Chapter 3

Tensor product

In Chapter 2 we have looked at the conjugation action of GL(V) on matrices. This action corresponds with the view of matrices as linear transformations. In Chapter 1 we have looked into the role of matrices for describing linear subspaces of $\mathbb{F}^n$. In Remark 1.1.2, we mentioned yet another interpretation of a matrix $A$, namely as data determining a bilinear map $\mathbb{F}^n \times \mathbb{F}^m \to \mathbb{F}$ by means of matrix multiplication:

$$(x,y) \mapsto (x_1, \ldots, x_n)^\top A(y_1, \ldots, y_m).$$

In Chapter 4 we will take a closer look at that interpretation. In this chapter we pave the way by taking a closer look at the notion ‘bilinear’. For this purpose we will be using a very general construction: tensor products of vector spaces.

The foundation of tensor products lies in a natural extension of the notion of linear map: the multilinear map. If we take multi equal to 1 then we get linear; equal to 2 gives us bilinear, 3 gives us trilinear, et cetera. We shall focus on bilinear maps. In this chapter $V, W, U, V_1, V_2, \ldots$ are vector spaces over $\mathbb{F}$. In general they are finite dimensional, but every now and then we will make an exception to this rule.

### 3.1 Bilinear maps

**Definition 3.1.1** A map $f : V \oplus W \to U$ with the property that, for every $v \in V$ and every $w \in W$, the maps $x \mapsto f(x,w)$ and $y \mapsto f(v,y)$ are linear, is called bilinear (with respect to $V$ and $W$).

In general, for every natural number $t$ a map $f : V_1 \oplus V_2 \oplus \cdots \oplus V_t \to U$ is called $t$-linear (with respect to $V_1, \ldots, V_t$) if, for every $i \in \{1, \ldots, t\}$ and every $v_j \in V_j$ ($j \neq i$), the map $x \mapsto f(v_1, \ldots, v_{i-1}, x, v_{i+1}, \ldots, v_t)$ is a linear map from $V_i$ to $U$.

The set of all $t$-linear maps $f : V_1 \oplus V_2 \oplus \cdots \oplus V_t \to U$ will be denoted by $\mathcal{ML}(V_1, V_2, \ldots, V_t; U)$.

Thus, if we view $f$ as a map with $t$ arguments, then $t$-linearity means that $f$ is linear in each of its $t$ arguments.

**Example 3.1.1** Here are two well-known special cases.

(i) Inner products on $V$ are bilinear maps $V \oplus V \to \mathbb{F}$.

(ii) The determinant, defined on the vector space $M_n(\mathbb{F})$ of $n \times n$ matrices over $\mathbb{F}$, can be viewed as a $n$-linear map from $\mathbb{F}^n \oplus \mathbb{F}^n \oplus \cdots \oplus \mathbb{F}^n$ ($n$ copies) to $\mathbb{F}$: the $n$ arguments are the the columns of the matrix in $M_n(\mathbb{F})$. 

1
Exercise 3.1 Often the word ‘product’ refers to a bilinear map. Think of the inner product described as above. Prove that also the following products are bilinear maps.

(i) Matrix multiplication $M_n(\mathbb{F}) \oplus M_n(\mathbb{F}) \rightarrow M_n(\mathbb{F})$.

(ii) Outer product $\mathbb{F}^3 \oplus \mathbb{F}^3 \rightarrow \mathbb{F}^3$, defined by $(x, y) \mapsto x \wedge y$, where

\[ x \wedge y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1). \]

Exercise 3.2 Show that $\mathcal{ML}(V_1, V_2, \ldots, V_t; U)$, with point-wise addition and scalar multiplication, is a vector space over $\mathbb{F}$. (Harder, and a consequence of the theory which we will develop shortly:) Verify that the dimension is $\dim(U) \prod_i \dim(V_i)$.

We construct a vector space which is so to speak a product of $V$ and $W$. One could think of the set $V \times W$ of all pairs $(v, w)$ with $v \in V$ and $w \in W$. But what should addition look like? If we demand $(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2)$ then we arrive at the direct sum of $V$ and $W$.

So the question for such a product has to be made precise before we can continue. The demand that the product is the ‘universal domain’ for bilinear maps on $V \oplus W$ ‘viewed as linear maps’, helps us further. The product we want to form is called the tensor product and is denoted by $V \otimes W$. It is characterised as ‘the’ vector space $T$ satisfying the following property.

