[r] \arrow[dr] & F \arrow[d] \\ & G \end{tikzcd}$$

By (i), every object $X \to F$ in this category can be obtained from the inclusion map $X \to C(X)$ 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 $X \to U$ is also universal, then the composites $\psi \circ \varphi$ and $\varphi \circ \psi$ of the induced linear maps $\varphi : C(X) \to U$ and $\psi : U \to C(X)$ are linear and extend the inclusion maps, so must be the identity maps on $C(X)$ and U by uniqueness; that is, φ and ψ 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 $\alpha : X \to Y$ to the linear map $\alpha_* : C(X) \to C(Y)$. 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 $\langle 1,1\rangle = \langle 1,0\rangle + \langle 0,1\rangle$ in E, but *not* in $C(E)$. The "formal" linear combination

$$\langle 1,1\rangle - \langle 1,0\rangle - \langle 0,1\rangle$$

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

In (v), we see that π_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.

§ 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

$$x(t) = (1-t)x + ty \qquad (0 \le t \le 1)$$

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 $\alpha : E \to \mathbb{R}$ by $x - \alpha(x)e \in E_1$ for all $x \in E$. It is clear that α 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 $\varphi_1^{-1} : E \leftarrow \operatorname{Im}\varphi$.[1]

§ 4

Remark. The direct sum $E \oplus F$ is a coproduct in the category of vector spaces in the following sense: if $\varphi : E \to G$ and $\psi : F \to G$ are linear maps, there is a unique linear map $\chi : E \oplus F \to G$ such that $\varphi = \chi \circ i_E$ and $\psi = \chi \circ i_F$, where i_E and i_F are the canonical injections:

$$\begin{tikzcd}
E \arrow[r, "i_E"] \arrow[dr, "\varphi"'] & E \oplus F \arrow[d, dashed, "\chi"'] & F \arrow[l, "i_F"'] \arrow[dl, "\psi"] \\
& G &
\end{tikzcd}$$

[1] See Corollary I to Proposition I in subsection 1.15.

Indeed, χ 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 φ and ψ. This property makes $E \oplus F$ unique up to a unique isomorphism.

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

$$\begin{tikzcd} & G \arrow[dl, "\varphi"'] \arrow[d, dashed, "\chi"'] \arrow[dr, "\psi"] & \\ E & E \oplus F \arrow[l, "\pi_E"'] \arrow[r, "\pi_F"] & F \end{tikzcd}$$

Indeed, χ is given by $\chi(x) = \varphi(x) + \psi(x)$, and "combines" φ and ψ. 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, σ is the product map and τ is the coproduct map. If $\varphi_1 : E_1 \to F_1$ and $\varphi_2 : E_2 \to F_2$ are linear maps, then $\varphi = \varphi_1 \oplus \varphi_2$ is both a coproduct and product map:

$$\begin{tikzcd} E_1 \arrow[r, shift left] \arrow[d, "\varphi_1"'] & E_1 \oplus E_2 \arrow[l, shift left] \arrow[r, shift left] \arrow[d, "\varphi"] & E_2 \arrow[l, shift left] \arrow[d, "\varphi_2"] \\ F_1 \arrow[r, shift left] & F_1 \oplus F_2 \arrow[l, shift left] \arrow[r, shift left] & F_2 \arrow[l, shift left] \end{tikzcd}$$

The structure of φ is completely determined by the structures of φ_1 and φ_2. In particular, φ is injective (surjective, bijective) if and only if φ_1 and φ_2 are.

§ 5

Remark. The definition of dual space is fundamentally *symmetrical* between E and E^*, as is the definition of dual mapping between φ and φ^*. 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 φ and φ^*. 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 section), at the expense of this symmetry.

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 $\varphi : E \to F$ is linear, then $\varphi^* : E^* \leftarrow F^*$ is dual to φ if and only if the following diagram commutes:

$$\begin{tikzcd}
F^* \times E \arrow[r, "\varphi^* \times \iota_E"] \arrow[d, "\iota_{F^*} \times \varphi"'] & E^* \times E \arrow[d, "\langle\, ,\, \rangle"] \\
F^* \times F \arrow[r, "\langle\, ,\, \rangle"'] & \Gamma
\end{tikzcd}$$

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_α is called the *α-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 α, 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 $i_k : E_k \to E$ is the k-th canonical injection and $\pi_k : E \to E_k$ 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 $F_k^\perp \cap E_k^\perp = 0$.[2] The induced injection $F_k^\perp \to L(E_k)$ is surjective since every linear function on E_k can be extended to a linear function on E which kills F_k.

Remark. For $\varphi : E \to F$ 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 $\varphi : E \to F$ is a linear map with restriction $\varphi_1 : E_1 \to F_1$ and dual map $\varphi^* : E^* \leftarrow F^*$, then φ^* 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 φ_1.

[2] 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 φ^* 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 $x \in E_1$,

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

□

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

Alternatively, with $E^* = L(E)$, let $f_1^*, \ldots, f_n^*$ be the dual basis of $f_1, \ldots, f_n$ in $E^{**} = L(L(E))$, so $\langle f_j^*, f_i \rangle = \delta_{ij}$ by (2.62). Let $e_1, \ldots, e_n \in E$ be defined by $\langle f_j^*, f \rangle = \langle f, e_j \rangle$ for all $f \in E^*$ (see § 5, problem 3). Then $\langle f_i, e_j \rangle = \langle f_j^*, f_i \rangle = \delta_{ij}$, so $f_1, \ldots, f_n$ is the dual basis of $e_1, \ldots, e_n$.

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

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

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

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 function notation (writing φx instead of $x\varphi$, and

$\varphi\psi$ to mean φ *after* ψ), 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 with left-hand notation. This has the following undesirable 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 also has the annoying habit of indexing 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.

§ 3

Remark. In subsection 3.13, note $(\alpha_\nu^\mu) = M(\iota; \bar{x}_\nu, x_\mu)$ by (3.22), $(\check{\alpha}_\nu^\mu) = M(\iota; x_\nu, \bar{x}_\mu)$ by (3.23), and $(\beta_\sigma^\varrho) = M(\iota; \bar{x}^{*\varrho}, x^{*\sigma})$ by (3.24). It follows from (3.4) that

$$\left(\beta_\nu^\mu\right) = \left(\check{\alpha}_\nu^\mu\right)^* = \left(\left(\alpha_\nu^\mu\right)^{-1}\right)^*$$

In other words, the matrix of the dual basis transformation $x^{*\nu} \mapsto \bar{x}^{*\nu}$ in E^* is the *transpose* of the inverse of the matrix of the basis transformation $x_\nu \mapsto \bar{x}_\nu$ in E, contrary to what the book says. It's easier to remember that the matrix of $x^{*\nu} \leftarrow\!\shortmid \bar{x}^{*\nu}$ (arrow reversed!) is the transpose of the matrix of $x_\nu \mapsto \bar{x}_\nu$.

