# On active and passive transformations

Costas J. Papachristou

Department of Physical Sciences, Hellenic Naval Academy, Piraeus, Greece

papachristou@hna.gr

The concepts of active and passive transformations on a vector space are discussed. Orthogonal coordinate transformations and matrix representations of linear operators are considered in particular.

## 1. Introduction

A physical situation may *appear* changing for two reasons: the physical system itself may pass from one state to another, or, the *same* state of the system may be viewed from two different points of view (e.g., by two different observers, using different frames of reference). The former case corresponds to an *"active"* view of the situation, while the latter one to a *"passive"* view.

Given that many physical quantities are vectors, of particular interest in Physics are linear transformations on vector spaces. Starting with the prototype transformation of rotation on a plane, we study both the active and the passive view of these transformations. In the case of a Euclidean space with Cartesian coordinates, a passive transformation corresponding to a change of basis is an orthogonal transformation. On the other hand, an active transformation on a vector space is produced by a linear operator, which is represented by a matrix in a given basis. A change of basis, leading to a different representation, is a passive transformation on this space.

## 2. Active view of transformations

Consider the *xy*-plane with Cartesian coordinates $(x, y)$ and basis unit vectors $\{\hat{u}_x, \hat{u}_y\}$. We call $\mathbf{R}(\theta)$ the rotation operator on this plane, i.e., the operator which rotates any vector $\vec{A}$ on the plane by an angle $\theta$ (see Fig. 2.1; by convention, $\theta>0$ for counterclockwise rotation while $\theta<0$ for clockwise rotation). This operator is linear, given that adding two vectors and then rotating the sum is the same as first rotating the vectors and then adding them.

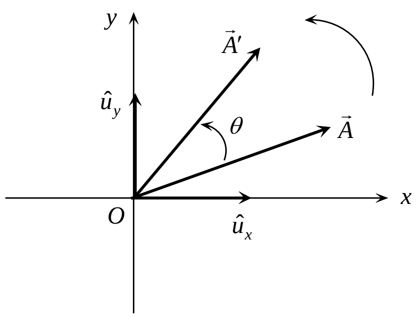


Figure 2.1

Imagine, in particular, that we rotate each vector in the basis $\{\hat{u}_x, \hat{u}_y\}$ by an angle $\theta$ to obtain a new set of vectors $\{\hat{u}_x{}', \hat{u}_y{}'\}$ (Fig. 2.2). The transformation equations describing these rotations are

$$\begin{aligned} \hat{u}_x{}' &= \mathbf{R}(\theta)\hat{u}_x = \cos\theta\,\hat{u}_x + \sin\theta\,\hat{u}_y \\ \hat{u}_y{}' &= \mathbf{R}(\theta)\hat{u}_y = -\sin\theta\,\hat{u}_x + \cos\theta\,\hat{u}_y \end{aligned} \tag{2.1}$$

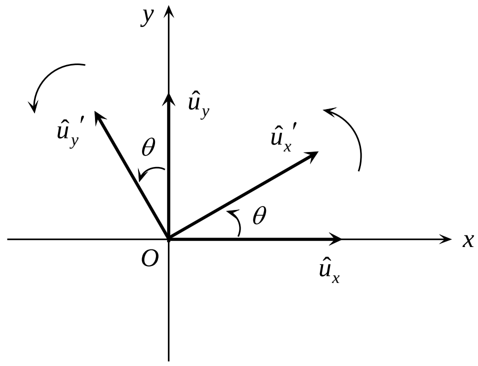


Figure 2.2

Now, let $\vec{A} = A_x \hat{u}_x + A_y \hat{u}_y$ be a vector on the *xy*-plane (see Fig. 2.1). The rotation operator $\mathbf{R}(\theta)$ will transform it into a new vector

$$\vec{A}' = \mathbf{R}(\theta)\vec{A} = A_x{}'\hat{u}_x + A_y{}'\hat{u}_y \tag{2.2}$$

We want to express the components $A_x{}'$ and $A_y{}'$ in terms of $A_x$, $A_y$ and $\theta$. By the linearity of $\mathbf{R}(\theta)$ and by using (2.1), we have:

$$\begin{aligned} \vec{A}' &= \mathbf{R}(\theta)\left(A_x\hat{u}_x + A_y\hat{u}_y\right) = A_x\mathbf{R}(\theta)\hat{u}_x + A_y\mathbf{R}(\theta)\hat{u}_y \\ &= \left(A_x\cos\theta - A_y\sin\theta\right)\hat{u}_x + \left(A_x\sin\theta + A_y\cos\theta\right)\hat{u}_y \end{aligned}$$