Definition 3.1.2 Let $T$ be a linear vector space over $\mathbb{F}$. We say that $T$ satisfies the characteristic property of the tensor product (with respect to $V$ and $W$) if there is a bilinear map $h : V \oplus W \rightarrow T$ such that the image of $h$ in $T$ spans all of the vector space $T$ and, if, for every bilinear map $f : V \oplus W \rightarrow U$ (with $U$ a vector space), there exists a linear map $g : T \rightarrow U$ with $f = gh$.

One also expresses this by saying that the following diagram commutes.

\[
\begin{array}{ccc}
V \oplus W & \rightarrow & U \\
\downarrow h & & \nearrow g \\
T & & 
\end{array}
\]

This property determines $T$ modulo an isomorphism. Strictly speaking the tensor product does not exist solely of the space $T$ but also of the map $h : V \otimes W \rightarrow T$. The fact that two tensor products are isomorphic should therefore be formulated a little bit more precise: a homomorphism $\psi : (h, T) \rightarrow (h', T')$ is a linear map $\psi : T \rightarrow T'$ such that $h' = \psi h$. The map $\psi$ is called an isomorphism if there is a homomorphism $\phi : (h', T') \rightarrow (h, T)$ which, as a linear map $T' \rightarrow T$, is the inverse of $\psi$.

Lemma 3.1.1 Any two vector spaces which satisfy the characteristic property of the tensor product of $V$ and $W$, are isomorphic.
3.1. Bilinear maps

**Proof:** Suppose both \((h, T)\) and \((h', T')\) satisfy the characteristic property of the tensor product with respect to \(V\) and \(W\). Apply the characteristic property of the tensor product to \((h, T)\) with \((f, U) = (h', T')\). It follows that there exists a linear map \(g : T \to T'\) with \(h' = gh\). Apply the property to \((h', T')\) with \((f, U) = (h, T)\). Then there exists a linear map \(g' : T' \to T\) with \(h = h'g\). Therefore we have both \(h = g'h\) and \(h' = g'h'\). Since \(h(V \oplus W)\) spans all of \(T\), and \(h\) is linear, we find \(g'g = \text{id}_T\), and likewise \(gg' = \text{id}_{T'}\). It follows that \(T\) and \(T'\), and even that \((h, T)\) and \((h', T')\) are isomorphic (by means of \(g\)). \(\square\)

We could take the characteristic property of the tensor product as the definition of the tensor product \(V \otimes W\), except that we have not shown yet its existence! A rather axiomatic definition of the tensor product is the following.

**Definition 3.1.3** Consider the formal vector space \(A\) with basis all pairs \((v, w)\) from \(V \times W\). Let \(B\) be the linear subspace of \(A\) generated by all vectors of the form

- \((v, w_1 + w_2) - (v, w_1) - (v, w_2)\)
- \((v_1 + v_2, w) - (v_1, w) - (v_2, w)\)
- \(\lambda(v, w) - (\lambda v, w)\)
- \(\lambda(v, w) - (v, \lambda w)\)

with \(\lambda \in \mathbb{F}\), \(v_1, v_2 \in V\) and \(w, w_1, w_2 \in W\). Then we define \(V \otimes W\), the tensor product of \(V\) and \(W\), as the quotient vector space \(A/B\).

**Exercise 3.3** Show that the vector space \(A\) of Definition 3.1.3 is finite dimensional if and only if \(\mathbb{F}\) is finite and \(V\) and \(W\) are finite dimensional. \(\square\)

To derive the characteristic property of the tensor product, we use the bilinear map \(h \in \mathcal{ML}(V, W; A/B)\) given by \(h(v, w) = (v, w) + B\).

**Proposition 3.1.1** The tensor product \(V \otimes W\) of \(V\) and \(W\) satisfies the characteristic property of the tensor product.

**Proof:** In the first place we see that the residue class \((v, w) + B\) in \(A/B\) of every basis element \((v, w)\) of \(A\) appears as an image under \(h\), so that indeed \(A/B\) is spanned by \(h(V \oplus W)\). Now let \(f : V \otimes W \to U\) be a bilinear map. To see that \(g\) can be found as in the commutative diagram of Definition 3.1.2, we use the well-known result:

If \(k : A \to U\) is a linear map with \(B\) contained in \(\ker(k)\), then there is a unique linear map \(g : A/B \to U\) with \(g(a + B) = k(a)\) for all \(a \in A\).