Remark. In subsection 3.13, we see that if a basis transformation is effected by τ, then the corresponding *coordinate* transformation is effected by τ^{-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, which introduce transposition into matrix-vector multiplication (see the remarks above).

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 $\Phi : X^0 \to Y$ can be identified with elements of Y.

- For vector spaces E and F, a map $\Phi : E^0 \to F$ is vacuously 0-linear. Since the only permutation in S_0 is the identity $\iota = \emptyset$, Φ is also trivially skew symmetric.

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

- Determinant functions in E are just scalars in Γ, 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.

- If E is real ($\Gamma = \mathbb{R}$), the orientations in E are represented by the scalars ± 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

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

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

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

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 $\Delta(x_1, \ldots, x_n) \neq 0$ if and only if $x_1, \ldots, x_n$ are linearly independent. By (4.8), it follows that $\det\varphi$ "determines" whether a linear transformation φ preserves linear independence, i.e. whether or not φ is invertible.

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

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

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

where x is in the j-th position on the right.[3] Therefore

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

Viewing this as a function of $x_1, \ldots, x_n$ (that is, a set map from $E^n \to L(E;E)$), it is obviously multilinear and skew symmetric (by Proposition I(ii)). Therefore if $x_1, \ldots, x_n$ are linearly dependent, it is zero (by Proposition I(iii)). If $x_1, \ldots, x_n$ are linearly independent (and hence a basis), then viewing it as a function of x, its value at x_i is just $\Delta(x_1, \ldots, x_n)x_i$ (by Proposition I(ii)), so it agrees on a basis with $\Delta(x_1, \ldots, x_n)x$ 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 Δ be a determinant function in E with $\Delta(e_1, \ldots, e_n) = 1$ where $e_1 \in E_1$. Then Δ induces a determinant function Δ_1 in E/E_1 by

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

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

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

[3] $\widehat{x_j}$ denotes deletion of x_j from the sequence on the left.

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

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

This result 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 to recursively *define* an n-dimensional determinant function in terms of an $(n-1)$-dimensional one.

The geometrical 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 closely related to this result.

§ 2

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

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

To obtain this, we can define

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

where x is in the j-th position on the right.[4] Then ψ obviously satisfies the above properties, by multilinearity and skew symmetry of Δ.

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

[4]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 $\varphi : E\to E$ be linear with E_1 a stable subspace. If $\varphi_1 : E_1\to E_1$ and $\overline{\varphi} : E/E_1\to E/E_1$ 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

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

is a determinant function in E_1 and

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

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

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

It follows from (2) and (3) that

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

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. □

§ 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

$$\det A = \det A_1 \cdot \det A_2 \tag{1}$$

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

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

If $\varphi_2^* : E_2^* \to E_2^*$ 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 Δ^* its dual in E^*. We claim

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

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,[5] it is equal to

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

[5] See subsection 4.9, item 4.

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

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

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 $\varphi : \Gamma^n \to \Gamma^n$ is defined by $M(\varphi) = (\alpha_k^j) = A$, $x = (\xi^i)$, and $y = (\eta^j)$. If $\det A \neq 0$, then φ is invertible and

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

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

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

Moreover, it follows from (4.38) that $\sum_j \operatorname{cof}(\alpha_i^j)\eta^j = \det A_i$ where A_i is the matrix obtained from A by replacing the i-th row with y.[6] 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.

[6] 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 $\varphi_1 : E_1 \to E_1$ and $\varphi_2 : E_2 \to E_2$ are linear, then

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

where $\chi(\varphi)$ denotes the characteristic polynomial of φ.

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

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

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

$$\chi(\varphi) = \chi(\varphi_1) \cdot \chi(\overline{\varphi})$$

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 $\overline{\varphi - \lambda\iota_E} = \overline{\varphi} - \lambda\iota_{E/E_1}$. $\square$

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

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

Proof. If φ 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 $\overline{\varphi} : E/E_1 \to E/E_1$ the induced map. Then $\overline{\varphi}^{k-1} = 0$ since

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

as $\varphi^{k-1}(x) \in E_1$. By the induction hypothesis, $\chi(\overline{\varphi}) = (-\lambda)^{n-p}$ 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 $(-\lambda)^n = (-\lambda)^p \chi(\overline{\varphi})$, which implies $\chi(\overline{\varphi}) = (-\lambda)^{n-p}$. By induction, $\overline{\varphi}$ is nilpotent. If $\overline{\varphi}^k = 0$, then $\varphi^{k+1} = 0$, so φ is nilpotent. $\square$

§ 7

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

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

Proof. Immediate since

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

□

§ 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 Δ_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 $\varphi : A \to B$ 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 $\varphi_1 : A_1 \to B_1$ of φ to A_1, B_1 is a homomorphism.

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

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

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

Exercise (12)**.** The mapping

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

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

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

Exercise (16)**.** Every algebra automorphism $\Phi : A(E;E) \to A(E;E)$ 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 (φ_{ij}) of $A(E;E)$ defined by $\varphi_{ij}(e_k) = \delta_{jk}e_i$. This basis satisfies

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

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 (φ_{ij}) as above. Since Φ 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 $\varphi \in A(E;E)$. □

Remark. The result is true for any nonzero endomorphism Φ, 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.[7]

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 $\varphi : E \to F$ is linear and homogeneous of degree k and φx 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 $\varphi : E \to E$ is linear and homogeneous with $\deg\varphi \neq 0$, then $\operatorname{tr}\varphi = 0$.

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 $1 \le i \le n$, (e_{ij}) is a basis of E_{k_i} for $1 \le j \le d_i$. Let Δ be a determinant function in E with $\Delta(e_{ij}) = 1$. Then

$$\operatorname{tr}\varphi = \sum_{i,j} \Delta(e_{11}, \dots, e_{1d_1}, \dots, \varphi(e_{ij}), \dots, e_{n1}, \dots, e_{nd_n})$$

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 $\varphi : E \to F$ and $\varphi^* : E^* \leftarrow F^*$ be dual linear maps. If φ is homogeneous of degree k, then φ^* is homogeneous of degree $-k$.

[7] See problem 2 in Chapter II, § 4.