By comparing this with (2.2), we get:

$$\begin{aligned} A_x{}' &= A_x\cos\theta - A_y\sin\theta \\ A_y{}' &= A_x\sin\theta + A_y\cos\theta \end{aligned} \tag{2.3}$$

We define the matrix

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{2.4}$$

The systems (2.1) and (2.3) are then rewritten in the form of matrix equations as

$$\begin{bmatrix} \hat{u}_x{}' \\ \hat{u}_y{}' \end{bmatrix} = M^T \begin{bmatrix} \hat{u}_x \\ \hat{u}_y \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A_x{}' \\ A_y{}' \end{bmatrix} = M \begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{2.5}$$

respectively, where $M^T$ is the transpose of $M$.

We note that the vectors $\vec{A}$ and $\vec{A}' = \mathbf{R}(\theta)\vec{A}$ are *different* geometrical objects, the latter one being a transformation of the former. On the other hand, the components of these vectors, connected by (2.3), are referred to the *same* basis $\{\hat{u}_x, \hat{u}_y\}$. This is the general idea of the *active view* of a linear transformation.

In a more abstract sense, we consider an $n$-dimensional vector space $\Omega$ with basis vectors $\{\hat{e}_1, \hat{e}_2, ..., \hat{e}_n\} \equiv \{\hat{e}_k\}$, and we let $\mathbf{R}$ be a linear operator on $\Omega$. We assume that the basis vectors transform under $\mathbf{R}$ as follows:

$$\hat{e}_i{}' = \mathbf{R}\,\hat{e}_i = \hat{e}_j R^j{}_i \quad (\text{sum on } j) \tag{2.6}$$

where the familiar summation convention for repeated upper and lower indices has been used. Thus, for each value of $i$, the right-hand side of (2.6) is actually a sum over all values of $j$, i.e., from $j$=1 to $j$=$n$. Explicitly,

$$\begin{aligned} \hat{e}_1{}' &= \hat{e}_1 R^1{}_1 + \hat{e}_2 R^2{}_1 + \cdots + \hat{e}_n R^n{}_1 \\ \hat{e}_2{}' &= \hat{e}_1 R^1{}_2 + \hat{e}_2 R^2{}_2 + \cdots + \hat{e}_n R^n{}_2 \\ &\vdots \\ \hat{e}_n{}' &= \hat{e}_1 R^1{}_n + \hat{e}_2 R^2{}_n + \cdots + \hat{e}_n R^n{}_n \end{aligned} \tag{2.7}$$

Now, let

$$\vec{V} = V^1 \hat{e}_1 + V^2 \hat{e}_2 + \cdots + V^n \hat{e}_n \equiv V^i \hat{e}_i \tag{2.8}$$

be a vector in $\Omega$, and let $\vec{V}' = \mathbf{R}\vec{V}$. We have:

$$\vec{V}' = \mathbf{R}(V^j \hat{e}_j) = V^j \mathbf{R}\,\hat{e}_j = V^j \hat{e}_i R^i{}_j \equiv V^{i\prime} \hat{e}_i \ .$$

Therefore the components of the original and the transformed vector are related by

$$V^{i\prime} = R^i{}_j V^j \tag{2.9}$$

or, explicitly,

$$
\begin{aligned}
V^{1\prime} &= R^1{}_1 V^1 + R^1{}_2 V^2 + \cdots + R^1{}_n V^n \\
V^{2\prime} &= R^2{}_1 V^1 + R^2{}_2 V^2 + \cdots + R^2{}_n V^n \\
&\vdots \\
V^{n\prime} &= R^n{}_1 V^1 + R^n{}_2 V^2 + \cdots + R^n{}_n V^n
\end{aligned}
\tag{2.10}
$$

Define the $n \times n$ matrix

$$
M = \left[ R^i{}_j \right] \quad \text{with} \quad M_{ij} = R^i{}_j \tag{2.11}
$$

The basis transformations (2.6) are then written as

$$
\begin{bmatrix} \hat{e}_1' \\ \vdots \\ \hat{e}_n' \end{bmatrix} = M^T \begin{bmatrix} \hat{e}_1 \\ \vdots \\ \hat{e}_n \end{bmatrix} \tag{2.12}
$$

while the component transformations (2.9) become

$$
\begin{bmatrix} V^{1\prime} \\ \vdots \\ V^{n\prime} \end{bmatrix} = M \begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} \tag{2.13}
$$

