Proposition. Let $V$, $W$, and $Z$ be vector spaces, and let $T \in \mathcal{L}(V,W)$ and $S \in \mathcal{L}(W,Z)$. Then the composition $S \circ T : V \to Z$ is linear.

For linear functions $T$ and $S$ as above, we usually write $ST$ rather than $S \circ T$.

Proof. For $x, y \in V$ and $c \in \mathbb{R}$ we have

$$(ST)(x + y) = S(T(x) + T(y)) = S(T(x)) + S(T(y)) = (ST)(x) + (ST)(y)$$

and

$$(ST)(cx) = S(T(cx)) = S(cT(x)) = cS(T(x)) = c(ST)(x).$$

QED

Example. Let $V$ and $W$ be vector spaces, let $\{u_1, \ldots, u_n\}$ be a basis of $V$, and let $v_1, \ldots, v_n \in W$. Then the unique $T \in \mathcal{L}(V,W)$ taking the $u_i$’s to the corresponding $v_i$’s is the composition of the following two functions: the coordinate vector function

$$x \mapsto [x]_E : V \to \mathbb{R}^n$$

followed by the “linear combination function”

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \mapsto \sum_{i=1}^n c_i u_i : \mathbb{R}^n \to W.$$

Let $V$ be a vector space and $T \in \mathcal{L}(V)$. We write

$$T^2 = TT,$$

and inductively we define

$$T^n = T(T^{n-1}) \quad \text{for } n = 3, 4, \ldots.$$

We also write $T^1 = T$ and $T^0 = I$.

Example. Let $V$ be the vector space of polynomials of degree at most $n$, and let $D \in \mathcal{L}(V)$ be the differentiation operator: $D(f) = f'$. Then $D^2$ is the 2nd-derivative operator:

$$D^2(f) = f'',$$

etc. We have $D^{n+1} = 0$.

Proposition. Let $U$, $V$, $W$, and $Z$ be vector spaces. Then:

1. For all $T \in \mathcal{L}(V,W)$ we have

$$I_Z T = TI_U = T, \quad 0T = 0, \quad \text{and} \quad T0 = 0.$$
2. For all \( T, S \in L(V, W) \), \( P \in L(U, V) \), and \( Q \in L(W, Z) \) we have
\[
(T + S)P = TP + SP \quad \text{and} \quad Q(T + S) = QT + QS.
\]
3. For all \( P \in L(U, V) \), \( T \in L(V, W) \), and \( S \in L(W, Z) \) we have
\[
(ST)P = S(TP).
\]
4. For all \( c \in \mathbb{R} \), \( T \in L(V, W) \), and \( S \in L(W, Z) \) we have
\[
c(ST) = (cS)T = S(cT).
\]

Note that in item 1, the first 0 is the zero function from \( W \) to \( Z \), the second is the zero function from \( V \) to \( Z \), the third is the zero function from \( U \) to \( V \), and the fourth is the zero function from \( U \) to \( W \). The notation for the identity functions is self-explanatory, but in the future we will often drop the subscript, so the vector space on which the identity function \( I \) operates will have to be determined from the context.

**Proof.** Exercise. \( \text{QED} \)

**Example.** As the above proposition shows, composition of linear functions should be regarded as a kind of multiplication. But it does not satisfy all the properties of multiplication of real numbers. For example, it is noncommutative: define \( T, S \in L(\mathbb{R}^2) \) by
\[
T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} \quad \text{and} \quad S \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}.
\]

Then
\[
(TS) \begin{pmatrix} x \\ y \end{pmatrix} = T \left( S \begin{pmatrix} x \\ y \end{pmatrix} \right) = T \begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix}
\]
and
\[
(ST) \begin{pmatrix} x \\ y \end{pmatrix} = S \left( T \begin{pmatrix} x \\ y \end{pmatrix} \right) = S \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\]
In particular, \((TS) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\) and \((ST) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}\), so \(TS \neq ST\).

Another interesting phenomenon exhibited by this example: we have \( S \neq 0 \) and \( T \neq 0 \) but \( ST = 0 \). This also shows that there is no “cancellation law” for composition of linear functions: we can have \( T \neq 0 \) and \( ST = RT \) but \( S \neq R \) (in this case with \( R = 0 \)).