We apply this result to the linear map \(k : A \to U\) determined by \(k(v, w) = f(v, w)\) \(((v, w) \in V \oplus W)\). (Since \(k\) is only given on a basis of \(A\), this determines a unique linear map.) The fact that \(f\) is bilinear implies that \(B \subseteq \ker(k)\). We conclude that there is a map \(g : A/B \to U\) with \(g((v, w) + B) = f(v, w)\) for all \(v \in V\) and \(w \in W\), so with \(f = gh\). \(\square\)

**Theorem 3.1.2** Suppose \(m = \dim(W)\) is finite. Then \(V' \otimes W \cong \mathcal{L}(V, W)\).
Proof: We show that $\mathcal{L}(V, W)$, equipped with the bilinear map

$$h : V' \oplus W \rightarrow \mathcal{L}(V, W) \quad \text{given by} \quad h(b, w) = (v \mapsto b(v)w),$$

satisfies the characteristic property for the tensor product of $V'$ and $W$. Then the theorem follows from Lemma 3.1.1.

Choose a basis $(w_i)_i$ of $W$. As $a \in \mathcal{L}(V, W)$, there are linear functionals $b_i \in V'$ such that $a = (v \mapsto \sum_i b_i(v)w_i)$, so $a = \sum_i h(b_i, w_i)$ is in the linear span in $\mathcal{L}(V, W)$ of the image of $h$.

Now let $f : V' \oplus W \rightarrow U$ be a bilinear map. Then the linear map $g : \mathcal{L}(V, W) \rightarrow U$ given by

$$g(v \mapsto \sum_i b_i(v)w_i) = \sum_i f(b_i, w_i)$$

makes the following diagram commutative:

$$\begin{align*}
V' \oplus W & \xrightarrow{f} U \\
\downarrow h & \quad \nearrow g \\
\mathcal{L}(V, W) & 
\end{align*}$$

We have seen that the pair $(h, \mathcal{L}(V, W))$ satisfies the characteristic property for the tensor product of $V'$ and $W$. This suffices for the proof of the theorem.

Exercise 3.4 Investigate where the theorem goes wrong if $W$ is not finite dimensional.

Remark 3.1.1 From the theorem it follows that $U' = \mathcal{L}(U, \mathbb{F}) \cong U' \otimes \mathbb{F}$. But this can also be seen directly. Verify this!

Remark 3.1.2 By Theorem 3.1.2 the linear map

$$V' \otimes W \rightarrow \mathcal{L}(V, W) \quad \text{determined by} \quad f \otimes w \mapsto (v \mapsto f(v)w)$$

is an isomorphism.

Likewise, for finite dimensional $V$, the linear map

$$V \otimes W \rightarrow \mathcal{L}(V', W) \quad \text{determined by} \quad v \otimes w \mapsto (f \mapsto f(v)w)$$

is an isomorphism.

Compare this with the comments following Definition 2.2.1.

Exercise 3.5 Show that the linear map

$$V \otimes W \rightarrow W \otimes V \quad \text{determined by} \quad v \otimes w \mapsto w \otimes v$$

is an isomorphism.

Exercise 3.6 Show that the linear map

$$(U \otimes V) \otimes W \rightarrow U \otimes (V \otimes W) \quad \text{determined by} \quad (u \otimes v) \otimes w \mapsto u \otimes (v \otimes w)$$

is an isomorphism. It therefore does not matter where the parentheses are (or where we put them in case they are missing) in a multiple tensor product.
Exercise 3.7 Prove
\[ \mathcal{M}(V_1, V_2, \ldots, V_t; U) \cong \mathcal{L}(V_1 \otimes V_2 \otimes \cdots \otimes V_t, U). \]

Proposition 3.1.2 (Basis of a tensor product) Suppose \( V \) and \( W \) are finite dimensional. If \((v_i)_i\) is a basis of \( V \) and \((w_j)_j\) is a basis of \( W \), then \((v_i \otimes w_j)_i,j\) is a basis of \( V \otimes W \). In particular we have
\[ \dim(V \otimes W) = \dim(V) \dim(W). \]

Proof: Since \( V \) is finite dimensional, we have \( V \cong V' \). Therefore we have by Theorem 3.1.2 that \( V \otimes W \cong \mathcal{L}(V', W) \). We can make the isomorphism explicit, as indicated in Remark 3.1.2:
\[ \psi : V \otimes W \to \mathcal{L}(V', W) \quad v \otimes w \mapsto (f \mapsto f(v)w) \quad (f \in V', v \in V, w \in W). \]

Now \((f \mapsto f(v_i)w_j)_{i,j}\) is a basis of \( \mathcal{L}(V', W) \), and so \((\psi^{-1}(f \mapsto f(v_i)w_j))_{i,j} = (v_i \otimes w_j)_{i,j}\) is a basis of \( V \otimes W \).