## 3. Passive view of transformations

Imagine that our previous $x$-$y$ system of axes on the plane, with basis unit vectors $\{\hat{u}_x, \hat{u}_y\}$, is rotated counterclockwise by an angle $\theta$ to obtain a new system of axes $x'$ and $y'$ with corresponding basis $\{\hat{u}_x', \hat{u}_y'\}$ (Fig. 3.1). As before, the two bases are related by the system of equations

$$
\begin{aligned}
\hat{u}_x' &= \cos\theta\, \hat{u}_x + \sin\theta\, \hat{u}_y \\
\hat{u}_y' &= -\sin\theta\, \hat{u}_x + \cos\theta\, \hat{u}_y
\end{aligned}
\tag{3.1}
$$

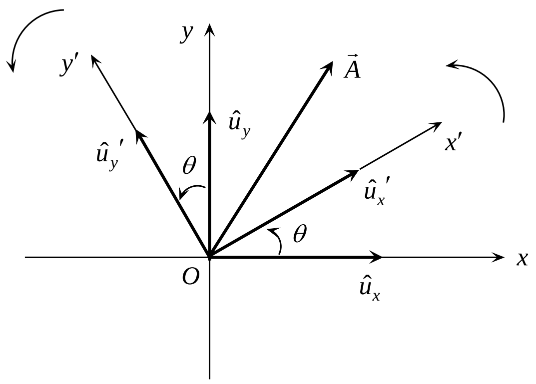


Figure 3.1

A vector $\vec{A}$ on the plane can be expressed in both these bases, as follows:

$$\vec{A} = A_x \hat{u}_x + A_y \hat{u}_y = A_x' \hat{u}_x' + A_y' \hat{u}_y' \tag{3.2}$$

Substituting the basis transformations (3.1) into the right-hand side of (3.2), and equating coefficients of similar unprimed basis vectors, we find:

$$\begin{aligned} A_x &= A_x' \cos\theta - A_y' \sin\theta \\ A_y &= A_x' \sin\theta + A_y' \cos\theta \end{aligned} \tag{3.3}$$

Solving this for the primed components, we get:

$$\begin{aligned} A_x' &= A_x \cos\theta + A_y \sin\theta \\ A_y' &= -A_x \sin\theta + A_y \cos\theta \end{aligned} \tag{3.4}$$

Notice that, in contrast to what we did in the previous section, here we keep the geometrical object $\vec{A}$ *fixed* and simply expand it in two *different* bases. This is the adopted practice in the *passive view* of a transformation.

Introducing the matrix

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

we rewrite our previous equations in the matrix forms

$$\begin{bmatrix} \hat{u}_x' \\ \hat{u}_y' \end{bmatrix} = M^T \begin{bmatrix} \hat{u}_x \\ \hat{u}_y \end{bmatrix} \tag{3.5}$$

and

$$\begin{bmatrix} A_x \\ A_y \end{bmatrix} = M \begin{bmatrix} A_x' \\ A_y' \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} A_x' \\ A_y' \end{bmatrix} = M^{-1} \begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{3.6}$$

where

$$M^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = M^T \tag{3.7}$$

Notice that the transformation matrix $M$ is *orthogonal*. As will be shown below, this is related to the fact that the transformation (rotation of axes) relates two Cartesian bases in a Euclidean space.

By comparing (2.3) and (3.4) it follows that the transformation equations of the passive view reduce to those of the active view upon replacing $\theta$ with $-\theta$. Physically this means that a passive transformation in which the vector $\vec{A}$ is fixed and the basis of our space is rotated *counterclockwise* is equivalent to an active transformation in which the basis is fixed and the vector $\vec{A}$ is rotated *clockwise*.

