§43
Linear Transformations and Matrices

In this paragraph, we learn to construct a new vector space from two given vector spaces V,W, namely the vector space of linear transformations from V into W. We introduce matrices and study the relationship between linear transformations and matrices.

Suppose V and W are vector spaces over a field K. We denote by $\boldsymbol{L}_K(V,W)$ the set of all K-linear mappings from V into W. This set $\boldsymbol{L}_K(V,W)$ is not empty, for at least the mapping $\begin{array}{c} V \to W \\ v \to 0 \end{array}$ is a K-linear transformation in $\boldsymbol{L}_K(V,W)$. We want to define an addition and a multiplication by scalars on $\boldsymbol{L}_K(V,W)$ and make $\boldsymbol{L}_K(V,W)$ into a K-vector space.

Let $T,S \in \boldsymbol{L}_K(V,W)$. How shall we define $T + S$? Well, the only natural way to define $T + S$ is to put $v(T + S) = vT + vS$ for all $v \in V$ (pointwise addition). What about multiplication by scalars? Given $\alpha \in K$ and $T \in \boldsymbol{L}(V,W)$, the mapping αT had better mean: first multiply by α, then apply T, so that $v(\alpha T) := (\alpha v)T$ (or, first apply T, then multiply by α, so that $v(\alpha T) := \alpha(vT)$, but this is the same definition as before).

43.1 Theorem: *Let V,W be vector spaces over a field K and let $\boldsymbol{L}_K(V,W)$ be the set of all K-linear transformations from V into W. For any T,S in $\boldsymbol{L}_K(V,W)$ and for any α in K, we write*

$$v(T + S) = vT + vS, \qquad v(\alpha T) = (\alpha v)T \qquad (v \in V).$$

Under this addition and multiplication by scalars, $\boldsymbol{L}_K(V,W)$ is a vector space over K.

Proof: We show first that $\boldsymbol{L}_K(V,W)$ is an abelian group under addition.

(i) Let $T,S \in \boldsymbol{L}_K(V,W)$. Then
$$\begin{aligned}(\alpha v_1 + \beta v_2)(T + S) &= (\alpha v_1 + \beta v_2)T + (\alpha v_1 + \beta v_2)S \\ &= \alpha(v_1T) + \beta(v_2T) + \alpha(v_1S) + \beta(v_2S) \\ &= \alpha(v_1T) + \alpha(v_1S) + \beta(v_2T) + \beta(v_2S)\end{aligned}$$

$$= \alpha(v_1T + v_1S) + \beta(v_2T + v_2S)$$
$$= \alpha\big(v_1(T + S)\big) + \beta\big(v_2(T + S)\big)$$

for all $\alpha,\beta \in K$ and $v_1,v_2 \in V$. Thus $T + S$ is K-linear and $T + S \in \boldsymbol{L}_K(V,W)$. Therefore $\boldsymbol{L}_K(V,W)$ is closed under addition.

(ii) Let T,S,R be arbitrary elements of $\boldsymbol{L}_K(V,W)$. Then

$$v\big((T + S) + R\big) = v(T + S) + vR = (vT + vS) + vR$$
$$= vT + (vS + vR) = vT + v(S + R) = v\big(T + (S + R)\big)$$

for all $v \in V$; hence $(T + S) + R = T + (S + R)$. Thus addition in $\boldsymbol{L}_K(V,W)$ is associative.

(iii) Let $0^*\colon V \to W$, $v \mapsto 0$. Then $(\alpha v_1 + \beta v_2)0^* = 0 = \alpha 0 + \beta 0$ $= \alpha(v_10^*) + \beta(v_20^*)$ for all $\alpha,\beta \in K$, $v_1,v_2 \in V$ and 0^* is in $\boldsymbol{L}_K(V,W)$. From

$$v(T + 0^*) = vT + v0^* = vT + 0 = vT \qquad (v \in V),$$

we obtain $T + 0^* = T$ for any $T \in \boldsymbol{L}_K(V,W)$. Thus 0^* is a right identity.

(iv) Any $T \in \boldsymbol{L}_K(V,W)$ has an opposite in $\boldsymbol{L}_K(V,W)$, namely the mapping $S\colon V \to W$, $v \mapsto -(vT)$. Indeed

$$v(T + S) = vT + vS = vT + \big(-(vT)\big) = 0 = v0^*$$

for all $v \in V$, so $T + S = 0^*$. Are we done? No! We should check that S is in fact in $\boldsymbol{L}(V,W)$, but this easy:

$$(\alpha v_1 + \beta v_2)S = -\big((\alpha v_1 + \beta v_2)T\big)$$
$$= \big(-(\alpha v_1 + \beta v_2)\big)T \qquad \text{(Lemma 41.5(2))}$$
$$= \big((-\alpha v_1) + (-\beta v_2)\big)T$$
$$= \big(\alpha(-v_1) + \beta(-v_2)\big)T$$
$$= \alpha\big((-v_1)\big)T + \beta\big((-v_2)\big)T$$
$$= \alpha\big(-(v_1T)\big) + \beta\big(-(v_2T)\big)$$
$$= \alpha(v_1S) + \beta(v_2S)$$

for all $\alpha,\beta \in K$ and $v_1,v_2 \in V$. Thus S is in $\boldsymbol{L}_K(V,W)$ and S is a right inverse of T.

(v) Finally, $T + S = S + T$ for any $T,S \in \boldsymbol{L}_K(V,W)$, because

$$v(T + S) = vT + vS = vS + vT = v(S + T)$$

for all $v \in V$. Hence $\boldsymbol{L}_K(V,W)$ is a commutative group under addition.

Now the properties of multiplication by scalars. First we note that αT is in $\boldsymbol{L}(V,W)$ whenever $\alpha \in K$ and $T \in \boldsymbol{L}_K(V,W)$, because

$$\begin{aligned}(v_1 + v_2)(\alpha T) = (\alpha(v_1 + v_2))T &= (\alpha v_1 + \alpha v_2)T \\ &= (\alpha v_1)T + (\alpha v_2)T = v_1(\alpha T) + v_2(\alpha T)\end{aligned}$$

for all $v_1, v_2 \in V$, so that αT is additive and

$$(\beta v)(\alpha T) = (\alpha(\beta v))T = ((\alpha\beta)v)T = ((\beta\alpha)v)T = ((\beta(\alpha v))T = \beta((\alpha v)T) = \beta(v(\alpha T))$$

for all $\beta \in K$, $v \in V$, so that αT is homogeneous. So αT belongs to $\boldsymbol{L}_K(V,W)$.

(1) $\alpha(T + S) = \alpha T + \alpha S$ for all $\alpha \in K$ and $T,S \in \boldsymbol{L}_K(V,W)$ since

$$v(\alpha(T + S)) = (\alpha v)(T + S) = (\alpha v)T + (\alpha v)S = v(\alpha T) + v(\alpha S) = v(\alpha T + \alpha S)$$

for any $v \in V$.