Proof. We have direct sum decompositions

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

and

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

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 $\pi_{n-k} : E \to E_{n-k}$ 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. □

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

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

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

$$F^{\perp\perp} = \sum_{m\in G} F^{\perp\perp} \cap E_m \tag{2}$$

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

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

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 $x^* \in (F \cap E_n)^{\perp_n} \subseteq E^{*n}$. 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 $F^{\perp\perp} = F$. □

§ 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 $k = 0$. □

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. □

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

$$\Phi : A_G(E; E) \to A_G(E^*; E^*)^{\mathrm{opp}}$$

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

Proof. Φ is well defined by problems 6 and 10 of § 1, and is an isomorphism by problem 12 of chapter V, § 1. □

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

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

where x, y are the unique representatives of $\overline{x}, \overline{y}$ 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. □

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 π is the canonical projection onto $E_1^\perp$. Then $y \in E_1$ 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. □

§ 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 $\overline{e_2}, \dots, \overline{e_n}$ 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

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

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 $\varphi : E \to E$ is linear, then the dual transformation $\varphi^* : E \to E$ satisfies

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

If φ preserves inner products, then it also preserves orthonormal bases and is invertible. In this case, φ^{-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 φ preserves inner products.

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

§ 3

Remark. A determinant function Δ 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 Δ 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 Δ_0^* is its dual, then $\Delta_0^* = \alpha\Delta_0$ for some real number α by uniqueness of Δ_0, so

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

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 $x_1,\dots,x_p$,

$$G(x_1,\dots,x_p) \leq |x_1|^2 \cdots |x_p|^2 \tag{1}$$

Additionally

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}^2 \leq \sum_{k=1}^{n} |a_{1k}|^2 \cdots \sum_{k=1}^{n} |a_{nk}|^2 \tag{2}$$

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. □

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*.

§ 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

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

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

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

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_ν) and (y_μ) are orthonormal, then they are self-dual,[8] so $M(\widetilde{\varphi}, y_\mu, x_\nu) = M(\varphi, x_\nu, y_\mu)^*$ by (3.4). It follows that

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

Remark. Recall from subsection 7.18 the natural isomorphism $\tau : E \to L(E)$ given by $x \mapsto (x, -)$, which maps each vector to its dual vector. In subsection 8.4, observe that there is a natural isomorphism

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

given by $\Phi \mapsto (x \mapsto (y \mapsto \Phi(x,y)))$. For a linear transformation $\varphi : E \to E$ and the bilinear function Φ defined by $\Phi(x,y) = (\varphi x, y)$, the linear map corresponding to Φ under this isomorphism is $\tau\varphi$:

$$\begin{tikzcd}
x \arrow[r, "\varphi"] \arrow[dr, "\Phi"'] & \varphi x \arrow[d, "\tau"] \\
& (\varphi x, -)
\end{tikzcd}$$

[8] See problem 2 in Chapter VII, § 2.

In this sense, Φ is naturally dual to φ. Properties of Φ naturally correspond to those of φ; for example, Φ is symmetric if and only if φ is symmetric (self-adjoint), and Φ is skew symmetric if and only if φ is skew symmetric. Since

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

we see that φ is "left-dual" to Φ, while $\widetilde{\varphi}$ is "right-dual" to Φ. It might be said that Φ, φ, and $\widetilde{\varphi}$ enter into a holy trinity.[9]

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

$$\widetilde{\varphi}_i = \widetilde{\varphi_i}$$

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

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

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

$$\varphi + \widetilde{\varphi} \qquad \varphi\widetilde{\varphi} \qquad \widetilde{\varphi}\varphi$$

are self-adjoint.

Remark. If $E = E_1 \oplus E_2$ with $E_1 \perp E_2$, and $\varphi_1 : E_1 \to E_1$ and $\varphi_2 : E_2 \to E_2$ 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 $x_i, y_i \in E_i$,

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

□

[9] No one says this.

Remark. In subsection 8.6, consider $E = \mathbb{R}^n$. Since e_1 minimizes the quadratic form $q(x) = (x, \varphi x)$ subject to the constraint $h(x) = 1 - (x, x) = 0$, there exists a Lagrange multiplier $\lambda \in \mathbb{R}$ such that

$$\nabla q(e_1) + \lambda \nabla h(e_1) = 0$$

Now $\nabla h(e_1) = -2e_1$, and since φ is self-adjoint it follows that $\nabla q(e_1) = 2\varphi e_1$. Therefore $\varphi e_1 = \lambda e_1$, so e_1 is an eigenvector of φ with eigenvalue λ.

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

Proof. There is an orthonormal basis $e_1, \ldots, e_n$ of eigenvectors of φ, 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 ψ is positive with $\psi^2 = \varphi$.

If ψ_1 is positive with $\psi_1^2 = \varphi$, then the eigenvalues of ψ_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 $\psi_1 = \psi$. □

§ 3

Remark. In subsection 2.19, we see that a linear transformation π 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 π is additionally an *orthogonal* projection if and only if $\ker \pi \perp \operatorname{Im} \pi$. In summary, for a projection operator π, the following are equivalent:

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

§ 4

Remark. Geometrically, a skew transformation, apart from possibly killing off part of the space, induces scaled 90-degree rotations[10] on orthogonal stable planes in the space.

[10] No one calls these "scrotations".

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

$$\varphi = \tfrac{1}{2}(\varphi + \widetilde{\varphi}) + \tfrac{1}{2}(\varphi - \widetilde{\varphi})$$

uniquely represents φ as a sum of self-adjoint and skew transformations.

Remark. If φ is a skew transformation of E and F is 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 ψ by

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

where e_i is the i-th standard basis vector, then ψ 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 $\mathbb{R}^4$.

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)

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

where E_i is the eigenspace of φ corresponding to λ_i. Observe that E_i is stable under ψ, 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 λ_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 ψ. 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 ψ by the remark above). If $\lambda_i = 0$, assume that $i = r$ and choose any orthonormal basis of E_r, noting that ψ 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 ψ has the form (8.35).[11]

Exercise (1)**.** If φ 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 φ 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.

(ii) For $a, b \in E$, $\varphi_{a \times b} = \varphi_a \varphi_b - \varphi_b \varphi_a$.

(iii) If φ 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 φ_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 $x \in E$.

[11] See `https://math.stackexchange.com/q/3402347`.

(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 Φ is clearly trilinear and skew symmetric. By the universal property of determinants if Δ 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 $a = \Phi(e_1, e_2, e_3)$.[12] By direct computation it is easily verified that $a \times e_i = \varphi e_i$, so $\varphi = \varphi_a$.