**Matrix multiplication.**

**Proposition.** Let
\[
E = \{u_1, \ldots, u_n\}, \quad F = \{v_1, \ldots, v_m\}, \quad \text{and} \quad G = \{w_1, \ldots, w_l\}
\]
be bases for vector spaces \( V, W, \) and \( Z \), respectively, and let \( T \in L(V, W) \) and \( S \in L(W, Z) \). Also let \( B, A, \) and \( C \) be the matrices representing \( T \) relative to \( E \) and \( F \), \( S \) relative to \( F \).
and $G$, and $ST$ relative to $E$ and $G$, respectively. Then for each $i = 1, \ldots, n$ and $j = 1, \ldots, l$ we have

$$C_{ij} = \sum_{k=1}^{m} A_{ik} B_{kj}.$$  

**Proof.** This follows from the following computation: for each $j = 1, \ldots, n$ we have

$$(ST)(u_j) = S(T(u_j))$$

$$= S\left(\sum_{k=1}^{m} B_{kj} v_k\right)$$

$$= \sum_{k=1}^{m} B_{kj} S(v_k)$$

$$= \sum_{k=1}^{m} B_{kj} \sum_{i=1}^{l} A_{ik} w_i$$

$$= \sum_{i=1}^{l} \left(\sum_{k=1}^{m} A_{ik} B_{kj}\right) w_i.$$  

QED

We use the above proposition to motivate the definition of matrix multiplication: for $A \in M_{m \times n}$ and $B \in M_{n \times l}$ we define the product $AB$ to be the $m \times l$ matrix $C$ with entries

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \quad \text{for } i = 1, \ldots, m, j = 1, \ldots, l.$$  

Thus, matrix multiplication is defined precisely so that, if matrices $A$ and $B$ represent composable linear functions $S$ and $T$, then the matrix product $AB$ represents the composition $ST$.

**Example.** Define $T \in L(\mathbb{R}^3, \mathbb{R})$ by $T \left(\begin{array}{c} x \\ y \\ z \end{array}\right) = 2x - 3y + z$ and $S \in L(\mathbb{R}, \mathbb{R}^2)$ by $S(t) = \left(\begin{array}{c} 4t \\ -3t \end{array}\right)$. Relative to the standard bases, the matrices representing $T$ and $S$ are

$$A = \begin{pmatrix} 2 & -3 & 1 \\ \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 4 \\ -3 \end{pmatrix},$$  

respectively. Thus the matrix representing the composition $ST \in L(\mathbb{R}^3, \mathbb{R}^2)$ is

$$BA = \begin{pmatrix} 4 \\ -3 \end{pmatrix} \begin{pmatrix} 2 & -3 & 1 \\ \end{pmatrix} = \begin{pmatrix} 8 & -12 & 4 \\ -6 & 9 & -3 \end{pmatrix}.$$
Alternatively, we can first compose the functions $T$ and $S$:

$$(ST) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = S \left( T \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right) = S(2x - 3y + z) = \begin{pmatrix} 8x - 12y + 4z \\ -6x + 9y - 3z \end{pmatrix},$$

and then compute

$$(ST) \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 8 \\ -6 \end{pmatrix}, \quad (ST) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -12 \\ 9 \end{pmatrix}, \quad \text{and} \quad (ST) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -3 \end{pmatrix},$$

so the matrix representing $ST$ is

$$\begin{pmatrix} 8 & -12 & 4 \\ -6 & 9 & -3 \end{pmatrix}.$$ 

Clearly the first method involved less work.

Note that for $A \in M_{m \times n}$ and $B \in M_{n \times l}$ the $ij$-entry of $AB$ can be regarded as the product of the $i$th row matrix of $A$ and the $j$th column matrix of $B$:

$$(AB)_{ij} = (A_{i1} \quad A_{i2} \quad \cdots \quad A_{in}) \begin{pmatrix} B_{1j} \\ B_{2j} \\ \vdots \\ B_{nj} \end{pmatrix} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$ 

Also, the $j$th column of $AB$ is the product of $A$ and the $j$th column of $B$:

$$\begin{pmatrix} (AB)_{1j} \\ \vdots \\ (AB)_{mj} \end{pmatrix} = A \begin{pmatrix} B_{1j} \\ \vdots \\ B_{mj} \end{pmatrix}.$$ 

Equivalently, writing $B$ in terms of its columns:

$$B = \begin{pmatrix} b_1 & \cdots & b_l \end{pmatrix},$$ 

where each $b_i$ is an $n \times 1$ matrix, we have

$$AB = \begin{pmatrix} Ab_1 & \cdots & Ab_l \end{pmatrix}.$$ 

Similarly for the rows: writing $A$ in terms of its row matrices:

$$A = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix},$$
where each $a_j$ is a $1 \times n$ matrix, we have

$$AB = \begin{pmatrix} a_1 B \\ \vdots \\ a_m B \end{pmatrix},$$

i.e., the $i$th row of $AB$ is the $i$th row of $A$ times $B$.