(2) $(\alpha + \beta)T = \alpha T + \beta T$ for all $\alpha,\beta \in K$ and $T \in \boldsymbol{L}_K(V,W)$ since

$$\begin{aligned}v((\alpha + \beta)T) &= ((\alpha + \beta)v)T = (\alpha v + \beta v)T \\ &= (\alpha v)T + (\beta v)T = v(\alpha T) + v(\beta T) = v(\alpha T + \beta T)\end{aligned}$$

for any $v \in V$.

(3) $(\alpha\beta)T = \alpha(\beta T)$ for all $\alpha,\beta \in K$ and $T \in \boldsymbol{L}_K(V,W)$ since

$$v((\alpha\beta)T) = ((\alpha\beta)v)T = ((\beta\alpha)v)T = (\beta(\alpha v))T = (\alpha v)(\beta T) = v(\alpha(\beta T))$$

for any $v \in V$.

(4) $1T = T$ for all $T \in \boldsymbol{L}_K(V,W)$ since

$$v(1T) = (1v)T = (v)T = vT$$

for any $v \in V$.

Thus $\boldsymbol{L}_K(V,W)$ is a K-vector space. □

Let us assume now that V is an n-dimensional K-vector space and W is an m-dimensional K-vector space ($n,m \in \mathbb{N}$). Let $B = \{v_1, v_2, \ldots, v_n\}$ be a K-basis of V and $B' = \{w_1, w_2, \ldots, w_m\}$ be a K-basis of W. Then, for any linear transformation T in $\boldsymbol{L}_K(V,W)$, we have

$$\begin{aligned}v_1T &= \alpha_{11}w_1 + \alpha_{12}w_2 + \cdots + \alpha_{1m}w_m \\ v_2T &= \alpha_{21}w_1 + \alpha_{22}w_2 + \cdots + \alpha_{2m}w_m \qquad (*) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ v_nT &= \alpha_{n1}w_1 + \alpha_{n2}w_2 + \cdots + \alpha_{nm}w_m\end{aligned}$$

where $\alpha_{11}, \alpha_{12}, \ldots, \alpha_{nm}$ are scalars in K. The arrangement of the scalars in (*) deserves a name.

43.2 Definition: Let K be a field and $n,m \in \mathbb{N}$. An *n by m matrix over K* is an array

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & . \; . \; . & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & . \; . \; . & \alpha_{2m} \\ . & . & . \; . \; . & . \\ \alpha_{n1} & \alpha_{n2} & . \; . \; . & \alpha_{nm} \end{pmatrix} \tag{1}$$

of nm elements $\alpha_{11}, \alpha_{12}, \dots, \alpha_{nm}$ of K, arranged in n rows and m columns, and enclosed within parentheses. The set of all n by m matrices over K will be denoted by $Mat_{n\times m}(K)$.

Sometimes we write "$n \times m$" instead of "n by m". The horizontal lines

$$\alpha_{i1} \; \alpha_{i2} \dots \; \alpha_{im} \tag{2}$$

of a matrix over K are called the *rows* of that matrix. More specifically, (2) is the i-th row of the matrix (1). The vertical lines

$$\begin{matrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{nj} \end{matrix} \tag{3}$$

of a matrix over K are called the *columns* of that matrix. More specifically, (3) is the j-th column of the matrix (1). The element α_{ij} is at the place where the i-th row and the j-th column meet. The first index i refers to the row, the second index j refers to the column. Aso, in the expression "n by m", the first number n specifies the number of rows, the second number m specifies the number of columns of the matrix. The elements α_{ij} are called the *entries* of the matrix (1). When $n = m$, the matrix (1) is said to be a *square* matrix. The set of all square matrices with n rows (or n columns) over K will be denoted by $Mat_n(K)$ (instead of $Mat_{n\times n}(K)$).

We will usually abbreviate the matrix (1) as (α_{ij}).

Two matrices $(\alpha_{ij}) \in Mat_{n\times m}(K)$ and $(\beta_{ij}) \in Mat_{n\times m'}(K)$ are declared to be *equal* if $n = n'$, $m = m'$ and $\alpha_{ij} = \beta_{ij}$ for all $i = 1,2, \dots ,n$; $j = 1,2, \dots ,m$. Thus two matrices are equal if and only if they have the same number of rows and columns, and have the same elements at corresponding places. We write then $(\alpha_{ij}) = (\beta_{ij})$. Otherwise, we put $(\alpha_{ij}) \neq (\beta_{ij})$.

We now make $Mat_{n\times m}(K)$ into a vector space over K.

43.3 Definition: Let K be a field, $\alpha \in K$ and let $A,B \in Mat_{n\times m}(K)$, say $A = (\alpha_{ij})$, $B = (\beta_{ij})$. We write

$A + B = C$, C being the matrix (γ_{ij}) in $Mat_{n\times m}(K)$, where $\gamma_{ij} = \alpha_{ij} + \beta_{ij}$,

and $\alpha A = E$, E being the matrix $(\mathbb{E}_{ij})$ in $Mat_{n\times m}(K)$, where $\mathbb{E}_{ij} = \alpha\alpha_{ij}$.

In other words, $(\alpha_{ij}) + (\beta_{ij}) = (\alpha_{ij} + \beta_{ij})$ and $\alpha(\alpha_{ij}) = (\alpha\alpha_{ij})$.

43.4 Theorem: *Let K be a field. Under the addition and multiplication by scalars of Definition 43.3, the set $Mat_{n\times m}(K)$ is a vector space over K.*

Proof: First we check that $Mat_{n\times m}(K)$ is an abelian group under addition.

(i) For any $A = (\alpha_{ij})$, $B = (\beta_{ij})$ in $Mat_{n\times m}(K)$, we have $A + B = (\alpha_{ij} + \beta_{ij})$ and each $\alpha_{ij} + \beta_{ij}$ is an element of K, because K is closed under addition ($i = 1,2, \ldots ,n$; $j = 1,2, \ldots ,m$). Hence $A + B \in Mat_{n\times m}(K)$ and $Mat_{n\times m}(K)$ is closed under addition.

(ii) For any $A = (\alpha_{ij})$, $B = (\beta_{ij})$, $C = (\gamma_{ij})$ in $Mat_{n\times m}(K)$, there holds

$$\begin{aligned}(A + B) + C &= \big((\alpha_{ij}) + (\beta_{ij})\big) + (\gamma_{ij}) = (\alpha_{ij} + \beta_{ij}) + (\gamma_{ij}) = \big((\alpha_{ij} + \beta_{ij}) + \gamma_{ij}\big)\\ &= \big(\alpha_{ij} + (\beta_{ij} + \gamma_{ij})\big) = (\alpha_{ij}) + (\beta_{ij} + \gamma_{ij}) = (\alpha_{ij}) + \big((\beta_{ij}) + (\gamma_{ij})\big) = A + (B + C)\end{aligned}$$

and addition in $Mat_{n\times m}(K)$ is associative.

(iii) Let $\tilde{0}$ be the n by m matrix whose entries are all equal to the zero element of K. Thus $\tilde{0} = (\zeta_{ij})$, where $\zeta_{ij} = 0 \in K$ for all i,j. Then

$$A + \tilde{0} = (\alpha_{ij}) + (\zeta_{ij}) = (\alpha_{ij} + \zeta_{ij}) = (\alpha_{ij} + 0) = (\alpha_{ij}) = A$$

for any $A = (\alpha_{ij}) \in Mat_{n\times m}(K)$. So $\tilde{0} \in Mat_{n\times m}(K)$ and $\tilde{0}$ is a right identity of $Mat_{n\times m}(K)$.