(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 Φ 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 φ 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 Δ is the normed determinant function representing the orientation in E. In other words, Δ_1 is just the 2-dimensional area function induced by the 3-dimensional volume function Δ and the vector a. Therefore if we *define* a 3-dimensional volume map Φ in terms of Δ_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.

[12] See Proposition III in subsection 4.3.

Exercise (3)**.** If $\varphi \neq 0$ and ψ 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

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

Then $\varphi \neq 0$ is skew, and ψ is skew since $(x, \psi x) = 0$ for all x. The kernel of φ is the line determined by $a_1 \times a_2$, which is killed by ψ. 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 φ satisfies $\tilde{\varphi} = \lambda\varphi$ for $\lambda \in \mathbb{R}$ if and only if φ 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 φ is self-adjoint or skew, respectively. The rest is obvious. □

Exercise (6)**.** If Φ 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 φ dual to Φ 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. □

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 φ 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 τ 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. □

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 $\varphi : E \to E$ is a mapping such that $\varphi(0) = 0$ and

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

for all $x, y \in E$, then φ 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

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

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 Θ 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 φx. Fix an orientation in E so that Θ is the *oriented* angle between x and φx. Then (8.40) holds, so $\cos\Theta = \frac{1}{2}\operatorname{tr}\varphi$. Clearly Θ is independent of x and depends only on φ, 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 ψ is skew, $\Psi(x,y) = (\psi x, y)$ measures oriented area in planes in E, and since ψ 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 Ψ actually measures oriented *volume* relative to u, so it follows that u must lie in E_1.

Note that ψ stabilizes $F = E_1^{\perp}$, and if ψ_1 denotes the restriction, then

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

where φ_1 denotes the restriction of φ. 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 Θ of rotation in that plane is the same. It follows that

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

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

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

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 Θ 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 $\varphi : E \to F$ of complex vector spaces E, F is *conjugate-linear* if

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

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

§ 1

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

§ 2

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

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

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

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

§ 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 $z \mapsto \overline{z}$ 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 $x_j, y_j \in F$,

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

and

$$(\overline{z_1}, \overline{z_2}) = (x_1, x_2) + (y_1, y_2) - i\left[(y_1, x_2) - (x_1, y_2)\right]$$

so $(\overline{z_1}, \overline{z_2}) = \overline{(z_1, z_2)}$ as required.[13]

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

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

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

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

where $\kappa_F : F \to \overline{F}$ is the identity map.

If $\varphi : E \to F$ is a linear map, then the dual map $\varphi^* : \overline{E} \leftarrow \overline{F}$ 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 $\widetilde{\varphi} = \kappa_E^{-1} \varphi^* \kappa_F$:

$$\begin{tikzcd}
E \arrow[r, "\varphi"] \arrow[d, "\kappa_E"'] & F \arrow[d, "\kappa_F"] \\
\overline{E} & \overline{F} \arrow[l, "\varphi^*"]
\end{tikzcd}$$

[13] See problem 2(i) in § 2.

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

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

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

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

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

Remark. If $z \mapsto \overline{z}$ is a conjugation in E and Δ is a determinant function in E, then

$$\overline{\Delta}(z_1,\dots,z_n) = \overline{\Delta(\overline{z_1},\dots,\overline{z_n})}$$

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 $\varphi : E \to E$ 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 $\widetilde{\varphi}$ represents the adjoint $\widetilde{\Phi}$.

Proof. The forward direction is trivial. If Φ 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 φx such that

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

and therefore

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

Clearly φ is linear and is uniquely determined by Φ. Also

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

Remark. This result shows that the linear transformation φ is naturally left-dual to the sesquilinear function Φ. Observe that Φ is Hermitian if and only if φ is Hermitian (self-adjoint) and Φ is skew-Hermitian if and only if φ is skew-Hermitian.

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

$$\lambda^1 = \xi^1\eta^1 \qquad \lambda^2 = \xi^1\eta^2 \qquad \lambda^3 = \xi^2\eta^1 \qquad \lambda^4 = \xi^2\eta^2$$

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 $\widetilde{\varphi}$, note that for each $y \in F$ the linear map $\varphi(-, y) : E \to G$ sends E_1 into G_1 and hence induces a linear map $\overline{\varphi}(-, y) : E/E_1 \to G/G_1$ by $\overline{\varphi}(\rho x, y) = \pi\varphi(x, y)$:

$$\begin{tikzcd}
E \arrow[r, "\varphi(-,y)"] \arrow[d, "\rho"'] & G \arrow[d, "\pi"] \\
E/E_1 \arrow[r, "\overline{\varphi}(-,y)"'] & G/G_1
\end{tikzcd}$$

Since $\overline{\varphi}(-, y)$ depends linearly on y, $\overline{\varphi}$ is bilinear and we define $\widetilde{\varphi} = \overline{\varphi}$.

In the construction of the second induced bilinear map $\widetilde{\varphi}$, note that the linear map $y \mapsto \overline{\varphi}(-, y)$ kills F_1, so it factors through σ:

$$\begin{tikzcd}
F \arrow[r, "y \mapsto \overline{\varphi}(-,y)"] \arrow[d, "\sigma"'] & L(E/E_1; G/G_1) \\
F/F_1 \arrow[ur, "\sigma y \mapsto \overline{\varphi}(-,y)"'] &
\end{tikzcd}$$

This allows us to define the bilinear map $\widetilde{\varphi}(\rho x, \sigma y) = \overline{\varphi}(\rho x, y) = \pi\varphi(x, y)$.

§ 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 ψ preserves linear relations satisfied by φ.

§ 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 $E \times F \to G$, and the arrows are linear maps $G \to H$ which respect the bilinear maps:

$$\begin{tikzcd} E \times F \arrow[r] \arrow[d] & H \\ G \arrow[ur] & \end{tikzcd}$$

Every object $E \times F \to G$ in this category can be obtained from the tensor product $\otimes : E \times F \to E \otimes F$ 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 $E \times F \to E$ defined by $(x, y) \mapsto g(y)x$. By the universal property of the tensor product, there is a linear map $h : E \otimes F \to E$ with $h(x \otimes y) = g(y)x$. Now

$$0 = h(\sum a_i \otimes b_i) = \sum h(a_i \otimes b_i) = \sum g(b_i)a_i$$

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 $h = \iota \otimes g : E \otimes F \to E \otimes \Gamma$ (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 $g : G \to E \otimes F$ with $g \circ f = \iota$. If $\psi : E \times F \to K$ is bilinear, there is a linear map $h : E \otimes F \to K$ with

$$\psi = h \circ \otimes = h \circ \iota \circ \otimes = h \circ g \circ f \circ \otimes = h \circ g \circ \varphi$$

So ψ factors through φ. Since ψ was arbitrary, φ satisfies $\bigotimes_2$.