Let us generalize to the case of an $n$-dimensional vector space $\Omega$ with basis $\{\hat{e}_1, \hat{e}_2, ..., \hat{e}_n\} \equiv \{\hat{e}_k\}$. Let $\{\hat{e}_k{}'\}$ be another basis related to the former one by

$$\hat{e}_i{}' = \hat{e}_j \Lambda^j{}_{i'} \tag{3.8}$$

(note sum on $j$ ). A vector $\vec{V}$ in $\Omega$ may be expressed in both these bases, as follows:

$$\vec{V} = V^i \hat{e}_i = V^{j\prime} \hat{e}_j{}' = V^{j\prime} \hat{e}_i \Lambda^i{}_{j'}$$

where use has been made of (3.8). This yields

$$V^i = \Lambda^i{}_{j'} V^{j\prime} \tag{3.9}$$

Introducing the $n\times n$ matrix

$$M = \left[\Lambda^i{}_{j'}\right] \quad \text{with} \quad M_{ij} = \Lambda^i{}_{j'} \tag{3.10}$$

we write

$$\begin{bmatrix} \hat{e}_1{}' \\ \vdots \\ \hat{e}_n{}' \end{bmatrix} = M^T \begin{bmatrix} \hat{e}_1 \\ \vdots \\ \hat{e}_n \end{bmatrix} \tag{3.11}$$

and

$$\begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} = M \begin{bmatrix} V^{1\prime} \\ \vdots \\ V^{n\prime} \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} V^{1\prime} \\ \vdots \\ V^{n\prime} \end{bmatrix} = M^{-1} \begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} \tag{3.12}$$

## 4. Orthogonal transformations in a Euclidean space

In this section the *passive* view of transformations will be adopted. Let $\Omega$ be an $n$-dimensional Euclidean space with Cartesian[1] coordinates $(x^1, x^2, ..., x^n) \equiv (x^k)$ and corresponding Cartesian basis $\{\hat{e}_k\}$. Let $(x^{k\prime})$ be another Cartesian coordinate system for

[1] Cartesian systems of coordinates exist only in Euclidean spaces. For example, you can define a system of Cartesian coordinates on a plane but you *cannot* define such coordinates on the surface of a sphere, which is a *non-Euclidean* space.

$\Omega$, with corresponding basis $\{\hat{e}_k{}'\}$. We assume that the two coordinate systems have a common origin $O\equiv(0,0,...,0)$. Both Cartesian bases are *orthonormal*, in the sense that

$$\hat{e}_i \cdot \hat{e}_j = \hat{e}_i{}' \cdot \hat{e}_j{}' = \delta_{ij} \tag{4.1}$$

Assuming that the *handedness* of the two coordinate systems is the same (e.g., for $n=3$, both coordinate systems are right-handed) it is apparent that a linear transformation from one basis to the other is a "rotation" in $\Omega$. Let us explore this in more detail.

*Definition:* A linear transformation from a Cartesian basis to another is said to be an *orthogonal transformation*.

*Proposition 4.1:* An orthogonal transformation is represented by an *orthogonal* matrix $M$:

$$M^{-1} = M^T \quad \Leftrightarrow \quad M^T M = M M^T = \mathbf{1} \tag{4.2}$$

*Proof:* Assume a linear basis transformation of the form (3.8): $\hat{e}_i{}' = \hat{e}_j \Lambda^j{}_{i'}$. Also, let $M$ be the transformation matrix defined in (3.10). We have:

$$\hat{e}_i{}' \cdot \hat{e}_j{}' = \left(\hat{e}_k \Lambda^k{}_{i'}\right) \cdot \left(\hat{e}_l \Lambda^l{}_{j'}\right) = \delta_{kl} \Lambda^k{}_{i'} \Lambda^l{}_{j'} = \sum_k \Lambda^k{}_{i'} \Lambda^k{}_{j'}$$
$$= \sum_k M_{ki} M_{kj} = \sum_k \left(M^T\right)_{ik} M_{kj} = \left(M^T M\right)_{ij}$$

where we have taken into account that the original (unprimed) basis is orthonormal. Given that the same is true for the transformed (primed) basis, we have:

$$\left(M^T M\right)_{ij} = \delta_{ij} \quad \Leftrightarrow \quad M^T M = \mathbf{1}.$$

The *magnitude* of a vector $\vec{V}$ is a non-negative quantity whose square is expressed in a Cartesian basis in terms of the scalar (dot) product, as follows:

$$\left|\vec{V}\right|^2 = \vec{V} \cdot \vec{V} = \left(V^i \hat{e}_i\right) \cdot \left(V^j \hat{e}_j\right) = V^i V^j \hat{e}_i \cdot \hat{e}_j = \delta_{ij} V^i V^j \tag{4.3}$$

[Obviously, the last term in (4.3) is the sum of the squares of the components of $\vec{V}$.]

*Proposition 4.2:* An orthogonal transformation preserves the Cartesian form (4.3) of the magnitude of a vector.

