By comparing the two sums above and using the commutativity of field multiplication, we see that for all i and j, $\gamma_{ij} = \delta_{ij}$. This computation proves the following result:

Theorem 12. With notations as above, $M_{\mathcal{D}}^{\mathcal{E}}(\varphi \circ \psi) = M_{\mathcal{B}}^{\mathcal{E}}(\varphi) M_{\mathcal{D}}^{\mathcal{B}}(\psi)$, i.e., with respect to a compatible choice of bases, the product of the matrices representing the linear transformations φ and ψ is the matrix representing the composite linear transformation $\varphi \circ \psi$.

Corollary 13. Matrix multiplication is associative and distributive (whenever the dimensions are such as to make products defined). An $n \times n$ matrix A is nonsingular if and only if it is invertible.

Proof: Let A, B and C be matrices such that the products $(AB)C$ and $A(BC)$ are defined, and let S, T and R denote the associated linear transformations. By Theorem 12, the linear transformation corresponding to AB is the composite $S \circ T$ so the linear transformation corresponding to $(AB)C$ is the composite $(S \circ T) \circ R$. Similarly, the linear transformation corresponding to $A(BC)$ is the composite $S \circ (T \circ R)$. Since function composition is associative, these two linear transformations are the same, and so $(AB)C = A(BC)$ by Theorem 10. The distributivity is proved similarly. Note also that it is possible to prove these results by straightforward (albeit tedious) calculations with matrices.

If A is invertible, then $Ax = 0$ implies $x = A^{-1}Ax = A^{-1}0 = 0$, so A is nonsingular. Conversely, if A is nonsingular, fix bases $\mathcal{B}$, $\mathcal{E}$ for V and let φ be the linear transformation of V to itself represented by A with respect to these bases. By Corollary 9, φ is an isomorphism of V to itself, hence has an inverse, φ^{-1}. Let B be the matrix representing φ^{-1} with respect to the bases $\mathcal{E}$, $\mathcal{B}$ (note the order). Then $AB = M_{\mathcal{B}}^{\mathcal{E}}(\varphi) M_{\mathcal{E}}^{\mathcal{B}}(\varphi^{-1}) = M_{\mathcal{E}}^{\mathcal{E}}(\varphi \circ \varphi^{-1}) = M_{\mathcal{E}}^{\mathcal{E}}(1) = I$. Similarly, $BA = I$ so B is the inverse of A.

Corollary 14.

(1) If $\mathcal{B}$ is a basis of the n-dimensional space V, the map $\varphi \mapsto M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ is a ring and a vector space isomorphism of $\text{Hom}_F(V, V)$ onto the space $M_n(F)$ of $n \times n$ matrices with coefficients in F.

(2) $GL(V) \cong GL_n(F)$ where $\dim V = n$. In particular, if F is a finite field the order of the finite group $GL_n(F)$ (which equals $|GL(V)|$) is given by the formula at the end of Section 1.

Proof: (1) We have already seen in Theorem 10 that this map is an isomorphism of vector spaces over F. Corollary 13 shows that $M_n(F)$ is a ring under matrix multiplication, and then Theorem 12 shows that multiplication is preserved under this map, hence it is also a ring isomorphism.

(2) This is immediate from (1) since a ring isomorphism sends units to units.

Definition. If A is any $m \times n$ matrix with entries from F, the *row rank* (respectively, *column rank*) of A is the maximal number of linearly independent rows (respectively, columns) of A (where the rows or columns of A are considered as vectors in affine n-space, m-space, respectively).

The relation between the rank of a matrix and the rank of the associated linear transformation is the following: the rank of φ as a linear transformation equals the column rank of the matrix $M_{\mathcal{B}}^{\mathcal{E}}(\varphi)$ (cf. the exercises). We shall also see that the row rank and the column rank of any matrix are the same.

We now consider the relation of two matrices associated to the same linear transformation of a vector space to itself but with respect to two different choices of bases (cf. the exercises for the general statement regarding a linear transformation from a vector space V to another vector space W).

Definition. Two $n \times n$ matrices A and B are said to be *similar* if there is an invertible (i.e., nonsingular) $n \times n$ matrix P such that $P^{-1}AP = B$. Two linear transformations φ and ψ from a vector space V to itself are said to be *similar* if there is a nonsingular linear transformation ξ from V to V such that $\xi^{-1}\varphi\xi = \psi$.