(iv) For any $A = (\alpha_{ij}) \in Mat_{n\times m}(K)$, let $B = (-\alpha_{ij}) \in Mat_{n\times m}(K)$. Then $A + B = (\alpha_{ij}) + (-\alpha_{ij}) = \big(\alpha_{ij} + (-\alpha_{ij})\big) = \tilde{0}$. Hence every element $A = (\alpha_{ij})$ in $Mat_{n\times m}(K)$ has an inverse $(-\alpha_{ij})$ in $Mat_{n\times m}(K)$.

(v) For all $A = (\alpha_{ij})$, $B = (\beta_{ij}) \in Mat_{n\times m}(K)$, we have

$$A + B = (\alpha_{ij}) + (\beta_{ij}) = (\alpha_{ij} + \beta_{ij}) = (\beta_{ij} + \alpha_{ij}) = (\beta_{ij}) + (\alpha_{ij}) = B + A$$

and addition on $Mat_{n\times m}(K)$ is commutative.

This proves that $Mat_{n\times m}(K)$ is an abelian group under addition. Now the properties of multiplication by scalars. For any $\alpha,\beta \in K$ and $A = (\alpha_{ij})$, $B = (\beta_{ij}) \in Mat_{n\times m}(K)$, we have

$$\begin{aligned}
&(1) \qquad & \alpha(A + B) &= \alpha\big((\alpha_{ij}) + (\beta_{ij})\big) = \alpha(\alpha_{ij} + \beta_{ij}) = \big(\alpha(\alpha_{ij} + \beta_{ij})\big) \\
&&&= (\alpha\alpha_{ij} + \alpha\beta_{ij}) = (\alpha\alpha_{ij}) + (\alpha\beta_{ij}) \\
&&&= \alpha(\alpha_{ij}) + \alpha(\beta_{ij}) = \alpha A + \alpha B, \\
&(2) \qquad & (\alpha + \beta)A &= (\alpha + \beta)(\alpha_{ij}) = \big((\alpha + \beta)\alpha_{ij}\big) = (\alpha\alpha_{ij} + \beta\alpha_{ij}) \\
&&&= (\alpha\alpha_{ij}) + (\beta\alpha_{ij}) = \alpha(\alpha_{ij}) + \beta(\alpha_{ij}) = \alpha A + \beta A, \\
&(3) \qquad & (\alpha\beta)A &= (\alpha\beta)(\alpha_{ij}) = \big((\alpha\beta)\alpha_{ij}\big) = \big(\alpha(\beta\alpha_{ij})\big) = \alpha(\beta\alpha_{ij}) \\
&&&= \alpha\big(\beta(\alpha_{ij})\big) = \alpha(\beta A), \\
&(4) \qquad & 1A &= 1(\alpha_{ij}) = (1\alpha_{ij}) = (\alpha_{ij}) = A.
\end{aligned}$$

Thus $Mat_{n\times m}(K)$ is a vector space over K. □

A convenient K-basis of $Mat_{n\times m}(K)$ is described in the next lemma.

43.5 Lemma: *Let E_{ij} be the matrix in $Mat_{n\times m}(K)$ all of whose entries are* 0*, except for the single entry in the i-th row, j-th column, which entry is the identity element of K. Then the nm matrices E_{ij} (where $i = 1,2, \dots ,n$; $j = 1,2, \dots ,m$) form a K-basis of $Mat_{n\times m}(K)$. In particular, $dim_K\big(Mat_{n\times m}(K)\big)$ is equal to nm.*

Proof: The matrices E_{ij} span $Mat_{n\times m}(K)$ over K because any $A = (\alpha_{ij})$ in $Mat_{n\times m}(K)$ can be written as a K-linear combination

$$A = (\alpha_{ij}) = \sum_{i,j} \alpha_{ij}E_{ij}$$

of them. Moreover, matrices E_{ij} are linearly independent over K, for if α_{ij} are scalars such that

$$\sum_{i,j} \alpha_{ij}E_{ij} = \tilde{0},$$

then $(\alpha_{ij}) = \tilde{0}$

and $\alpha_{ij} = 0$ for all i,j. Therefore $\{E_{ij}\colon\ i = 1,2, \ldots, n;\ j = 1,2, \ldots, m\}$ is a K-basis of $Mat_{n\times m}(K)$. In particular, $dim_K\big(Mat_{n\times m}(K)\big) = nm$. □

We relate $\boldsymbol{L}(V,W)$ to $Mat_{n\times m}(K)$. This relation is implicit in (*). We state this relation as a definition and prove that $\boldsymbol{L}_K(V,W)$ and $Mat_{n\times m}(K)$ are isomorphic K-vector spaces.

43.6 Definition: Let V be an n-dimensional and W be an m-dimensional vector space over a field K, where $n,m \in \mathbb{N}$. Let $B = \{v_1, v_2, \ldots, v_n\}$ be a K-basis of V and $B' = \{w_1, w_2, \ldots, w_m\}$ be a K-basis of W.
Let T be a K-linear transformation in $\boldsymbol{L}_K(V,W)$ and let

$$v_iT = \sum_{j=1}^{m} \alpha_{ij}w_j \qquad (i = 1,2, \ldots, n), \qquad (*)$$

where $\alpha_{ij} \in K$.
The $n \times m$ matrix (α_{ij}) over K will be called the *matrix associated with T* (*relative to the bases B and B´*), and will be written $\mathrm{M}^B_{B'}(T)$.

In the following discussion, the bases will be fixed and we simply write $\mathrm{M}(T)$ instead of $\mathrm{M}^B_{B'}(T)$. The role of the bases will be discused at the end of this paragraph.

43.7 Theorem: *Let V be an n-dimensional and W be an m-dimensional vector space over a field K, where n,m* $\in \mathbb{N}$. *Then, for any T,S* $\in \boldsymbol{L}_K(V,W)$ *and* $\alpha \in K$, *we have*

$$\mathrm{M}(T + S) = \mathrm{M}(T) + \mathrm{M}(S) \qquad \textit{and} \qquad \mathrm{M}(\alpha T) = \alpha\mathrm{M}(T)$$

(all associated matrices are taken relative to the same pair of K-bases).
In other words, $\mathrm{M}\colon \boldsymbol{L}_K(V,W) \to Mat_{n\times m}(K)$ *is a K-linear transformation.*

Proof: Let $B = \{v_1, v_2, \ldots, v_n\}$ be the K-basis of V and $B' = \{w_1, w_2, \ldots, w_m\}$ be the K-basis of W relative to which the associated matrices are taken, so that $\mathrm{M}(T) = (\alpha_{ij})$ and $\mathrm{M}(S) = (\beta_{ij})$, where

$$v_iT = \sum_{j=1}^{m} \alpha_{ij}w_j \qquad \text{and} \qquad v_iS = \sum_{j=1}^{m} \beta_{ij}w_j.$$