Remark 3.1.3 Let \( V = \mathbb{F}^n \). Using the fact that \( \mathcal{L}(V) \cong V' \otimes V \) we can now interpret \( n \times n \) matrices as vectors in \( V' \otimes V \). The tensor \( y^\top \otimes x \) of an \( n \)-dimensional column vector \( x \) from \( V = \mathbb{F}^n \) and an \( n \)-dimensional row vector \( y^\top \) from \( V' \) corresponds with the linear transformation \( z \mapsto (y^\top z)x \in \mathcal{L}(V) \), i.e., with the rank 1 matrix \( xy^\top \) (for \( (y^\top z)x = x(y^\top z) \)).

Corollary 3.1.3 A linear transformation on a finite dimensional vector space \( V \) has rank at most \( r \) if and only if the corresponding element in \( V' \otimes V \) is the sum of at most \( r \) distinct matrices of rank 1.

Proof: A matrix has rank \( r \) if and only if the image of the associated linear transformation has dimension \( r \).
In the first place every rank 1 matrix is of the form \( xy^\top = \psi(y \otimes x) \) in the notation of the proof of Proposition 3.1.2. So the statement we are trying to prove holds in the case \( r = 1 \). From this it follows directly that a vector \( g \in V' \otimes V \) is the sum of at most \( r \) simple tensors if and only if the associated linear transformation \( \psi(g) \in \mathcal{L}(V) \) is spanned by at most \( r \) distinct vectors.

Exercise 3.8 Write the matrix
\[ A = \begin{pmatrix} 74 & 88 & 102 \\ 82 & 98 & 114 \\ 90 & 108 & 126 \end{pmatrix} \]
as a sum of two simple tensors, or, equivalently, as a sum of two rank 1 matrices.

Exercise 3.9 Write an algorithm which decomposes a square matrix \( A \) into a sum of rank(\( A \)) matrices of rank 1.
Exercise 3.10 For vector spaces \( V \) and \( W \) over \( \mathbb{F} \), prove the following assertions.

(i) \( \mathcal{ML}(V,W;\mathbb{F}) \cong (V \otimes W)' \).

(ii) \( (V \otimes W)' \cong V' \otimes W' \).

3.2 Tensor product of linear maps

Now that we have constructed the tensor product \( U \otimes V \), we would of course like to relate linear maps on \( U \) and \( V \) to linear maps on \( U \otimes V \).

Theorem 3.2.1 For \( A \in \mathcal{L}(U,W) \) and \( B \in \mathcal{L}(V,T) \) the requirements
\[
 u \otimes v \mapsto Au \otimes Bv \quad (u \in U, v \in V)
\]
determine a linear map in \( \mathcal{L}(U \otimes V, W \otimes T) \). This map only depends on \( A \otimes B \in \mathcal{L}(U,W) \otimes \mathcal{L}(V,T) \), so there exists a linear map \( \mathcal{L}(U,W) \otimes \mathcal{L}(V,T) \to \mathcal{L}(U \otimes V, W \otimes T) \) determined by
\[
 A \otimes B \mapsto \left( \sum_i u_i \otimes v_i \mapsto \sum_i Au_i \otimes Bv_i \right).
\]

This map is an isomorphism.

Proof: The map \( f : U \otimes V \to W \otimes T \) with \( f(u,v) = Au \otimes Bv \) is bilinear. The characteristic property of the tensor product (Definition 3.1.2), applied to the standard bilinear map \( h : U \otimes V \to U \otimes V \), yields a linear map \( g : U \otimes V \to W \otimes T \) with \( f = gh \). In particular, for \( u \in U \) and \( v \in V \),
\[
 g(u \otimes v) = gh(u,v) = f(u,v) = Au \otimes Bv,
\]
proving the first statement.

The map sending \( (A,B) \) to the endomorphism \( g \), is itself a bilinear map \( \mathcal{L}(U,W) \otimes \mathcal{L}(V,T) \to \mathcal{L}(U \otimes V, W \otimes T) \). So again applying the characteristic property of the tensor product shows that \( g \) only depends on \( A \otimes B \).