*Proof:* By using the transformation formula (3.9) for components of vectors, derived in the previous section, we have:

$$\delta_{ij}V^iV^j = \delta_{ij}\left(\Lambda^i{}_{k'}V^{k'}\right)\left(\Lambda^j{}_{l'}V^{l'}\right) = \left(\sum_i \Lambda^i{}_{k'}\Lambda^i{}_{l'}\right)V^{k'}V^{l'}$$
$$= \left(\sum_i M_{ik}M_{il}\right)V^{k'}V^{l'} = \left(\sum_i \left(M^T\right)_{ki}M_{il}\right)V^{k'}V^{l'}$$
$$= \left(M^TM\right)_{kl}V^{k'}V^{l'} = \delta_{kl}V^{k'}V^{l'}$$

For a more compact proof, define the matrices

$$\left[V^k\right] \equiv \begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} \quad \text{and} \quad \left[V^k\right]^T \equiv \begin{bmatrix} V^1 & \cdots & V^n \end{bmatrix}$$

and similarly for the corresponding primed quantities. Then, in the unprimed basis,

$$\left|\vec{V}\right|^2 = \left[V^k\right]^T\left[V^k\right].$$

Using the fact that, by (3.12), $\left[V^k\right] = M\left[V^{k'}\right]$, we have:

$$\left[V^k\right]^T\left[V^k\right] = \left(M\left[V^{k'}\right]\right)^T M\left[V^{k'}\right] = \left[V^{k'}\right]^T M^TM\left[V^{k'}\right]$$
$$= \left[V^{k'}\right]^T\left[V^{k'}\right]$$

*Comment:* The above proof suggests an alternate definition of an orthogonal transformation as a linear transformation in a Euclidean space that preserves the Cartesian form of the magnitude of vectors. In fact, this is the way orthogonal transformations are usually defined in textbooks.

Now, let $P$ be a point in $\Omega$, with Cartesian coordinates $(x^1, x^2, ..., x^n) \equiv (x^k)$. In this system of coordinates the position vector of $P$ can be written as $\vec{r} = x^i\hat{e}_i$. Since this vector is a geometrical object independent of the system of coordinates, we can write:

$$\vec{r} = x^i\hat{e}_i = x^{j\,\prime}\hat{e}_j{}' .$$

By using (3.8) we find, as in Sec. 3,

$$x^i = \Lambda^i{}_{j'}\,x^{j\,\prime} \tag{4.4}$$

which is the analog of (3.9). If $M$ is the matrix defined in (3.10), and if $[x^k]$ is the column vector of the $x^k$, then by the general matrix relation (3.12) we have:

$$\left[x^k\right] = M\left[x^{k'}\right] \quad \Rightarrow \quad \left[x^{k'}\right] = M^{-1}\left[x^k\right] = M^T\left[x^k\right] \tag{4.5}$$

where the orthogonality condition (4.2) has been used. Let us call

$$M^T \equiv L \quad \text{with} \quad L_{ij} = M_{ji} = \Lambda^j{}_{i'} \tag{4.6}$$

Then the matrix relation (4.5) can be written as a system of $n$ linear equations of the form

$$\begin{aligned} x^{1\prime} &= L_{11}\,x^1 + L_{12}\,x^2 + \cdots + L_{1n}\,x^n \\ x^{2\prime} &= L_{21}\,x^1 + L_{22}\,x^2 + \cdots + L_{2n}\,x^n \\ &\vdots \\ x^{n\prime} &= L_{n1}\,x^1 + L_{n2}\,x^2 + \cdots + L_{nn}\,x^n \end{aligned} \tag{4.7}$$

which equations represent an orthogonal coordinate transformation in $\Omega$.

As an example for $n$=2, let $\Omega$ be a plane with Cartesian coordinates $(x^1, x^2) \equiv (x, y)$. A position vector in $\Omega$ is written: $\vec{r} = x\hat{u}_x + y\hat{u}_y$. As seen in Sec. 3, the transformation matrix $M$ for a rotation of the basis vectors by an angle $\theta$ is

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \Rightarrow \quad L = M^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