Then $v_i(T + S) = v_iT + v_iS = \sum_{j=1}^{m} \alpha_{ij}w_j + \sum_{j=1}^{m} \beta_{ij}w_j = \sum_{j=1}^{m} (\alpha_{ij} + \beta_{ij})w_j$

and therefore $\mathrm{M}(T + S) = (\alpha_{ij} + \beta_{ij}) = (\alpha_{ij}) + (\beta_{ij}) = \mathrm{M}(T) + \mathrm{M}(S)$. Also

$$v_i(\alpha T) = (\alpha v_i)T = \alpha(v_iT) = \alpha \sum_{j=1}^{m} \alpha_{ij}w_j = \sum_{j=1}^{m} \alpha\alpha_{ij}w_j$$

and therefore $\mathrm{M}(\alpha T) = (\alpha\alpha_{ij}) = \alpha(\alpha_{ij}) = \alpha\mathrm{M}(T)$. □

43.8 Theorem: *Let V be an n-dimensional and W be an m-dimensional vector space over a field K, where $n,m \in \mathbb{N}$. Then $\boldsymbol{L}_K(V,W)$ is isomorphic to $Mat_{n\times m}(K)$ (as K-vector spaces). In particular, $\dim_K \boldsymbol{L}_K(V,W) = nm$.*

Proof: We know that $\mathrm{M}: \boldsymbol{L}_K(V,W) \to Mat_{n\times m}(K)$ (in the notation of Theorem 43.7) is a vector space homomorphism. We will prove that M is in fact an isomorphism.

We prove that M is one-to-one and onto. Let (γ_{ij}) be any matrix in $Mat_{n\times m}(K)$. We want to find a T in $\boldsymbol{L}_K(V,W)$ such that $\mathrm{M}(T) = (\gamma_{ij})$. Such a K-linear transformation T should satisfy

$$v_iT = \sum_{j=1}^{m} \gamma_{ij}w_j$$

and
$$\Big(\sum_{i=1}^{n} \alpha_iv_i\Big)T = \sum_{i=1}^{n} \alpha_i(v_iT) = \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{m} \gamma_{ij}w_j = \sum_{j=1}^{m} \Big(\sum_{i=1}^{n} \alpha_i\gamma_{ij}\Big)w_j$$

for any vector $v = \alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n$ in V. Thus there is at most one T in $\boldsymbol{L}_K(V,W)$ with $\mathrm{M}(T) = (\gamma_{ij})$. Hence M is one-to-one.

With the hindsight gained from the chain of equations above, given any (γ_{ij}) in $Mat_{n\times m}(K)$, we *define* a function $T: V \to W$ by

$$\Big(\sum_{i=1}^{n} \alpha_iv_i\Big)T = \sum_{j=1}^{m} \Big(\sum_{i=1}^{n} \alpha_i\gamma_{ij}\Big)w_j.$$

Then, for any $\alpha,\beta \in K$ and $v = \sum_{i=1}^{n} \alpha_iv_i$, $v' = \sum_{i=1}^{n} \beta_iv_i \in V$, we have

$$(\alpha v + \beta v')T = \Big(\alpha\sum_{i=1}^{n} \alpha_iv_i + \beta\sum_{i=1}^{n} \beta_iv_i\Big)T = \Big(\sum_{i=1}^{n} (\alpha\alpha_i + \beta\beta_i)v_i\Big)T$$

$$= \sum_{j=1}^{m} \Big(\sum_{i=1}^{n} (\alpha\alpha_i + \beta\beta_i)\gamma_{ij}\Big) w_j = \sum_{j=1}^{m} \Big(\alpha\sum_{i=1}^{n}\alpha_i\gamma_{ij} + \beta\sum_{i=1}^{n}\beta_i\gamma_{ij}\Big) w_j$$

$$= \alpha\sum_{j=1}^{m} \Big(\sum_{i=1}^{n}\alpha_i\gamma_{ij}\Big) w_j + \beta\sum_{j=1}^{m} \Big(\sum_{i=1}^{n}\beta_i\gamma_{ij}\Big) w_j = \alpha(vT) + \beta(v'T)$$

and T is K-linear. Thus $T \in \boldsymbol{L}_K(V,W)$. When we put $\alpha_{i_0} = 1$ and $\alpha_i = 0$ for $i \neq i_0$, we obtain

$$v_{i_0}T = \sum_{j=1}^{m} \gamma_{i_0 j} w_j, \qquad (i_0 = 1,2, \dots ,n)$$

so $M(T) = (\gamma_{ij})$. Thus every $(\gamma_{ij}) \in Mat_{n\times m}(K)$ is the image, under M, of at least one $T \in \boldsymbol{L}_K(V,W)$ and so M is onto. Consequently M is a vector space isomorphism: $\boldsymbol{L}_K(V,W) \cong Mat_{n\times m}(K)$. From Theorem 42.18 and Lemma 43.5, we get $dim_K\boldsymbol{L}_K(V,W) = dim_K Mat_{n\times m}(K) = nm$. □

Now let U be a vector space over the field K, with $dim_K U = k \in \mathbb{N}$, and let $B'' = \{u_1,u_2,\dots,u_k\}$ be a basis of U over K. If $T\colon V \to W$ and $S\colon W \to U$ are K-linear transformations, whose associated matrices [relative to the K-bases $B = \{v_1,v_2, \dots ,v_n\}$, $B' = \{w_1,w_2, \dots ,w_m\}$ of V and W, and relative to the K-bases B', B'' of W and U] are $(\alpha_{ij}) \in Mat_{n\times m}(K)$ and $(\beta_{jl}) \in Mat_{m\times k}(K)$, so that

$$v_iT = \sum_{j=1}^{m} \alpha_{ij} w_j, \qquad w_jT = \sum_{l=1}^{k} \beta_{jl} u_l,$$

then $TS\colon V \to U$ is a K-linear transformation (Theorem 41.7) and

$$v_i(TS) = (v_iT)S = \Big(\sum_{j=1}^{m} \alpha_{ij} w_j\Big)S = \sum_{j=1}^{m} \alpha_{ij}(w_jS)$$

$$= \sum_{j=1}^{m} \alpha_{ij} \sum_{l=1}^{k} \beta_{jl}u_l = \sum_{l=1}^{k} \Big(\sum_{j=1}^{m} \alpha_{ij}\beta_{jl}\Big) u_l$$

so that the matrix associated with TS [relative to the K-bases B,B''] is the $n \times k$ matrix whose i-th row, l-th column entry is $\sum_{j=1}^{m} \alpha_{ij}\beta_{jl}$. This leads us to the following definition.

43.9 Definition: Let $A = (\alpha_{ij})$ be an $n \times m$ matrix and let $B = (\beta_{ij})$ be an $m \times k$ matrix, with entries from a field K. Then the *product of A and B*, denoted by AB, is the $n \times k$ matrix (γ_{ij}) over K, where $\gamma_{ij} = \sum_{j=1}^{m} \alpha_{ij}\beta_{jl}$. Stated otherwise

$$(\alpha_{ij})(\beta_{ij}) := (\sum_{j=1}^{m} \alpha_{ij}\beta_{jl})$$

Before studying the properties of this matrix multiplication, we summarize the discussion preceding Definition 43.9. Although matrix multiplication is defined in such a way as to make it true, the following theorem is by no means obvious (cf. Remark 43.18).