Suppose $A$ is $m \times n$ and $x$ is $n \times 1$, i.e., $x \in \mathbb{R}^n$. Write $A$ in terms of its columns:

$$A = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$$

and write $x = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. Then

$$Ax = \sum_{i=1}^n c_i a_i,$$

the linear combination of the columns of $A$ whose coefficients are the coordinates of the vector $x$. Similarly for rows. Thus, for any multipliable matrices $A$ and $B$, the columns of $AB$ are linear combinations of the columns of $A$, and the rows of $AB$ are linear combinations of the rows of $B$.

**Example.** Let’s illustrate the above remarks with the matrices

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 3 & 2 \\ -2 & 0 \end{pmatrix}.$$ 

We have

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ -2 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 1 & 6 \end{pmatrix}.$$ 

The 21-entry of $AB$ is

$$(3 \ 4) \begin{pmatrix} 3 \\ -2 \end{pmatrix} = 1.$$ 

The 1st column of $AB$ is

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 3 \end{pmatrix} - 2 \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$ 

The 2nd row of $AB$ is

$$(3 \ 4) \begin{pmatrix} 3 & 2 \\ -2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 6 \end{pmatrix} = 3 \begin{pmatrix} 3 \\ 2 \end{pmatrix} + 4 \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$$ 

Because of the correspondence between linear functions and matrices, we immediately deduce the following properties of matrix multiplication:

**Proposition.** For all matrices where the operations are defined, we have:

1. $IA = AI = A$;
2. $0A = 0$ and $A0 = 0$;
3. $(A + B)C = AC + BC$;
4. $A(B + C) = AB + AC$;
5. $(AB)C = A(BC)$;
6. $t(AB) = (tA)B = A(tB)$ (where $t \in \mathbb{R}$).

**Theorem.** Let $V$ and $W$ be finite-dimensional vector spaces, with bases $E$ and $F$, respectively, let $T \in L(V, W)$, and let $x \in V$. Let $A$ be the matrix representing $T$ relative to $E$ and $F$. Then for all $x \in V$ we have

$$[T(x)]_F = A[x]_E.$$ 

**Proof.** Let $E = \{u_1, \ldots, u_n\}$. Since both sides are linear functions of $x$, it suffices to verify the equation for $x = u_1, \ldots, u_n$. For each $j = 1, \ldots, n$, if $x = u_j$, the left-hand side is the $j$th column of $A$. On the right-hand side, we have $[u_j]_E = e_j$ (the $j$th standard basis vector of $\mathbb{R}^n$), and by the properties of matrix multiplication the product $Ae_j$ is the $j$th column of $A$.

**QED**

**Example.** Let $V$ be the vector space of polynomials of degree at most 3, and define $T: V \to \mathbb{R}^2$ by

$$T(f) = \begin{pmatrix} f(0) \\ f(2) \end{pmatrix}.$$ 

Let $E$ and $F$ be the standard bases of $V$ and $\mathbb{R}^2$, respectively. We have

$$T(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad T(x) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad T(x^2) = \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \quad \text{and} \quad T(x^3) = \begin{pmatrix} 0 \\ 8 \end{pmatrix},$$

so the matrix representing $T$ relative to the bases $E$ and $F$ is

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 \end{pmatrix}.$$ 

Let $f = 3 - 2x + x^3$. Then

$$[f]_E = \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \end{pmatrix},$$

so we have

$$A[f]_E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 4 & 8 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}.$$ 

On the other hand,

$$f(0) = 3 \quad \text{and} \quad f(2) = 3 - 2(2) + 2^3 = 7,$$
so by direct calculation we have

\[ T(f) = \begin{pmatrix} 3 \\ 7 \end{pmatrix}. \]

Since the vector \((\frac{3}{7})\) \(\in\) \(\mathbb{R}^2\) is equal to its coordinate vector relative to the standard basis \(F\), we have verified the above theorem in this case.

**Example.** Let \(A \in M_{m \times n}\), and define \(T \in L(\mathbb{R}^n, \mathbb{R}^m)\) by \(T(x) = Ax\). Then \(A\) is the matrix representing \(T\) relative to the standard bases.