Remains to prove that this map is an isomorphism. Since domain and range have the same dimension (this follows from Theorem 3.1.2 and Proposition 3.1.2), it suffices to see that every \( g \in \mathcal{L}(U \otimes V, W \otimes T) \) is in the image. Prove this yourself. \( \square \)

If we take \( U = W \) and \( V = T \), then we get

Corollary 3.2.2 \( \mathcal{L}(U) \otimes \mathcal{L}(V) \cong \mathcal{L}(U \otimes V) \).

Remark 3.2.1 Let \( e_1, \ldots, e_n \) be a basis of \( V \) and let \( f_1, \ldots, f_m \) be a basis of \( W \). Then \( (e_i \otimes f_j)_{i,j} \) is a basis for \( V \otimes W \). If we order the indices lexicographically (first with respect to \( i \), then with respect to \( j \)), then the matrix of the linear transformation in \( \mathcal{L}(V \otimes W) \) which belongs to the matrices \( A \) and \( B \) of linear transformations in \( \mathcal{L}(V) \) respectively \( \mathcal{L}(W) \), is
\[
 A \otimes B = \begin{pmatrix}
 A_{1,1}B & A_{1,2}B & \ldots & A_{1,n}B \\
 A_{2,1}B & A_{2,2}B & \ldots & A_{2,n}B \\
 \vdots & \vdots & \ddots & \vdots \\
 A_{n,1}B & A_{n,2}B & \ldots & A_{n,n}B
\end{pmatrix}.
\]
3.2. Tensor product of linear maps

For example, if

\[
A = \begin{pmatrix} u & v \\ x & y \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix},
\]

then \( A \otimes B = \begin{pmatrix} u a & u b & u c & v a & v b & v c \\ u d & u e & u f & v d & v e & v f \\ u g & u h & u k & v g & v h & v k \\ x a & x b & x c & y a & y b & y c \\ x d & x e & x f & y d & y e & y f \\ x g & x h & x k & y g & y h & y k \end{pmatrix} \).

This formula identifies the element \( A \otimes B \) from \( \mathcal{L}(\mathbb{F}^2) \otimes \mathcal{L}(\mathbb{F}^3) \) with the 6 \times 6-matrix from \( \mathcal{L}(\mathbb{F}^2 \otimes \mathbb{F}^3) = \mathcal{L}(\mathbb{F}^6) \).

\[ \triangle \]

**Corollary 3.2.3** If \( A, C \in M_m(\mathbb{F}) \) and \( B, D \in M_n(\mathbb{F}) \), then

\[
(A \otimes B)(C \otimes D) = AC \otimes BD.
\]

**Exercise 3.11** Prove the following statements for \( A \in \mathcal{L}(V) \) and \( B \in \mathcal{L}(W) \).

(i) If \( A \) is nilpotent, then so is \( A \otimes B \).

(ii) If \( A \) and \( B \) are diagonalisable, then so is \( A \otimes B \).

(iii) \( \text{rank}(A \otimes B) = \text{rank}(A) \text{rank}(B) \).

(iv) \( \det(A \otimes B) = \det(A)^n \det(B)^m \) where \( m = \dim(V) \) and \( n = \dim(W) \).

(v) \( \text{tr}(A \otimes B) = \text{tr}(A) \text{tr}(B) \).  

\[ \square \]

**Exercise 3.12** Is, for \( A, B \in \mathcal{L}(V) \), the minimal polynomial of \( A \otimes B \in \mathcal{L}(V) \) uniquely determined by \( A \) and \( B \)?  

\[ \square \]

**Exercise 3.13** This exercise is experimental.

(i) For a 4 \times 4 matrix \( A \), determine when it is of the form \( B \otimes C \) with \( B \) and \( C \) both 2 \times 2 matrices. You may give as an answer a system of equations in the coefficients of \( A \). [Hint: The eigenvalues (see below) have to satisfy certain conditions; this can be expressed in conditions on the coefficients of \( A \). Use computer algebra software to experiment.]  

(ii) Determine the smallest natural number \( k \) with the property that every 4 \times 4 matrix \( A \) over \( \mathbb{F} \) can be written as a sum of at most \( k \) tensors of two 2 \times 2 matrices.

\[ \square \]

**Exercise 3.14** Show that for two square matrices \( A \) and \( B \) it holds that

\[
(A \otimes B)^\top = A^\top \otimes B^\top.
\]

\[ \square \]
3.3 Hadamard matrices

As an application we now look at a special type of matrix that has done well in statistics.