43.10 Theorem: *Let V,W,U be vector spaces over a field K, of nonzero finite dimensions n,m,k, respectively. Let B, B', B'' be fixed K-bases of V,W,U, respectively. If, relative to these bases, $T \in \mathbf{L}_K(V,W)$ has the associated matrix A, and $S \in \mathbf{L}_K(W,U)$ has the associated matrix C, then $TS \in \mathbf{L}_K(V,U)$ has the associated matrix AC. Equivalently,*

$$\mathrm{M}(TS) = \mathrm{M}(T)\mathrm{M}(S). \qquad \square$$

The product of an $n \times m$ matrix by an $m \times k$ matrix is an $n \times k$ matrix. Notice that the number of columns in A has to be equal to the number of rows in B in order AB to make sense. The product of an $n \times m$ matrix by an $m' \times k$ matrix is *not* defined unless $m = m'$.

Matrix multiplication is associative whenever it is possible. That is to say, $(AB)C = A(BC)$ for any matrices A,B,C with entries from a field K, provided the sizes of A,B,C are such that the products AB and BC are defined (then $(AB)C$ and $A(BC)$ are defined, too). More precisely, if A is an $n \times m$ matrix, B is an $m \times k$ matrix and C is an $k \times s$ matrix, then the two $n \times s$ matrices $(AB)C$, $A(BC)$ are equal. To prove this, let us put $A = (\alpha_{ij})$, $B = (\beta_{jl})$, $C = (\gamma_{lr})$, where $i = 1,2,\ldots,n$; $j = 1,2,\ldots,m$; $r = 1,2,\ldots,s$. Then

$$AB = E = (\mathbb{E}_{il}), \qquad \text{where } \mathbb{E}_{il} = \sum_{j=1}^{m} \alpha_{ij}\beta_{jl}$$

$$BC = F = (\phi_{jr}), \qquad \text{where } \phi_{jr} = \sum_{l=1}^{k} \beta_{jl}\gamma_{lr}$$

and from $(AB)C = EC = \Big(\sum_{l=1}^{k} \mathbb{E}_{il}\gamma_{lr} \Big)$

$$= \Big(\sum_{l=1}^{k} \big(\sum_{j=1}^{m} \alpha_{ij}\beta_{jl} \big)\gamma_{lr} \Big) = \Big(\sum_{l=1}^{k} \sum_{j=1}^{m} (\alpha_{ij}\beta_{jl})\gamma_{lr} \Big),$$

$$A(BC) = AF = \Big(\sum_{j=1}^{m} \alpha_{ij}\phi_{jr} \Big) = \Big(\sum_{j=1}^{m} \alpha_{ij} \big(\sum_{l=1}^{k} \beta_{jl}\gamma_{lr} \big) \Big)$$

$$= \Big(\sum_{j=1}^{m} \sum_{l=1}^{k} \alpha_{ij}(\beta_{jl}\gamma_{lr}) \Big) = \Big(\sum_{l=1}^{k} \sum_{j=1}^{m} \alpha_{ij}(\beta_{jl}\gamma_{lr}) \Big)$$

$$= \Big(\sum_{l=1}^{k} \sum_{j=1}^{m} (\alpha_{ij}\beta_{jl})\gamma_{lr} \Big),$$

we conclude that $(AB)C = A(BC)$.

However, there is no hope for commutativity. For one thing, the product BA need not be defined even if the product AB happens to be defined. For instance, if A is a 2×3 and B is a 3×4 matrix, then AB is a 2×4 matrix, but BA is not even defined, let alone is equal to AB. But also in cases where both products AB and BA are defined, they will, generally speaking, have diferent sizes, so they will fail to be equal on dimension grounds. For instance, if A is a 2×3 matrix and B is a 3×2 matrix, then AB is a 2×2 matrix and BA is a 3×3 matrix and $AB \neq BA$, since a 2×2 matrix cannot be equal to a 3×3 matrix. Even if both AB and BA are defined and have the same size (this occurs only in case A and B are square matrices with the same number of rows), it usually happens that $AB \neq BA$. For example $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

Let I_m be the square matrix over K with m rows, whose entries are all equal to $0 \in K$, except for those on the main diagonal, which are all equal to $1 \in K$ (the main diagonal in any $n \times m$ matrix consists of the places where the i-th row and the i-th column intersect, ($i = 1,2, \ldots ,\min\{n,m\}$)). It is easily verified that $AI_m = A$ for any $A \in Mat_{n\times m}(K)$. Likewise $I_n A = A$ for any $A \in Mat_{n\times m}(K)$.

Let $0_{m\times k}$ be the $m \times k$ matrix over K all of whose entries are $0 \in K$. One checks easily that $A0_{m\times k} = 0_{n\times k}$ and $0_{k\times n}A = 0_{k\times m}$ for any $A \in Mat_{n\times m}(K)$.

Multiplication of matrices is distributive over addition. Indeed, for all $A = (\alpha_{ij}) \in Mat_{n\times m}(K)$ and $B = (\beta_{jl})$, $C = (\gamma_{jl}) \in Mat_{m\times k}(K)$, we have

$$A(B + C) = (\alpha_{ij})(\beta_{jl} + \gamma_{jl}) = \Bigl(\sum_{j=1}^{m} \alpha_{ij}(\beta_{jl} + \gamma_{jl})\Bigr) = \Bigl(\sum_{j=1}^{m} (\alpha_{ij}\beta_{jl} + \alpha_{ij}\gamma_{jl})\Bigr)$$

$$= \Bigl(\sum_{j=1}^{m} \alpha_{ij}\beta_{jl} + \sum_{j=1}^{m} \alpha_{ij}\gamma_{jl}\Bigr) = \Bigl(\sum_{j=1}^{m} \alpha_{ij}\beta_{jl}\Bigr) + \Bigl(\sum_{j=1}^{m} \alpha_{ij}\gamma_{jl}\Bigr) = AB + AC.$$

In like manner, one proves $(B + C)A = BA + CA$ for all $A \in Mat_{n\times k}(K)$ and $B,C \in Mat_{m\times n}(K)$.