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 $\varphi : C(S \times T) \to C(S) \otimes C(T)$ with $\varphi(s, t) = s \otimes t$ for $s \in S$ and $t \in T$:

$$\begin{tikzcd}
S \times T \arrow[r, "\otimes"] \arrow[d, "i"'] & C(S) \otimes C(T) \\
C(S \times T) \arrow[ur, "\varphi"'] &
\end{tikzcd}$$

Observe that φ is injective since

$$0 = \varphi\Big(\sum_i \lambda_i (s_i, t_i)\Big) = \sum_i \lambda_i \, s_i \otimes t_i$$

implies $\lambda_i = 0$ by linear independence of the s_i in S and the t_i in T (1.5.1). Also φ is surjective since

$$\Big(\sum_i \lambda_i s_i\Big) \otimes \Big(\sum_j \mu_j t_j\Big) = \sum_{i,j} \lambda_i \mu_j \, s_i \otimes t_j = \sum_{i,j} \lambda_i \mu_j \, \varphi(s_i, t_j)$$

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*:

$$\Big(\bigoplus_\alpha E_\alpha\Big) \otimes \Big(\bigoplus_\beta F_\beta\Big) = \bigoplus_{\alpha,\beta} E_\alpha \otimes F_\beta$$

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

$$E \otimes F \cong \Big(\bigoplus_\alpha E_\alpha\Big) \otimes \Big(\bigoplus_\beta F_\beta\Big) = \bigoplus_{\alpha,\beta} E_\alpha \otimes F_\beta$$

by the result for external direct sums in the previous section. Note 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

$$x_s = \sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^{s-1} \mu_j x'_j \qquad (\lambda_i, \mu_j \in \Gamma)$$

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. □

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. □

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 $s \leq r$. □

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 $\varphi : E \times F \to G$ satisfies $\bigotimes_2$ 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 $f : E \otimes F \to G$ is the induced linear map with $\varphi = f \circ \otimes$, then φ satisfies $\bigotimes_2$ 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 φ does, since $\otimes$ does (1.5.1). □

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 $\dim A = 1$. □

Exercise (5)**.** If E, E^* and F, F^* are pairs of dual spaces of finite dimension and $\beta : E \times F \to B(E^*, F^*)$ 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 $x_1, \ldots, x_n$ be a basis in E with dual basis $x^{*1}, \ldots, x^{*n}$ in E^* and let $y_1, \ldots, y_m$ be a basis in F with dual basis $y^{*1}, \ldots, y^{*m}$ in F^*. Let $\varphi^{*kl} : E^* \times F^* \to \Gamma$ 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 $\operatorname{Im}\beta = B(E^*, F^*)$, so β satisfies $\bigotimes_1$.

If $\varphi : E \times F \to G$ is bilinear, define $f : B(E^*, F^*) \to G$ 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 β satisfies $\bigotimes_2$. □

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 $x_i \to \lambda_i$ and $y_i \to \mu_i$ 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 $\varphi : E \to E'$ and $\psi : F \to F'$ to the arrow $\varphi \otimes \psi : E \otimes F \to E' \otimes F'$.

Remark. Corollary III to Proposition 1.16.1 is not established in section 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.[14]

§ 19

Remark. The proof of (1.12) is captured in the following diagram:

$$\begin{tikzcd}
E \otimes F \arrow[rr, "\varphi \otimes \psi"] \arrow[dd, "\pi_1 \otimes \pi_2"'] \arrow[ddrr, "\pi"] & & E' \otimes F' \\
& & \\
\overline{E} \otimes \overline{F} \arrow[rr, "g"'] \arrow[uurr, "\overline{\varphi} \otimes \overline{\psi}"] & & \overline{E \otimes F} \arrow[uu, "\chi"']
\end{tikzcd}$$

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 φ is not.

Exercise (2)**.** If $\varphi : E \to E$ and $\psi : F \to F$ are linear with $\dim E = n$ and $\dim F = m$, then

$$\operatorname{tr}(\varphi \otimes \psi) = \operatorname{tr} \varphi \cdot \operatorname{tr} \psi$$

and

$$\det(\varphi \otimes \psi) = (\det \varphi)^m \cdot (\det \psi)^n$$

[14] 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

$$\varphi e_i = \sum_{k=1}^{n} \alpha_i^k e_k$$
$$\psi f_j = \sum_{l=1}^{m} \beta_j^l f_l$$

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

$$\operatorname{tr}(\varphi \otimes \psi) = \sum_{i,j} \alpha_i^i \beta_j^j = \big(\sum_i \alpha_i^i\big)\big(\sum_j \beta_j^j\big) = \operatorname{tr}\varphi \cdot \operatorname{tr}\psi$$

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 $\det(\iota_E \otimes \psi) = (\det\psi)^n$. □

Exercise (3)**.** If $\alpha, \beta : E \to E$ are linear with $\dim E = n$, and $\Phi : L(E;E) \to L(E;E)$ is defined by

$$\Phi\sigma = \alpha \circ \sigma \circ \beta$$

then

$$\operatorname{tr}\Phi = \operatorname{tr}\alpha \cdot \operatorname{tr}\beta$$

and

$$\det\Phi = \det(\alpha \circ \beta)^n$$

Proof. Let $e_1, \dots, e_n$ be a basis of E and let $\psi_{ij} : E \to E$ be the induced basis transformation defined by $\psi_{ij} e_k = \delta_k^i e_j$. The isomorphism $\Psi : E \otimes E \to L(E;E)$ defined by $\Psi(e_i \otimes e_j) = \psi_{ij}$ induces an isomorphism

$$\widehat{\Psi} : L(E \otimes E; E \otimes E) \cong L(L(E;E); L(E;E))$$

by $\widehat{\Psi}(\sigma \otimes \tau) = \Psi \circ (\sigma \otimes \tau) \circ \Psi^{-1}$:

$$
\begin{tikzcd}
E \otimes E \arrow[r, "\Psi"] \arrow[d, "\sigma \otimes \tau"'] & L(E;E) \arrow[d, "\widehat{\Psi}(\sigma \otimes \tau)"] \\
E \otimes E \arrow[r, "\Psi"'] & L(E;E)
\end{tikzcd}
$$