Suppose $\mathcal{B}$ and $\mathcal{E}$ are two bases of the same vector space V and let $\varphi \in \text{Hom}_F(V, V)$. Let I be the identity map from V to V and let $P = M_{\mathcal{E}}^{\mathcal{B}}(I)$ be its associated matrix (in other words, write the elements of the basis $\mathcal{E}$ in terms of the basis $\mathcal{B}$ — note the order — and use the resulting coordinates for the columns of the matrix P). Note that if $\mathcal{B} \neq \mathcal{E}$ then P is *not* the identity matrix. Then $P^{-1}M_{\mathcal{B}}^{\mathcal{B}}(\varphi)P = M_{\mathcal{E}}^{\mathcal{E}}(\varphi)$. If $[v]_{\mathcal{B}}$ is the $n \times 1$ matrix of coordinates for $v \in V$ with respect to the basis $\mathcal{B}$, and similarly $[v]_{\mathcal{E}}$ is the $n \times 1$ matrix of coordinates for $v \in V$ with respect to the basis $\mathcal{E}$, then $[v]_{\mathcal{B}} = P[v]_{\mathcal{E}}$. The matrix P is called the *transition* or *change of basis* matrix from $\mathcal{B}$ to $\mathcal{E}$ and this similarity action on $M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ is called a *change of basis*. This shows that the matrices associated to the same linear transformation with respect to two different bases are similar.

Conversely, suppose A and B are $n \times n$ matrices similar by a nonsingular matrix P. Let $\mathcal{B}$ be a basis for the n-dimensional vector space V. Define the linear transformation φ of V (with basis $\mathcal{B}$) to V (again with basis $\mathcal{B}$) by equation (3) using the given matrix A, i.e.,

$$\varphi(v_j) = \sum_{i=1}^{n} \alpha_{ij} v_i.$$

Then $A = M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ by definition of φ. Define a new basis $\mathcal{E}$ of V by using the i^{th} column of P for the coordinates of w_i in terms of the basis $\mathcal{B}$ (so $P = M_{\mathcal{E}}^{\mathcal{B}}(I)$ by definition). Then $B = P^{-1}AP = P^{-1}M_{\mathcal{B}}^{\mathcal{B}}(\varphi)P = M_{\mathcal{E}}^{\mathcal{E}}(\varphi)$ is the matrix associated to φ with respect to the basis $\mathcal{E}$. This shows that any two similar $n \times n$ matrices arise in this fashion as the matrices representing the same linear transformation with respect to two different choices of bases.

Note that change of basis for a linear transformation from V to itself is the same as conjugation by some element of the group $GL(V)$ of nonsingular linear transformations of V to V. In particular, the relation "similarity" is an equivalence relation whose equivalence classes are the orbits of $GL(V)$ acting by conjugation on $\text{Hom}_F(V, V)$. If

$\varphi \in GL(V)$ (i.e., φ is an invertible linear transformation), then the similarity class of φ is none other than the conjugacy class of φ in the group $GL(V)$.

Example

Let $V = \mathbb{Q}^3$ and let φ be the linear transformation

$$\varphi(x, y, z) = (9x + 4y + 5z, -4x - 3z, -6x - 4y - 2z), \qquad x, y, z \in \mathbb{Q}$$

from V to itself we considered in an earlier example. With respect to the standard basis, $\mathcal{B}$, $b_1 = (1, 0, 0)$, $b_2 = (0, 1, 0)$, $b_3 = (0, 0, 1)$ we saw that the matrix A representing this linear transformation is

$$A = M_{\mathcal{B}}^{\mathcal{B}}(\varphi) = \begin{pmatrix} 9 & 4 & 5 \\ -4 & 0 & -3 \\ -6 & -4 & -2 \end{pmatrix}.$$

Take now the basis, $\mathcal{E}$, $e_1 = (2, -1, -2)$, $e_2 = (1, 0, -1)$, $e_3 = (3, -2, -2)$ for V (we shall see that this is in fact a basis momentarily). Since

$$\begin{aligned}
\varphi(e_1) &= \varphi(2, -1, -2) = (4, -2, -4) = 2 \cdot e_1 + 0 \cdot e_2 + 0 \cdot e_3 \\
\varphi(e_2) &= \varphi(1, 0, -1) = (4, -1, -4) = 1 \cdot e_1 + 2 \cdot e_2 + 0 \cdot e_3 \\
\varphi(e_3) &= \varphi(3, -2, -2) = (9, -6, -6) = 0 \cdot e_1 + 0 \cdot e_2 + 3 \cdot e_3,
\end{aligned}$$

the matrix representing φ with respect to this basis is the matrix

$$B = M_{\mathcal{E}}^{\mathcal{E}}(\varphi) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Writing the elements of the basis $\mathcal{E}$ in terms of the basis $\mathcal{B}$ we have

$$\begin{aligned}
e_1 &= 2b_1 - b_2 - 2b_3 \\
e_2 &= b_1 - b_3 \\
e_3 &= 3b_1 - 2b_2 - 2b_3
\end{aligned}$$

so the matrix $P = M_{\mathcal{E}}^{\mathcal{B}}(I) = \begin{pmatrix} 2 & 1 & 3 \\ -1 & 0 & -2 \\ -2 & -1 & -2 \end{pmatrix}$ with inverse $P^{-1} = \begin{pmatrix} -2 & -1 & -2 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ conjugates A into B, i.e., $P^{-1}AP = B$, as can easily be checked. (Note incidentally that since P is invertible this proves that $\mathcal{E}$ is indeed a basis for V.)