One checks easily that, for all $\alpha \in K$, $A \in Mat_{n\times m}(K)$, $B \in Mat_{m\times k}(K)$

$$(\alpha A)B = \alpha(AB) = A(\alpha B). \qquad \text{(e)}$$

On writing $A = (\alpha_{ij})$, $B = (\beta_{jl})$, we get indeed

$$(\alpha A)B = (\alpha\alpha_{ij})(\beta_{jl}) = \Bigl(\sum_{j=1}^{m} (\alpha\alpha_{ij})\beta_{jl}\Bigr) = \Bigl(\sum_{j=1}^{m} \alpha(\alpha_{ij}\beta_{jl})\Bigr)$$

$$= \Bigl(\alpha\sum_{j=1}^{m} \alpha_{ij}\beta_{jl}\Bigr) = \alpha\Bigl(\sum_{j=1}^{m} \alpha_{ij}\beta_{jl}\Bigr) = \alpha(AB)$$

and

$$A(\alpha B) = (\alpha_{ij})(\alpha\beta_{jl}) = \Bigl(\sum_{j=1}^{m} \alpha_{ij}(\alpha\beta_{jl})\Bigr) = \Bigl(\sum_{j=1}^{m} \alpha(\alpha_{ij}\beta_{jl})\Bigr)$$

$$= \Bigl(\alpha\sum_{j=1}^{m} \alpha_{ij}\beta_{jl}\Bigr) = \alpha\Bigl(\sum_{j=1}^{m} \alpha_{ij}\beta_{jl}\Bigr) = \alpha(AB).$$

Let us now consider the set $Mat_n(K)$ of square matrices over a field K. From Theorem 43.3, we know that $Mat_n(K)$ is an abelian group under addition. The product of any two $n \times n$ matrices is an $n \times n$ matrix. Since the matrix multiplication is associative and distributive over addition, $Mat_n(K)$ is a ring. We also know that $AI_n = I_nA = A$ for any $n \times n$ matrix A. Thus we proved

43.11 Theorem: *Let K be a field and $n \in \mathbb{N}$. Then, under matrix addition and matrix multiplication, $Mat_n(K)$ is a ring with identity I_n.*

□

The counterpart of Theorem 43.11 for linear transformations is also valid.

43.12 Theorem: *Let V be a vector space over a field K and let $\mathbf{L}_K(V,V)$ be the set of all K-linear mappings from V into V. Then, under the pointwise addition and composition of K-linear transformations, $\mathbf{L}_K(V,V)$ is a ring with identity. The identity mapping $\iota: V \to V$ is the identity element of this ring $\mathbf{L}_K(V,V)$.* [Notice that there is no hypothesis about $dim_K V$.]

Proof: We must check the ring axioms. From Theorem 43.1, we know that $\mathbf{L}_K(V,V)$ is an abelian group under addition. Also, (1) $\mathbf{L}_K(V,V)$ is closed under the composition of mappings (Theorem 41.7), and (2) composition of mappings (whether K-linear or not) is associative (Theorem 3.10), and (D) composition is distributive over addition: when T,S,R are arbitrary elements of $\mathbf{L}_K(V,V)$, then

$$v\big(T(S+R)\big) = (vT)(S+R) = \big((vT)S\big) + \big((vT)R\big) = \big(v(TS)\big) + \big(v(TR)\big) = v(TS+TR)$$

and $$v\big((S+R)T\big) = \big(v(S+R)\big)T = (vS+vR)T = (vS)T + (vR)T = v(ST) + v(RT) = v(ST+RT)$$

for all $v \in V$, hence $T(S+R) = TS + TR$ and $(S+R)T = ST + RT$. So $\mathbf{L}_K(V,V)$ is a ring. Finally, the identity mapping ι is clearly a K-linear transformation, so $\iota \in \mathbf{L}_K(V,V)$ and as $T\iota = T = \iota T$ for all $T \in \mathbf{L}_K(V,V)$, we conclude that $\mathbf{L}_K(V,V)$ is a ring with identity ι. □

43.13 Theorem: *Let V be a vector space over a field K with $dim_K V = n$, where $n \in \mathbb{N}$. Then $\mathbf{L}_K(V,V) \cong Mat_n(K)$ (ring isomorphism).*

Proof: We fix a K-basis of V and use the mapping $\mathrm{M}: \mathbf{L}_K(V,V) \to Mat_n(K)$ of Theorem 43.7, so that $\mathrm{M}(T)$ is the associated matrix of the K-linear transformation $T \in \mathbf{L}_K(V,V)$. By Theorem 43.8, M is an isomorphism of abelian groups (in fact of K-vector spaces, but we do not need this now) and by Theorem 43.10, M preserves multiplication as well. Hence M is a ring isomorphism. □

Let us recall that a unit in a ring with identity is an element of that ring possessing a (unique) right inverse which is also a left inverse. What are the units of $\boldsymbol{L}_K(V,V)$? The units in $\boldsymbol{L}_K(V,V)$ are, by definition, those K-linear transformations T with the inverse T^{-1} in $\boldsymbol{L}_K(V,V)$. The inverse of T in $\boldsymbol{L}_K(V,V)$, whenever it exists, is in $\boldsymbol{L}_K(V,V)$ by Lemma 41.10(2). Thus the units of $\boldsymbol{L}_K(V,V)$ are the K-linear transformations in $\boldsymbol{L}_K(V,V)$ which are one-to-one and onto: the units of $\boldsymbol{L}_K(V,V)$ are the vector space isomorphisms from V onto V. The set of all isomorphisms from V onto V will be denoted by $GL(V)$. This is a group under the composition of mappings, called the *general linear group of* V. Thus $\boldsymbol{L}_K(V,V)^{\times} = GL(V)$.

The units in $Mat_n(K)$ are the invertible matrices, that is to say, matrices A in $Mat_n(K)$ for which an $A^{-1} \in Mat_n(K)$ exists such that $AA^{-1} = \boldsymbol{I}_n = A^{-1}A$. These are the matrices associated with isomorphisms from V onto V. In the next paragraph, we will give a necessary and sufficient condition for a matrix to be invertible (Theorem 44.20). The set of all invertible matrices in $Mat_n(K)$ will be denoted by $GL(n,K)$. This is a group under the multiplication of matrices, called the *general linear group of degree n over K*. Thus $Mat_n(K)^{\times} = GL(n,K)$. When V is an n-dimensional vector space over K, the group $GL(V)$ is isomorphic to the group $GL(n,K)$.

We return to the more general case of $\boldsymbol{L}_K(V,W)$ and $Mat_{n\times m}(K)$. Suppose again that V and W are K-vector spaces of K-dimensions n and m, respectively, where $n,m \in \mathbb{N}$. Let $B = \{v_1, v_2, \ldots, v_n\}$ and $B^* = \{v_1^*, v_2^*, \ldots, v_n^*\}$ be K-bases of V and let $B' = \{w_1, w_2, \ldots, w_m\}$ and $B'^* = \{w_1^*, w_2^*, \ldots, w_m^*\}$ be K-bases of W. With each K-linear transformation $T: V \rightarrow W$, there is associated a matrix $M_{B'}^{B}(T)$ relative to the bases B and B', and a matrix $M_{B'^*}^{B^*}(T)$ relative to the bases B^* and B'^*. We want to study the relationship between $M_{B'}^{B}(T)$ and $M_{B'^*}^{B^*}(T)$.