**Definition 3.3.1** A Hadamard matrix is a square matrix $H$ of dimension $n$ satisfying

$$HH^\top = nI_n,$$

and with each coefficients either 1 or $-1$.

**Example 3.3.1**

For $n = 1$ the only two Hadamard matrices are $H = (1)$ and $H = (-1)$.

For $n = 2$

$$H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

is a Hadamard matrix.

**Exercise 3.15**

Prove that if $H \in M_n(\mathbb{Q})$ is a Hadamard matrix with $n \geq 4$, then we have $n \equiv 0 \pmod{4}$. 

It is an open problem whether, for each $t$, there exists a Hadamard matrix of dimension $4t$. Tensor products help to make a Hadamard matrix of dimension $mn$ if there exist Hadamard matrices $A$ and $B$ of dimension $m$ and $n$.

**Theorem 3.3.1** If $A$ and $B$ are Hadamard matrices, then so is $A \otimes B$.

**Proof:**

\[
(A \otimes B)(A \otimes B)^\top = (A \otimes B)(A^\top \otimes B^\top) \quad \text{(Exercise 3.14)} \\
= AA^\top \otimes BB^\top \quad \text{(Corollary 3.2.3)} \\
= mI_m \otimes nI_n \\
= mnI_{mn}.
\]

3.4 Bilinear forms and matrices

We have now reached the stage where we can return to our point of departure: understanding a matrix as a data structure for bilinear forms. The bilinear forms on a vector space $V$ are the elements of $\mathcal{ML}(V,V;\mathbb{F})$, but they can also be viewed as element of $(V \otimes V)'$, or of $V' \otimes V'$ (see Exercise 3.10), or as an element of $\mathcal{L}(V,V')$.

Let $f$ be a bilinear form on $V$. If we choose a basis $e_i$ for $V$, then $f$ is determined by the numbers $f_{i,j} = f(e_i,e_j) \in \mathbb{F}$. If we identify $V$ with $\mathbb{F}^n$ using our basis $(e_i)_i$, then the matrix $A_f = (f_{i,j})_{i,j}$ satisfies

$$f(x,y) = x^\top A_f y.$$

We want to investigate how $A_f$ changes under a coordinate transformation. Since a coordinate transformation is an element of the group $\text{GL}(V)$, we shall first investigate how the action of $\text{GL}(V)$ on $V'$ can be described.

To start with, we explain what we mean by an action.
Definition 3.4.1 An action of a group $G$ on a vector space $V$ is a map $G \times V \to V$ such that

(i) $v \mapsto g \cdot v \in \mathcal{L}(V)$ for all $g \in \text{GL}(V)$,

(ii) $\text{id} \cdot v = v$ for all $v \in V$, and

(iii) $(gh) \cdot v = g \cdot (h \cdot v)$ for all $g, h \in \text{GL}(V)$ and $v \in V$.

Example 3.4.1 Thus, $(g, v) \mapsto g(v)$ is an action of $\text{GL}(V)$ on $V$. This is the natural action. If a group $G$ admits actions on $V$ and $W$, then

- the conjugation action of $G$ on $\mathcal{L}(V, W)$ is the map given by $(g, f) \mapsto gfg^{-1}$;

- the diagonal action of $G$ on $V \otimes W$ is the map $G \times V \otimes W \to V \otimes W$ given by

  $$(g, \sum_i v_i \otimes w_i) \mapsto \sum_i gv_i \otimes gw_i;$$

- the dual action of $G$ on $V'$ is the map given by $(g, f) \mapsto fg^{-1}$. (For the dual action, see also Definition 1.1.5.)

\[\triangle\]

Verify that the dual action is a special case of the conjugation action. Indeed, $V' = \mathcal{L}(V, \mathbb{F})$, and $G$ acts trivial on $\mathbb{F}$ (i.e., via $(g, \lambda) \mapsto \lambda$), so that $gfg^{-1} = fg^{-1}$.