It is easy to verify that

$$\widehat{\Psi}(\beta^* \otimes \alpha) = \Phi$$

where β^* is the transpose of β. Therefore by the previous exercise,

$$\operatorname{tr}\Phi = \operatorname{tr}(\beta^* \otimes \alpha) = \operatorname{tr}(\beta^*) \cdot \operatorname{tr}\alpha = \operatorname{tr}\alpha \cdot \operatorname{tr}\beta$$

and

$$\det\Phi = \det(\beta^* \otimes \alpha) = (\det\beta^*)^n (\det\alpha)^n = \det(\alpha \circ \beta)^n \qquad \square$$

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

$$
\begin{aligned}
\alpha \circ \sigma \circ \beta &= (a^* \otimes b) \circ (x^* \otimes y) \circ (c^* \otimes d) \\
&= (a^* \otimes b) \circ \big(\langle x^*, d\rangle (c^* \otimes y)\big) \\
&= \langle x^*, d\rangle \langle a^*, y\rangle (c^* \otimes b) \\
&= \big(\langle x^*, d\rangle c^*\big) \otimes \big(\langle a^*, y\rangle b\big) \\
&= \beta^* x^* \otimes \alpha y \\
&= (\beta^* \otimes \alpha)\sigma
\end{aligned}
$$

so $\Phi = \beta^* \otimes \alpha$. Here $\beta^* = d \otimes c^*$.[15] Alternately, we could use the isomorphism Ω from Proposition 1.29.1.

§ 20

Remark. A *linear* map is universal if and only if it is an isomorphism. In fact, it satisfies $\bigotimes_1$ if and only if it is surjective, and it satisfies $\bigotimes_2$ 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.

[15] See problem 1.29.1.

Remark. To see that there is a unique isomorphism

$$f : (E_1 \otimes \cdots \otimes E_p) \otimes (E_{p+1} \otimes \cdots \otimes E_{p+q}) \to E_1 \otimes \cdots \otimes E_{p+q}$$

with

$$f((x_1 \otimes \cdots \otimes x_p) \otimes (x_{p+1} \otimes \cdots \otimes x_{p+q})) = x_1 \otimes \cdots \otimes x_{p+q} \tag{1}$$

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 $\varphi_{x_1 \dots x_p} : E_{p+1} \times \cdots \times E_{p+q} \to E_1 \otimes \cdots \otimes E_{p+q}$ 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 $f_{x_1 \dots x_p} : E_{p+1} \otimes \cdots \otimes E_{p+q} \to E_1 \otimes \cdots \otimes E_{p+q}$ 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

$$\widehat{f} : E_1 \otimes \cdots \otimes E_p \to L(E_{p+1} \otimes \cdots \otimes E_{p+q}; E_1 \otimes \cdots \otimes E_{p+q})$$

with $\widehat{f}(x_1 \otimes \cdots \otimes x_p) = f_{x_1 \dots x_p}$. Define

$$\varphi : (E_1 \otimes \cdots \otimes E_p) \times (E_{p+1} \otimes \cdots \otimes E_{p+q}) \to E_1 \otimes \cdots \otimes E_{p+q}$$

by $\varphi(x, y) = \widehat{f}(x)(y)$. Then φ 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

$$\psi : E_1 \times \cdots \times E_{p+q} \to (E_1 \otimes \cdots \otimes E_p) \otimes (E_{p+1} \otimes \cdots \otimes E_{p+q})$$

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

$$g : E_1 \otimes \cdots \otimes E_{p+q} \to (E_1 \otimes \cdots \otimes E_p) \otimes (E_{p+1} \otimes \cdots \otimes E_{p+q})$$

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_ν^i are linearly independent in E_i $(i = 1, \ldots, p)$, then their tensor products $a_{\nu_1}^1 \otimes \cdots \otimes a_{\nu_p}^p$ are linearly independent in $E_1 \otimes \cdots \otimes E_p$. Indeed, if $p = 2$, this follows from Lemma 1.5.1. If $p > 2$, we use the natural isomorphism

$$E_1 \otimes \cdots \otimes E_p \cong (E_1 \otimes \cdots \otimes E_{p-1}) \otimes E_p$$

If

$$\sum_{\nu_1, \ldots, \nu_p} \lambda_{\nu_1 \ldots \nu_p} a_{\nu_1}^1 \otimes \cdots \otimes a_{\nu_p}^p = 0$$

then

$$\sum_{\nu_1, \ldots, \nu_p} (\lambda_{\nu_1 \ldots \nu_p} a_{\nu_1}^1 \otimes \cdots \otimes a_{\nu_{p-1}}^{p-1}) \otimes a_{\nu_p}^p = 0$$

Since the $a_{\nu_p}^p$ are linearly independent in E_p, it follows from Lemma 1.5.1 that

$$\lambda_{\nu_1 \ldots \nu_p} a_{\nu_1}^1 \otimes \cdots \otimes a_{\nu_{p-1}}^{p-1} = 0$$

for all $\nu_1, \ldots, \nu_p$. But $a_{\nu_1}^1 \otimes \cdots \otimes a_{\nu_{p-1}}^{p-1} \neq 0$ since $a_\nu^i \neq 0$ for all i, ν (see problem 1), so $\lambda_{\nu_1 \ldots \nu_p} = 0$ for all $\nu_1, \ldots, \nu_p$.

If the vectors a_ν^i span E_i $(i = 1, \ldots, p)$, then clearly their tensor products span $E_1 \otimes \cdots \otimes E_p$. It follows that if the vectors a_ν^i form a basis of E_i $(i = 1, \ldots, p)$, then their tensor products form a basis of $E_1 \otimes \cdots \otimes E_p$.

Remark. To see that for $\varphi_i : E_i \to F_i$ the map $(\varphi_1, \ldots, \varphi_p) \mapsto \varphi_1 \otimes \cdots \otimes \varphi_p$ induces an injection

$$L(E_1; F_1) \otimes \cdots \otimes L(E_p; F_p) \to L(E_1 \otimes \cdots \otimes E_p; F_1 \otimes \cdots \otimes F_p)$$

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

$$\begin{aligned} L(E_1; F_1) \otimes \cdots \otimes L(E_p; F_p) &= \big(L(E_1; F_1) \otimes \cdots \otimes L(E_{p-1}; F_{p-1})\big) \otimes L(E_p; F_p) \\ &\subseteq L(E_1 \otimes \cdots \otimes E_{p-1}; F_1 \otimes \cdots \otimes F_{p-1}) \otimes L(E_p; F_p) \\ &\subseteq L((E_1 \otimes \cdots \otimes E_{p-1}) \otimes E_p; (F_1 \otimes \cdots \otimes F_{p-1}) \otimes F_p) \\ &= L(E_1 \otimes \cdots \otimes E_p; F_1 \otimes \cdots \otimes F_p) \end{aligned}$$