We recall that $M_{B'}^{B}(T) = (\alpha_{ij})$ and $M_{B'^*}^{B^*}(T) = (\beta_{ij})$, where

$$v_i T = \sum_{j=1}^{m} \alpha_{ij} w_j \qquad \text{and} \qquad v_i^* T = \sum_{j=1}^{m} \beta_{ij} w_j^*. \qquad (i = 1,2, \ldots ,n)$$

We introduce transition matrices which describe the change of bases. Writing

$$v_i = \sum_{k=1}^{n} \gamma_{ik} v_k^* \qquad (i = 1,2, \ldots ,n)$$

we obtain a matrix (γ_{ik}) in $Mat_n(K)$, called the *transition matrix from the K-basis B to the K-basis B* of V*. Of course, $(\gamma_{ik}) = \mathrm{M}_{B^*}^{B}(\iota)$, where ι is the identity mapping on V. We have the schema

$$\begin{array}{ccccccl} & \iota & & \iota & & & \text{mappings} \\ V & \rightarrow & V & \rightarrow & V & & \text{vector spaces} \\ B^* & & B & & B^* & & \text{bases} \\ & \mathrm{M}_{B}^{B^*}(\iota) & & \mathrm{M}_{B^*}^{B}(\iota) & & & \text{matrices} \end{array}$$

Now the composition $\iota\iota$ is the identity mapping ι. Relative to the bases B^* and B^*, the matrix associated with ι is the identity matrix I_n. This matrix is also equal to $\mathrm{M}_{B}^{B^*}(\iota)\mathrm{M}_{B^*}^{B}(\iota)$ by Theorem 43.10: $\mathrm{M}_{B}^{B^*}(\iota)\mathrm{M}_{B^*}^{B}(\iota) = I_n$. Thus the transition matrix from B to B^* is the inverse of the transition matrix from B^* to B.

$$\begin{array}{ccc} & \iota & \\ V & \rightarrow & V \\ B & & B^* \\ & \mathrm{M}_{B^*}^{B}(\iota) & \end{array} \qquad\qquad \begin{array}{ccc} & \iota & \\ V & \rightarrow & V \\ B^* & & B \\ & \mathrm{M}_{B}^{B^*}(\iota) = (\mathrm{M}_{B^*}^{B}(\iota))^{-1} & \end{array}$$

43.14 Theorem: *With the foregoing notation, let P be the transition matrix from B to B*, and let Q be the transition matrix from B´ to B´*. If* $T\colon V \rightarrow W$ *is any K-linear mapping, then the matrices* $\mathrm{M}_{B'}^{B}(T)$ *and* $\mathrm{M}_{B'^*}^{B^*}(T)$ *are connected by* $\mathrm{M}_{B'^*}^{B^*}(T) = P^{-1}\mathrm{M}_{B'}^{B}(T)Q$.

Proof: We have the following schema

$$\begin{array}{ccc} & T & \\ V & \rightarrow & W \\ B & & B' \\ & \mathrm{M}_{B'}^{B}(T) & \end{array}$$

The K-linear transformation T can be desribed also as follows.

$$\begin{array}{ccccccc} & \iota_V & & T & & \iota_W & \\ V & \rightarrow & V & \rightarrow & W & \rightarrow & W \\ B^* & & B & & B' & & B'^* \\ & \mathrm{M}_{B}^{B^*}(\iota_V) & & \mathrm{M}_{B'}^{B}(T) & & \mathrm{M}_{B'^*}^{B'}(\iota_W) & \end{array}$$

Now $\mathrm{M}_{B}^{B*}(\iota_V) = [\mathrm{M}_{B*}^{B}(\iota_V)]^{-1} = P^{-1}$ and $\mathrm{M}_{B'*}^{B'}(\iota_W) = Q$. By Theorem 43.10, the matrix associated with T relative to the bases $B*$ and $B'*$ is $P^{-1}\mathrm{M}_{B'}^{B}(T)Q$. Hence $\mathrm{M}_{B'*}^{B*}(T) = P^{-1}\mathrm{M}_{B'}^{B}(T)Q$.

$$\begin{array}{ccccccc} & \iota_V & & T & & \iota_W & \\ V & \to & V & \to & W & \to & W \\ B* & & B & & B' & & B'* \\ & P^{-1} & & \mathrm{M}_{B'}^{B}(T) & & Q & \end{array}$$

□

43.15 Theorem: *Let V be an n-dimensional vector space over a field K, B and $B*$ be K-bases of V, and let P be the transition matrix from B to $B*$. Suppose T is any K-linear mapping from V into V. If* $\mathrm{M}(T)$ *is the matrix associated with T relative to the bases B and B, and if* $\mathrm{M}^*(T)$ *is the matrix associated with T relative to the bases $B*$ and $B*$, then*

$$\mathrm{M}^*(T) = P^{-1}\mathrm{M}(T)P.$$

Proof: This is a special case of Theorem 43.14. Using the diagram

$$\begin{array}{ccccccc} & \iota & & T & & \iota & \\ V & \to & V & \to & V & \to & V \\ B* & & B & & B & & B'* \\ & P^{-1} & & \mathrm{M}(T) & & P & \end{array}$$

the proof follows immediately from Theorem 43.10. □

43.16 Definition: Let K be a field and let $A = (\alpha_{ij})$ be an $n \times m$ matrix with entries from K. Then the $m \times n$ matrix (β_{ij}), where $\beta_{ij} = \alpha_{ji}$ for all i,j, is called the *transpose of A*, and is written A^t.

Hence A^t is obtained from A by changing rows to columns and columns to rows. For instance, the transpose of $\begin{pmatrix} 0 & 1 & 2 \\ 1 & 3 & 5 \end{pmatrix}$ is $\begin{pmatrix} 0 & 1 \\ 1 & 3 \\ 2 & 5 \end{pmatrix}$. It follows from the definition that $(A^t)^t = A$ for any matrix A.

43.17 Lemma: *Let K be a field and let $A,B \in Mat_{n\times m}(K)$, $C \in Mat_{m\times k}(K)$, $\alpha \in K$. Then*

$$(A+B)^t = A^t + B^t, \qquad (\alpha A)^t = \alpha(A^t), \qquad (AC)^t = C^tA^t.$$

Proof: Let $A = (\alpha_{ij})$, $B = (\beta_{ij})$ and $C = (\gamma_{jl})$, where $i = 1,2,\dots,n$; $j = 1,2,\dots,m$ and $l = 1,2,\dots,k$. Let us put $A^t = (\alpha'_{ij})$, $B^t = (\beta'_{ij})$, $C^t = (\gamma'_{ij})$ and $\alpha A = (\phi_{ij})$. Then $\alpha'_{ij} = \alpha_{ji}$, $\beta'_{ij} = \beta_{ji}$, $\gamma'_{ij} = \gamma_{ji}$, and $\phi_{ij} = \alpha\alpha_{ij}$ for all i,j. So

$$\begin{aligned}(A+B)^t &= \big((\alpha_{ij}) + (\beta_{ij})\big)^t = (\alpha_{ij} + \beta_{ij})^t \\ &= \text{matrix whose } i\text{-th row, } j\text{-th column entry is } \alpha_{ji} + \beta_{ji} \\ &= (\text{matrix whose } i\text{-th row, } j\text{-th column entry is } \alpha_{ji}) \\ &\qquad + (\text{matrix whose } i\text{-th row, } j\text{-th column entry is } \beta_{ji}) \\ &= (\alpha'_{ij}) + (\beta'_{ij}) = A^t + B^t,\end{aligned}$$

$$\begin{aligned}(\alpha A)^t = (\phi_{ij})^t &= \text{matrix whose } i\text{-th row, } j\text{-th column entry is } \phi_{ji} \\ &= \text{matrix whose } i\text{-th row, } j\text{-th column entry is } \alpha\alpha_{ji} \\ &= \alpha(\text{matrix whose } i\text{-th row, } j\text{-th column entry is } \alpha_{ji}) \\ &= \alpha(\alpha'_{ij}) = \alpha A^t,\end{aligned}$$