The key observation in understanding the action of $\text{GL}(V)$ on $V'$ is that the evaluation function

$$V' \otimes V \to \mathbb{F}, \quad \text{determined by} \quad f \otimes v \mapsto f(v)$$

is not allowed to depend on a coordinate transformation. That means that, if $g \cdot f \in V'$ is the image of $f \in V'$ under $g \in \text{GL}(V)$, then we must have

$$(g \cdot f)(gv) = f(v) \quad \text{for all} \quad v \in V, \ f \in V'.$$

From this it immediately follows that

Lemma 3.4.1 Necessary and sufficient for the evaluation $V' \otimes V \to \mathbb{F}$ (determined by $f \otimes v = f(v)$) to be invariant under the group $\text{GL}(V)$ in her natural action on $V$ is that $\text{GL}(V)$ acts in the dual action on $V'$; i.e., via

$$g \cdot f = f \circ g^{-1} \quad \text{for all} \quad f \in V'.$$

We can also formulate this in the language of matrices. Below we shall denote by the exponent $-\top$ the ‘transposed inverse’: first take the transposed and then the inverse (or the other way around, the order does not matter).

Lemma 3.4.2 The action of $\text{GL}(n, \mathbb{F})$ on $\mathbb{F}^n$ dual to the natural action is given by

$$(T, v) \mapsto T^{-\top} v \quad (T \in \text{GL}(n, \mathbb{F}), v \in \mathbb{F}^n).$$
Chapter 3. Tensor product

Proof: If \( f = y^\top \), then we find \((T \cdot f)(x) = f \circ T^{-1}(x) = y^\top T^{-1}x \) for all \( x \in \mathbb{F}^n \). So the action of \( T \in \text{GL}(n, \mathbb{F}) \) on \((\mathbb{F}^n)'\) (consisting of row vectors, see Definition 1.1.5) which leaves evaluation invariant, is given by

\[
y^\top \mapsto y^\top T^{-1} \quad (T \in \text{GL}(n, \mathbb{F}), \ y \in \mathbb{F}^n)\).
\]

The lemma now follows by choosing the basis \( e'_i \) for \((\mathbb{F}^n)'\), and determining the matrix of the linear transformation \( y^\top \mapsto y^\top T^{-1} \) with respect to this basis. \( \square \)

Exercise 3.16 Suppose we are given two actions of a group \( G \): one on \( V \) and one on \( W \). A map \( \phi : V \to W \) such that \( \phi g = g \phi \) for all \( g \in G \) is called \( G \)-equivariant. Show that there are \( G \)-equivariant isomorphisms in the following cases:

(i) \( \phi : V' \otimes W \to \mathcal{L}(V, W) \).

(ii) \( \phi : \mathcal{ML}(V, W; \mathbb{F}) \to \mathcal{L}(V, W') \).

\( \square \)

Lemma 3.4.3 If \( V = \mathbb{F}^n \) and \( T \in \text{GL}(V) \), then the action of \( T \) on \( f \in \mathcal{ML}(V, V; \mathbb{F}) \) with matrix \( A_f \) is given by

\[
T \mapsto T^{-\top} A_f T^{-1}.
\]

Proof:

\[
(T \cdot f)(x, y) = f(T^{-1}x, T^{-1}y) = (T-1x)^\top A_f T^{-1}y = x^\top (T^{-\top} A_f T^{-1})y
\]

So \( T \) transforms \( A_f \) to \( T^{-\top} A_f T^{-1} \). \( \square \)

Now we really are only interested in bilinear forms modulo coordinate transformations, i.e. in the orbits of \( \text{GL}(V) \) on \( \mathcal{ML}(V, V; \mathbb{F}) \). For this it does not matter whether we look at this action or the dual, i.e., one composed with the automorphism \( g \mapsto g^{-\top} \).

Definition 3.4.2 If we take the dual action of Lemma 3.4.3, then we are dealing with the action of \( \text{GL}(n, \mathbb{F}) \) given by

\[
A \mapsto SAS^\top \quad \text{for} \ S \in \text{GL}(n, \mathbb{F}).
\]

We call this action the congruency action. Matrices (bilinear forms) in the same \( \text{GL}(n, \mathbb{F}) \)-orbit under this action are called congruent.

Notice that this action is indeed different from the conjugation action. In the next chapter we shall study the orbits of \( \text{GL}(V) \) under this action.