We observe in passing that the matrix B representing this linear transformation φ is much simpler than the matrix A representing φ. The study of the simplest possible matrix representing a given linear transformation (and which basis to choose to realize it) is the study of *canonical forms* considered in the next chapter.

Linear Transformations on Tensor Products of Vector Spaces

For convenience we reiterate Corollaries 18 and 19 of Section 10.4 for the special case of vector spaces.

Proposition 15. Let F be a subfield of the field K. If W is an m-dimensional vector space over F with basis $w_1, \dots, w_m$, then $K \otimes_F W$ is an m-dimensional vector space over K with basis $1 \otimes w_1, \dots, 1 \otimes w_m$.

Proposition 16. Let V and W be finite dimensional vector spaces over the field F with bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ respectively. Then $V \otimes_F W$ is a vector space over F of dimension nm with basis $v_i \otimes w_j$, $1 \le i \le n$ and $1 \le j \le m$.

Remark: If v and w are nonzero elements of V and W, respectively, then it follows from the proposition that $v \otimes w$ is a nonzero element of $V \otimes_F W$, because we may always build bases of V and W whose first basis vectors are v, w, respectively. In a tensor product $M \otimes_R N$ of two R-modules where R is not a field it is in general substantially more difficult to determine when the tensor product $m \otimes n$ of two nonzero elements is zero.

Now let V, W, X, Y be finite dimensional vector spaces over F and let

$$\varphi : V \to X \qquad \text{and} \qquad \psi : W \to Y$$

be linear transformations. We compute a matrix of the linear transformation

$$\varphi \otimes \psi : V \otimes W \to X \otimes Y.$$

Let $\mathcal{B}_1 = \{v_1, \ldots, v_n\}$ and $\mathcal{B}_2 = \{w_1, \ldots, w_m\}$ be (ordered) bases of V and W respectively, and let $\mathcal{E}_1 = \{x_1, \ldots, x_r\}$ and $\mathcal{E}_2 = \{y_1, \ldots, y_s\}$ be (ordered) bases of X and Y respectively. Let $\mathcal{B} = \{v_i \otimes w_j\}$ and $\mathcal{E} = \{x_i \otimes y_j\}$ be the bases of $V \otimes W$ and $X \otimes Y$ given by Proposition 16; we shall order these shortly. Suppose

$$\varphi(v_i) = \sum_{p=1}^{r} \alpha_{pi} x_p \qquad \text{and} \qquad \psi(w_j) = \sum_{q=1}^{s} \beta_{qj} y_q.$$

Then

$$\begin{aligned}(\varphi \otimes \psi)(v_i \otimes w_j) &= (\varphi(v_i)) \otimes (\psi(w_j)) \\ &= \Big(\sum_{p=1}^{r} \alpha_{pi} x_p\Big) \otimes \Big(\sum_{q=1}^{s} \beta_{qj} y_q\Big) \\ &= \sum_{p=1}^{r} \sum_{q=1}^{s} \alpha_{pi} \beta_{qj} (x_p \otimes y_q).\end{aligned} \tag{11.8}$$

In view of the order of summation in (11.8) we order the basis $\mathcal{E}$ into r ordered sets, with the p^{th} list being $x_p \otimes y_1, x_p \otimes y_2, \ldots, x_p \otimes y_s$, and similarly order the basis $\mathcal{B}$. Then equation (8) determines the column entries for the corresponding matrix of $\varphi \otimes \psi$. The resulting matrix $M_{\mathcal{B}}^{\mathcal{E}}(\varphi \otimes \psi)$ is an $r \times n$ block matrix whose p, q block is the $s \times m$ matrix $\alpha_{p,q} M_{\mathcal{B}_2}^{\mathcal{E}_2}(\psi)$. In other words, the matrix for $\varphi \otimes \psi$ is obtained by taking the matrix for φ and multiplying each entry by the matrix for ψ. Such matrices have a name:

Definition. Let $A = (\alpha_{ij})$ and B be $r \times n$ and $s \times m$ matrices, respectively, with coefficients from any commutative ring. The *Kronecker product* or *tensor product* of A and B, denoted by $A \otimes B$, is the $rs \times nm$ matrix consisting of an $r \times n$ block matrix whose i, j block is the $s \times m$ matrix $\alpha_{ij} B$.

With this terminology we have