Exercise (1)**.** In $E_1 \otimes \cdots \otimes E_p$:

(a) $x_1 \otimes \cdots \otimes x_p = 0$ 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 $i = 1, \ldots, p-1$ 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,

$$y_1 \otimes \cdots \otimes y_{p-1} = \lambda_p^{-1} x_1 \otimes \cdots \otimes x_{p-1} \qquad \text{and} \qquad y_p = \lambda_p x_p$$

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 $i = 1, \ldots, p-1$ with $\mu_1 \cdots \mu_{p-1} = 1$ by induction. Setting $\lambda_i = \mu_i \mu$ for $i = 1, \ldots, p-1$, we have $y_i = \lambda_i x_i$ for $i = 1, \ldots, p$ and $\lambda_1 \cdots \lambda_p = 1$ as desired. □

§ 21

Remark. For $p, q \geq 2$, let

$$\varphi_i : \prod_{j=1}^{p} E_j^i \to E_{p+1}^i \qquad (i = 1, \ldots, q)$$

be a family of q p-linear maps. Then there is a unique p-linear map

$$\varphi = \varphi_1 \otimes \cdots \otimes \varphi_q : \prod_{j=1}^{p} \left(\bigotimes_{i=1}^{q} E_j^i \right) \to \bigotimes_{i=1}^{q} E_{p+1}^i$$

with

$$\varphi(x_1^1 \otimes \cdots \otimes x_1^q, \ldots, x_p^1 \otimes \cdots \otimes x_p^q) = \varphi_1(x_1^1, \ldots, x_p^1) \otimes \cdots \otimes \varphi_q(x_1^q, \ldots, x_p^q)$$

Moreover,

$$\ker_j \varphi = \sum_{i=1}^{q} E_j^1 \otimes \cdots \otimes \ker_j \varphi_i \otimes \cdots \otimes E_j^q \qquad (j = 1, \ldots, p)$$

In particular, if each φ_i is nondegenerate, then φ is nondegenerate.

§ 22

Remark. If Φ is bilinear in $E \times E'$ and Ψ is bilinear in $F \times F'$, then nondegeneracy of $\Phi \otimes \Psi$ does not imply nondegeneracy of Φ and Ψ 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 Φ is degenerate, but $E \otimes F = 0$ and $E' \otimes F' = 0$, so $\Phi \otimes \Psi = 0$ is nondegenerate.[16]

§ 26

Remark. If $e_1, \ldots, e_n$ is a basis in E and $f_1, \ldots, f_n$ is its dual basis in $L(E)$, then the basis transformation φ_{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 $T : L(E) \otimes E \to L(E; E)$ 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.

Exercise (1)**.** For the bilinear map

$$\gamma : L(E, E'; E'') \times L(F, F'; F'') \to L(E \otimes F, E' \otimes F'; E'' \otimes F'')$$

with

$$\gamma(\varphi, \psi) : (x \otimes y, x' \otimes y') \mapsto \varphi(x, x') \otimes \psi(y, y')$$

the pair $(\operatorname{Im} \gamma, \gamma)$ is the tensor product of $L(E, E'; E'')$ and $L(F, F'; F'')$.

[16]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:

$$\begin{tikzcd}
L(E \otimes E'; E'') \otimes L(F \otimes F'; F'') \arrow[r, "f"] & L((E \otimes E') \otimes (F \otimes F'); E'' \otimes F'') \arrow[d, "\cong"] \\
 & L((E \otimes F) \otimes (E' \otimes F'); E'' \otimes F'') \arrow[d, "\cong"] \\
L(E, E'; E'') \otimes L(F, F'; F'') \arrow[uu, "\cong"'] \arrow[r, "g"] & L(E \otimes F, E' \otimes F'; E'' \otimes F'') \\
L(E, E'; E'') \times L(F, F'; F'') \arrow[u, "\otimes"'] \arrow[ur, "\gamma"'] &
\end{tikzcd}$$

Note g is injective since f is injective (1.16.1), so γ satisfies $\bigotimes_2$ (1.8.1). Since γ also satisfies $\bigotimes_1$, it follows that γ 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

$$(E^* \otimes F^*)/(E_1^\perp \otimes F^* + E^* \otimes F_1^\perp) \qquad \text{and} \qquad E_1 \otimes F_1$$

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. □

Exercise (3)**.** If E, E^* are dual spaces and $\varphi : E \to E$ 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 $(\varphi \otimes \varphi^*)^* = \varphi \otimes \varphi^*$. □

§ 28

In this section, 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 $\operatorname{tr}(\varphi \otimes \psi) = \operatorname{tr} \varphi \cdot \operatorname{tr} \psi$. 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

$$\operatorname{tr}(\varphi \otimes \psi) = \langle a^* \otimes b^*, a \otimes b \rangle = \langle a^*, a \rangle \langle b^*, b \rangle = \operatorname{tr} \varphi \cdot \operatorname{tr} \psi$$

§ 29

In this section, 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 $a^* \otimes b : E \to F$ is defined by $x \mapsto \langle a^*, x\rangle b$ and $b \otimes a^* : F^* \to E^*$ is defined by $y^* \mapsto \langle y^*, b\rangle a^*$. For $y^* \in F^*$ and $x \in E$,

$$\langle (b \otimes a^*)y^*, x\rangle = \langle y^*, b\rangle\langle a^*, x\rangle = \langle y^*, (a^* \otimes b)x\rangle$$ □

Exercise (2)**.** If $E, F \neq 0$ are Euclidean spaces with $\varphi : E \to E$ and $\psi : F \to F$, then $\varphi \otimes \psi : E \otimes F \to E \otimes F$ is a rotation if and only if $\varphi = \lambda\tau_E$ and $\psi = \lambda^{-1}\tau_F$ where τ_E and τ_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

$$\widetilde{\varphi} \otimes \widetilde{\psi} = \widetilde{\varphi \otimes \psi} = (\varphi \otimes \psi)^{-1} = \varphi^{-1} \otimes \psi^{-1}$$

It follows that $\widetilde{\varphi} = \mu\varphi^{-1}$ and $\widetilde{\psi} = \mu^{-1}\psi^{-1}$ for some $\mu \neq 0$.[17] 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 τ_E is a rotation, and similarly for τ_F. □

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 $\varphi, \psi : E \to E$ 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 φ and ψ either both preserve orientation or both reverse orientation.