The coordinate transformation equations (4.7) are written here as

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned}$$

*Exercise:* By using the relations $\vec{V} = V^j\hat{e}_j$ and $\hat{e}_j{}' = \hat{e}_l\Lambda^l{}_{j'}$, together with (3.10) and (4.1), show the following:

$$V^i = \hat{e}_i \cdot \vec{V}\,,$$

$$M_{ij} = \hat{e}_i \cdot \hat{e}_j{}'\,.$$

Under an orthogonal transformation from one Cartesian system of coordinates to another, the components $V^k$ of a vector transform like the coordinates $x^k$ themselves. That is,

$$V^{i\prime} = L_{ij}\,V^j\,.$$

From (4.7) we have that

$$L_{ij} = \frac{\partial x^{i'}}{\partial x^j} .$$

Therefore,

$$V^{i'} = \frac{\partial x^{i'}}{\partial x^j} V^j \quad \text{and, conversely,} \quad V^i = \frac{\partial x^i}{\partial x^{j'}} V^{j'} \tag{4.8}$$

## 5. Active and passive view combined

Let $\Omega$ be an $n$-dimensional vector space with basis $\{\hat{e}_k\}$ $(k = 1, 2, \ldots, n)$. Let **A** be a linear operator on $\Omega$. The action of **A** on the basis vectors is given by

$$\mathbf{A}\,\hat{e}_j = \sum_i \hat{e}_i A_{ij} \equiv \hat{e}_i A_{ij} \tag{5.1}$$

(Note a slight change in the summation convention; in this section subscripts only will be used.) The $n\times n$ matrix $A=[A_{ij}]$ is the *matrix representation of the operator* **A** in the basis $\{\hat{e}_k\}$.

A vector in $\Omega$ is written:

$$\vec{x} = \sum_i x_i\,\hat{e}_i \equiv x_i\,\hat{e}_i \tag{5.2}$$

Let $\vec{y} = \mathbf{A}\,\vec{x}$. If $\vec{y} = y_i\,\hat{e}_i$, then, by the linearity of **A** and by using (5.1) and (5.2) we find that

$$y_i = A_{ij}\,x_j \quad (\text{sum on } j) \tag{5.3}$$

which represents a system of $n$ linear equations for $i=1,...,n$. In matrix form,

$$[y_k] = A\,[x_k] \tag{5.4}$$

where $[x_k]$ and $[y_k]$ are column vectors.

Now, let **A** and **B** be linear operators on $\Omega$. We define their product **C**=**AB** by

$$\mathbf{C}\,\vec{x} = (\mathbf{AB})\,\vec{x} \equiv \mathbf{A}\,(\mathbf{B}\vec{x}) \ , \quad \forall\,\vec{x} \in \Omega \tag{5.5}$$

Then, in the basis $\{\hat{e}_k\}$,

$$\mathbf{C}\,\hat{e}_j = \mathbf{A}\,(\mathbf{B}\,\hat{e}_j) = \mathbf{A}\,(\hat{e}_l\,B_{lj}) = B_{lj}\,(\mathbf{A}\,\hat{e}_l) = A_{il}\,B_{lj}\,\hat{e}_i \equiv \hat{e}_i\,C_{ij}$$

where

$$C_{ij} = A_{il}B_{lj} \quad \text{or, in matrix form,} \quad C = AB \tag{5.6}$$

That is, in any basis of *Ω*,

> *the matrix of the product of two operators is the product of the matrices of these operators.*

Consider now a change of basis (passive transformation) with transformation matrix $T=[T_{ij}]$:

$$\hat{e}_j{}' = \hat{e}_i \, T_{ij} \tag{5.7}$$

The inverse transformation is

$$\hat{e}_j = \hat{e}_i' \left(T^{-1}\right)_{ij} \tag{5.8}$$

The same vector may be expressed in both these bases as $\vec{x} = x_i \, \hat{e}_i = x_j{}' \, \hat{e}_j{}'$, from which we get, by using (5.7) and (5.8),

$$x_i = T_{ij} \, x_j{}' \quad \text{and} \quad x_i' = \left(T^{-1}\right)_{ij} x_j \tag{5.9}$$

In matrix form,

$$[x_k] = T\,[x_k{}'] \quad \text{and} \quad [x_k{}'] = T^{-1}[x_k] \tag{5.10}$$

How do the matrix elements of a linear operator **A** transform under a change of basis of the form (5.7)? In other words, how does the matrix of an active transformation transform under a passive transformation? Let $\vec{y} = \mathbf{A}\,\vec{x}$. By combining (5.10) with (5.4), we have:

$$[y_k{}'] = T^{-1}[y_k] = T^{-1}A[x_k] = T^{-1}AT\,[x_k{}'] \equiv A'[x_k{}'] \;\Rightarrow$$