Exercise 3.17 Determine the orbits of \( \text{GL}(3, \mathbb{F}_2) \) on \( M_3(\mathbb{F}_2) \) in the congruency action. \( \square \)

Exercise 3.18 The natural \( \text{GL}(V) \)-action on \( V' \otimes V' \) is given as before: \( g \) sends \( f \otimes h \) to \( f \circ g^{-1} \otimes h \circ g^{-1} \). Does this action coincide with the action on \((V \otimes V)'\)? In other words, give a \( \text{GL}(V) \)-equivariant isomorphism \( V' \otimes V' \to (V \otimes V)' \). With what action of \( \text{GL}(V) \) on \( \mathcal{L}(V, V') \) does the action on \( V' \times V' \) coincide? \( \square \)
### Exercise 3.19
Give a solution of the congruency problem for bilinear forms in the one-dimensional case.

In fact we will not investigate all \( \text{GL}(V) \)-orbits under congruency on bilinear forms, but only those orbits that belong to special bilinear forms: the alternating bilinear forms and the symmetric bilinear forms.

**Theorem 3.4.4** Suppose \( \mathbb{F} \) has characteristic not equal to 2. Then the linear space \( \mathcal{ML}(V,V;\mathbb{F}) \) of all bilinear forms on \( V \) is a direct sum of the linear spaces

\[
\mathcal{S}(V) = \{ f \in \mathcal{ML}(V,V;\mathbb{F}) \mid f(x,y) = f(y,x) \text{ for all } x,y \in V \}
\]

and

\[
\mathcal{A}(V) = \{ f \in \mathcal{ML}(V,V;\mathbb{F}) \mid f(x,y) = -f(y,x) \text{ for all } x,y \in V \}.
\]

Both subspaces \( \mathcal{S}(V) \) and \( \mathcal{A}(V) \) are invariant under the congruency action of \( \text{GL}(V) \).

**Proof:** An arbitrary element \( f \in \mathcal{ML}(V,V;\mathbb{F}) \) can be written as

\[(x,y) \mapsto (f(x,y) + f(y,x))/2 + (f(x,y) - f(y,x))/2.
\]

The first summand belongs to \( \mathcal{S}(V) \), the second to \( \mathcal{A}(V) \). These two subsets of \( \mathcal{ML}(V,V;\mathbb{F}) \) are linear subspaces. Their intersection is zero: if \( f \in \mathcal{S}(V) \cap \mathcal{A}(V) \), then we have \( f(y,x) = f(x,y) = -f(y,x) \), so \( 2f(y,x) = 0 \), so \( f(y,x) = 0 \) for all \( x,y \in V \).

To prove the statement at the end of the theorem we choose \( \epsilon = \pm 1 \) such that \( f(x,y) = \epsilon f(y,x) \) for all \( x,y \in V \). Then it follows that for \( g \in \text{GL}(V) \)

\[(g \cdot f)(x,y) = f(g^{-1}x,g^{-1}y) = \epsilon f(g^{-1}y,g^{-1}x) = \epsilon(g \cdot f)(y,x)
\]

so that \( g \cdot f \in \mathcal{A}(V) \) if \( f \in \mathcal{A}(V) \) (with \( \epsilon = -1 \)) and \( g \cdot f \in \mathcal{S}(V) \) if \( f \in \mathcal{S}(V) \) (with \( \epsilon = 1 \)). \( \square \)

**Definition 3.4.3** Let \( \mathbb{F} \) be a field of arbitrary characteristic. A bilinear form \( f \in \mathcal{ML}(V,V;\mathbb{F}) \) is called alternating if \( f(x,x) = 0 \) for all \( x \in V \), and \( f \) is called symmetric if \( f(x,y) = f(y,x) \) for all \( x,y \in V \).

In matrix terms: if \( A \) represents the linear form \( f \) (so \( f(x,y) = xAy^\top \)), then \( f \) is symmetric if and only if \( A \) is symmetric (i.e., \( A^\top = A \)), and \( f \) is alternating if and only if \( A \) is alternating (i.e., \( A^\top = -A \) and \( A_{i,i} = 0 \) for all \( i \)).

**Remark 3.4.1** It is clear (by definition) that \( f \in \mathcal{S}(V) \) if and only if \( f \) is symmetric is. Verify that, if \( \mathbb{F} \) is a field of characteristic not equal to 2, it also holds that \( f \in \mathcal{A}(V) \) if and only if \( f \) is alternating. \( \triangle \)

**Exercise 3.20** Prove that if \( \dim(V) = n \), then

\[
\dim(\mathcal{S}(V)) = \binom{n+1}{2} \quad \text{and} \quad \dim(\mathcal{A}(V)) = \binom{n}{2}.
\]

\( \square \)