Proof. This follows from $\det(\varphi \otimes \psi) = (\det\varphi \cdot \det\psi)^n$.[18] □

Chapter 2

§ 2

Remark. If A and B are algebras and $A \otimes B$ denotes the tensor product of the underlying *vector spaces*, define $\varphi : A \times B \times A \times B \to A \otimes B$ by

$$\varphi(x_1, y_1, x_2, y_2) = x_1 x_2 \otimes y_1 y_2$$

Then φ is multilinear, so it induces a linear map $f : (A \otimes B) \otimes (A \otimes B) \to A \otimes B$ 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.

[17] See problem 1.8.4.
[18] See problem 1.19.2.

§ 3

Remark. If $\varphi_1 : A_1 \to B_1$ and $\varphi_2 : A_2 \to B_2$ 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 φ is linear, it follows that φ is an algebra homomorphism. This proof avoids the use of structure maps.

Chapter 3

§ 3

Remark. The tensor algebra $\bigotimes E$ 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 $E \to A$, for associative unital algebras A, and the arrows are algebra homomorphisms $A \to B$ which preserve the identity and respect the linear maps from E:

$$\begin{tikzcd} E \arrow[r] \arrow[d] & B \\ A \arrow[ur] & \end{tikzcd}$$

Every object in this category can be obtained from the tensor algebra $\bigotimes E$ in a unique way. This is why $\bigotimes E$ is said to satisfy the "universal property". This is only possible because the elements of $\bigotimes E$ satisfy only those properties that are required to make $\bigotimes E$ into an associative unital algebra containing E. By category theoretic abstract nonsense, $\bigotimes E$ 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 $\bigotimes E$ and $\bigotimes F$, and which sends a linear map $\varphi : E \to F$ to the algebra homomorphism $\varphi_\otimes : \bigotimes E \to \bigotimes F$.

§ 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 $\varphi : E \to E$ is linear, let $\Delta : E^n \to \Gamma$ 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 $\bigotimes^n E$, there is an induced linear function $d : \bigotimes^n E \to \Gamma$ with $d(x_1 \otimes \cdots \otimes x_n) = \Delta(x_1, \dots, x_n)$. Now

$$\begin{aligned}
\operatorname{tr}\varphi \cdot d(x_1 \otimes \cdots \otimes x_n) &= \sum_{i=1}^{n} d(x_1 \otimes \cdots \otimes \varphi x_i \otimes \cdots \otimes x_n) \\
&= d(\sum_{i=1}^{n} x_1 \otimes \cdots \otimes \varphi x_i \otimes \cdots \otimes x_n) \\
&= d(\theta_{ⓝ}(\varphi)(x_1 \otimes \cdots \otimes x_n))
\end{aligned}$$

It follows that

$$\operatorname{tr}\varphi \cdot d = d \circ \theta_{ⓝ}(\varphi)$$

§ 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. □

§ 11

Remark. In the finite-dimensional case, the following are familiar types of mixed tensors:

Bidegree	**Type**
(0,0)	Scalar
(0,1)	Vector
(1,0)	Function
(1,1)	Transformation

§ 12

Remark. We have

$$\bigotimes(E^*,E)=\sum_{p,q}\bigotimes{}_q^p(E^*,E)$$

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.

§ 14

Remark. The mapping $T:GL(E)\to GL(\bigotimes_q^p(E^*,E))$ given by $\alpha\mapsto T_\alpha$ is a group homomorphism. The map T_α is tensorial if α 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 $\bigotimes_q^p(E^*,E)\otimes\bigotimes_q^p(E^*,E)=\bigotimes_q^p(E^*\otimes E^*,E\otimes E)$ (see problem 10), we have $T_\alpha\otimes T_\beta=T_{\alpha\otimes\beta}$. Writing $\bigotimes_q^p(E^*,E)\otimes\bigotimes_s^r(E^*,E)=\bigotimes_{q+s}^{p+r}(E^*,E)$, we have $T_\alpha\otimes T_\alpha=T_\alpha$.

Remark. If e_i,e^{*i} and $\bar e_i,\bar e^{*i}$ are pairs of dual bases in E,E^* and α is the change of basis transformation $e_i\mapsto\bar e_i$ in E, then α^{*-1} is the corresponding change of basis transformation $e^{*i}\mapsto\bar e^{*i}$ in E^*.[19] Therefore in $\bigotimes_q^p(E^*,E)$,

$$T_\alpha(e_{\mu_1\cdots\mu_q}^{\nu_1\cdots\nu_p})=\bar e_{\mu_1\cdots\mu_q}^{\nu_1\cdots\nu_p}$$

In other words, T_α is the induced change of basis transformation in $\bigotimes_q^p(E^*,E)$. It follows that a mapping is tensorial if and only if its matrix is invariant under this type of change of basis.

[19] See the remark in Chapter III, § 3 of [1] above.

Exercise (3)**.** For $u^* \in \bigotimes^p E^*$ and $u \in \bigotimes^p E$,

$$\langle u^*, u \rangle = (C_1^1)^p (u^* \otimes u)$$

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

$$\begin{aligned} (C_1^1)^p (u^* \otimes u) &= (C_1^1)^{p-1} (C_1^1 (u^* \otimes u)) \\ &= \langle u^{*1}, u_1 \rangle (C_1^1)^{p-1} (\tilde{u}^* \otimes \tilde{u}) \\ &= \langle u^{*1}, u_1 \rangle \langle \tilde{u}^*, \tilde{u} \rangle \\ &= \langle u^*, u \rangle \end{aligned} \qquad \square$$

Exercise (4)**.** If the mappings

$$\Phi : \bigotimes_q^p (E^*, E) \to \bigotimes_s^r (E^*, E) \quad \text{and} \quad \Phi^* : \bigotimes_p^q (E^*, E) \leftarrow \bigotimes_r^s (E^*, E)$$

are dual and Φ is tensorial, then Φ^* is tensorial.

Proof. If α is an automorphism of E, then by a remark above

$$\Phi^* \circ T_\alpha = \Phi^* \circ T^*_{\alpha^{-1}} = (T_{\alpha^{-1}} \circ \Phi)^* = (\Phi \circ T_{\alpha^{-1}})^* = T^*_{\alpha^{-1}} \circ \Phi^* = T_\alpha \circ \Phi^* \qquad \square$$

Exercise (5)**.** The sum, composite, and tensor product of tensorial mappings are tensorial.

Proof. If Φ and Ψ 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. $\square$

References

[1] Greub, W. *Linear Algebra*, 4th ed. Springer, 1975.

[2] Greub,W. *Multilinear Algebra*, 2nd ed. Springer, 1978.