$$A' = T^{-1}A\,T \tag{5.11}$$

For an alternative proof, note that

$$\mathbf{A}\,\hat{e}_j{}' = \mathbf{A}(\hat{e}_i T_{ij}) = T_{ij}\,\mathbf{A}\,\hat{e}_i = T_{ij}\,\hat{e}_l\,A_{li} = A_{li}T_{ij}\,\hat{e}_k{}'\left(T^{-1}\right)_{kl}$$

$$= \left(T^{-1}AT\right)_{kj}\hat{e}_k{}' \equiv \hat{e}_k{}' A'_{kj} \;\Rightarrow\; A' = T^{-1}AT$$

as before. A transformation of the form (5.11) is called a *similarity transformation*.

By applying the properties of the trace and the determinant of a matrix to (5.11) it is not hard to show that, under basis transformations, *the trace and the determinant of the matrix representation of an operator remain unchanged:* tr$A$=tr$A'$, det$A$=det$A'$. This means that the trace and the determinant are basis-independent quantities that are properties of the operator itself, rather than properties of its representation.

*Definition:* A vector $\vec{x} \neq 0$ is said to be an *eigenvector* of the linear operator **A** if a constant $\lambda$ exists such that

$$\mathbf{A}\vec{x} = \lambda \vec{x} \tag{5.12}$$

The constant $\lambda$ is an *eigenvalue* of **A**, to which eigenvalue this eigenvector belongs. Note that, in general, more than one eigenvector may belong to the same eigenvalue.

In a given basis $\{\hat{e}_k\}$, the linear system (5.3) corresponding to the eigenvalue equation (5.12) takes on the form

$$A_{ij}\, x_j = \lambda\, x_i \quad \text{or} \quad (A_{ij} - \lambda\, \delta_{ij})\, x_j = 0 \tag{5.13}$$

where $[A_{ij}]=A$ is the matrix of the operator **A** in the given basis. This is a homogeneous linear system of equations, which has a nontrivial solution for the eigenvector components iff

$$\det[A_{ij} - \lambda\, \delta_{ij}] = 0 \quad \text{or} \quad \det(A - \lambda 1) = 0 \tag{5.14}$$

where 1 here is the $n$-dimensional unit matrix. This polynomial equation determines the eigenvalues $\lambda_i$ (not necessarily all different from each other) of the operator **A**.

Now, in general, for any value of the constant $\lambda$ the matrix $(A–\lambda 1)$ is the representation of the operator $(\mathbf{A}–\lambda\mathbf{1})$ in the considered basis $\{\hat{e}_k\}$. Under a basis transformation to $\{\hat{e}_k{}'\}$ this matrix transforms according to (5.11):

$$(A–\lambda 1)' = T^{-1}(A–\lambda 1)\, T = T^{-1} A\, T - \lambda 1 \equiv A' - \lambda 1 \; .$$

On the other hand, by the invariance of the determinant under this transformation,

$$\det(A' - \lambda 1) = \det(A - \lambda 1) \; .$$

In particular, if $\lambda$ is an eigenvalue of the operator **A**, the right-hand side of the above equation vanishes in view of (5.14) and, therefore, the same must be true for the left-hand side *for the same value of* $\lambda$. That is, the polynomial equation (5.14) determines the eigenvalues of **A** uniquely, regardless of the chosen representation. We conclude that

> *the eigenvalues of an operator are a property of the operator itself and do not depend on the choice of basis of the space Ω.*

If we can find $n$ linearly independent eigenvectors $\{\vec{x}_k\}$ of **A**, belonging to the corresponding eigenvalues $\lambda_k$ (not necessarily all different) we can use these vectors to define a basis of $\Omega$. The matrix representation of **A** in this basis is given by (5.1): $\mathbf{A}\vec{x}_j = \vec{x}_i A_{ij}$. On the other hand, if $\lambda_j \equiv \lambda'$, then $\mathbf{A}\vec{x}_j = \lambda'\vec{x}_j = \lambda'\delta_{ij}\vec{x}_i$. Therefore, since the $\vec{x}_k$ are linearly independent, we must have $A_{ij}=\lambda'\delta_{ij}$ . We conclude that, in the eigenvector basis the matrix representation of the operator **A** has the *diagonal* form

$$A = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \; .$$

Moreover, by the above formula and by the fact that the quantities tr$A$, det$A$ and $\lambda_k$ are basis-independent (i.e., invariant under basis transformations) it follows that, in *any* basis of $\Omega$,

$$\mathrm{tr}\,A = \lambda_1 + \lambda_2 + \ldots + \lambda_n \quad , \qquad \det A = \lambda_1\, \lambda_2 \ldots \lambda_n \tag{5.15}$$