$$\begin{aligned}(AC)^t &= \text{matrix whose } l\text{-th row, } i\text{-th column entry is the } i\text{-th} \\ &\qquad\qquad \text{row, } l\text{-th column entry in } AC \\ &= \text{matrix whose } l\text{-th row, } i\text{-th column entry is } \sum_j^m \alpha_{ij}\gamma_{jl} \\ &= \text{matrix whose } l\text{-th row, } i\text{-th column entry is } \sum_j^m \gamma_{jl}\alpha_{ij} \\ &= \text{matrix whose } l\text{-th row, } i\text{-th column entry is } \sum_j^m \gamma'_{ij}\alpha'_{ij} \\ &= (\text{matrix whose } l\text{-th row, } j\text{-th column entry is } \gamma'_{ij})\text{times} \\ &\qquad (\text{matrix whose } j\text{-th row, } i\text{-th column entry is } \alpha'_{ij}) \\ &= (\gamma'_{ij})(\alpha'_{ij}) = C^tA^t.\end{aligned}$$

□

43.18 Remark: The results in this paragraph are very natural. All operations discussed here are natural, and the vector spaces and rings of this paragraph arise naturally. Another natural item is the isomorphism in Theorem 43.10.

There is, however, a subtle point here. Theorem 43.10 is true only because we write the functions on the right of the elements on which they act! If we had written them on the left, Theorem 43.10 would read: M(TS) = M(S)M(T). Of course this is not as good as M(TS) = M(T)M(S). For this reason, people who write functions on the left define the associated matrices differently. If $T \in \boldsymbol{L}_K(V,W)$ and $v_iT = \sum_{j=1}^{m} \alpha_{ij} w_j$ as in Definition 43.6, they define the matrix associated with T (relative to the fixed K-bases $\{v_1, v_2, \dots, v_n\}$ of V and $\{w_1, w_2, \dots, w_m\}$ of W) to be $(\alpha_{ij})^{\mathrm{t}}$. Thus their $\mathrm{M}(T)$ is our $\mathrm{M}(T)^{\mathrm{t}}$, and

$$\text{their } \mathrm{M}(TS) = \text{their } \mathrm{M}(\text{first } S, \text{ then } T) = \text{our } \mathrm{M}(\text{first } S, \text{ then } T)^{\mathrm{t}} = \text{our}$$
$$\mathrm{M}(ST)^{\mathrm{t}} = \text{our } \big(\mathrm{M}(S)\mathrm{M}(T)\big)^{\mathrm{t}} = \text{our } \mathrm{M}(T)^{\mathrm{t}}\mathrm{M}(S)^{\mathrm{t}} = \text{their } \mathrm{M}(T)\mathrm{M}(S)$$

so that Theorem 43.10 is true in their notation, too. In some books, the forming of the transpose is included in the notation for the associated matrix. More clearly, some people write

$$v_iT = \sum_{j=1}^{m} \alpha_{ji} w_j$$

and define the matrix associated with T to be (α_{ji}). Then $\mathrm{M}(TS) = \mathrm{M}(T)\mathrm{M}(S)$ as before, but the equations above are not very sensible, for α_{ji} depends primarily on i and secondarily on j, so the indices occupy wrong places. In our notation, there is no need for the artificial transpositions, nor do we write the indices in the wrong order.

Exercises

1. Compute $A + B$ and AB when

$$A = \begin{pmatrix} 0 & 1 & 3 & 4 \\ -2 & 4 & -2 & 1 \\ 1 & 0 & 3 & 0 \\ 6 & 1 & -1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 7 & 1 & -3 \\ 1 & 0 & 3 & -4 \\ 2 & 0 & 2 & 1 \\ 0 & -2 & 1 & 0 \end{pmatrix} \text{ in } Mat_4(\mathbb{R}),$$

and when

$$A = \begin{pmatrix} \bar{5} & \bar{3} & \bar{2} & \bar{4} \\ \bar{3} & \bar{4} & \bar{2} & \bar{1} \\ \bar{1} & \bar{2} & \bar{3} & \bar{1} \\ \bar{0} & \bar{4} & \bar{6} & \bar{1} \end{pmatrix}, \qquad B = \begin{pmatrix} \bar{1} & \bar{2} & \bar{5} & \bar{1} \\ \bar{0} & \bar{3} & \bar{3} & \bar{6} \\ \bar{2} & \bar{6} & \bar{1} & \bar{4} \\ \bar{0} & \bar{4} & \bar{5} & \bar{0} \end{pmatrix} \text{ in } Mat_4(\mathbb{Z}_7).$$

2. Let K be a field and $A,B \in Mat_n(K)$ with $AB = BA$. Prove that $(A + B)^2 = A^2 + 2AB + B^2$ and $(A + B)(A - B) = A^2 - B^2$. Show that these equations need not hold if $AB \neq BA$.

3. Evaluate $A, A^2, A^3, \dots$, where A is given by

$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ and by $A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$. Generalize to square matrices of n rows.

4. Evaluate $A, A^2, A^3, \ldots$, where A is given by
$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ and by $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$. Generalize to square matrices of n rows.

4. The *trace* of a matrix $A = (\alpha_{ij}) \in Mat_n(K)$ is defined to be the sum of the entries on the main diagonal of A, denoted by $tr(A)$, so that $tr(A) = \alpha_{11} + \alpha_{22} + \ldots + \alpha_{nn}$. Prove that $tr(A) = tr(A^t)$, that $tr(AB) = tr(BA)$ and that $tr(C^{-1}AC) = tr(A)$ for any $A,B \in Mat_n(K)$, $C \in GL(n,K)$.

6. Let V,W be vector spaces over a field K, let $\{v_i: i \in \boldsymbol{I}\}$ be a K-basis of V and let $T,S \in \boldsymbol{L}_K(V,W)$. Prove that $T = S$ if and only if $v_iT = v_iS$ for all $i \in \boldsymbol{I}$.

7. Let V be a vector space over a field K, with $dim_K V = n \in \mathbb{N}$. Prove that $GL(V)$ is isomorphic to $GL(n,K)$.

8. Let $\varphi: \mathbb{R}^3 \to \mathbb{R}^3$ be the $\mathbb{R}$-linear mapping for which $u_1\varphi = (1,0,2)$, $u_2\varphi = (0,1,1)$, $u_3\varphi = (1,0,1)$, where, as usual, $u_1 = (1,0,0)$, $u_2 = (0,1,0)$, $u_3 = (0,0,1)$. We put $v_1 = (-1,1,0)$, $v_2 = (1,2,3)$, $v_3 = (0,1,2)$. Let $B = \{u_1,u_2,u_3\}$, and $B^* = \{v_1,v_2,v_3\}$. Show that B^* is an $\mathbb{R}$-basis of $\mathbb{R}^3$ and find the matrix of the $\mathbb{R}$-linear transformation φ relative to the bases (a) B and B; (b) B and B^*; (c) B^* and B; (d) B^* and B^*.