*Proposition 5.1:* Let **A** and **B** be two linear operator on $\Omega$. We assume that **A** and **B** have a common set of $n$ linearly independent eigenvectors $\{\vec{x}_k\}$. Then the operators **A** and **B** *commute*:

$$\mathbf{AB} = \mathbf{BA} \quad \Leftrightarrow \quad [\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA} = 0$$

where $[\mathbf{A}, \mathbf{B}]$ denotes the *commutator* of **A** and **B**.

*Proof:* Since the $n$ vectors $\{\vec{x}_k\}$ are linearly independent, they define a basis of $\Omega$. By assumption, for each value of $k$ the vector $\vec{x}_k$ is an eigenvector of both **A** and **B**, with corresponding eigenvalues, say, $\alpha$ and $\beta$. Then,

$$(\mathbf{AB})\,\vec{x}_k \equiv \mathbf{A}\,(\mathbf{B}\,\vec{x}_k) = \mathbf{A}\,(\beta\vec{x}_k) = \beta\,(\mathbf{A}\,\vec{x}_k) = \beta\alpha\,\vec{x}_k$$

and similarly, $(\mathbf{BA})\,\vec{x}_k = \alpha\beta\vec{x}_k$. Thus,

$$(\mathbf{AB})\,\vec{x}_k = (\mathbf{BA})\,\vec{x}_k \quad \Leftrightarrow \quad [\mathbf{A}, \mathbf{B}]\,\vec{x}_k = 0\,,$$

for all $k$=1,...,$n$. Now, let $\vec{\Psi} = \xi_i\,\vec{x}_i$ be an arbitrary vector in $\Omega$. Then,

$$[\mathbf{A}, \mathbf{B}]\,\vec{\Psi} = [\mathbf{A}, \mathbf{B}](\xi_i\,\vec{x}_i) = \xi_i\,[\mathbf{A}, \mathbf{B}]\vec{x}_i = 0\,, \quad \forall\,\vec{\Psi} \in \Omega\,.$$

This means that $[\mathbf{A}, \mathbf{B}]$=0.

*Definition:* An operator **A** is said to be *nonsingular* if det$A$≠0 (note that this is a *basis-independent* property). A nonsingular operator is *invertible*, in the sense that an inverse linear operator $\mathbf{A}^{-1}$ on $\Omega$ exists such that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{1}_{\mathrm{op}}$, where $\mathbf{1}_{\mathrm{op}}$ is the unit operator. This allows us to write

$$\vec{y} = \mathbf{A}\,\vec{x} \quad \Leftrightarrow \quad \vec{x} = \mathbf{A}^{-1}\,\vec{y}\,.$$

By (5.4) it follows that, if $A$ is the matrix representation of the nonsingular operator **A** in some basis, then *the matrix of the inverse operator* $\mathbf{A}^{-1}$ *is the inverse* $A^{-1}$ *of* $A$. As is well known, the matrix $A$ may have an inverse iff det$A$≠0, whence the definition of a nonsingular operator. In view of the second relation in (5.15),

> *all eigenvalues of a nonsingular operator are nonzero.*

Indeed, if even one eigenvalue vanishes, then det$A$=0 in *any* representation.

## 6. Comments

Both the active and the passive view are of importance in Physics. Let us see some examples:

1. The *Galilean transformation* of Classical Mechanics and the *Lorentz transformation* of Relativity[2] are *passive* transformations connecting different inertial frames of reference. When expressed in terms of mathematical equations, all physical laws are required to be invariant in form upon passing from one inertial frame to another.

2. The operators of Quantum Mechanics[3] are *active* transformations from a quantum state to a new state. On the other hand, both states and operators may be represented by matrices in different bases, the transformation from one basis to another being a *passive* transformation. Typically, the basis vectors of the quantum-mechanical space are chosen to be eigenvectors of linear operators representing physical quantities such as energy, angular momentum, etc. In such a basis the related operator is represented by a *diagonal* matrix, the diagonal elements being the *eigenvalues* of the operator. Physically, these eigenvalues give the possible values that a measurement of the associated physical quantity may yield in an experiment.

[2] H. Goldstein, *Classical Mechanics*, 2nd Ed. (Addison-Wesley, 1980).
[3] E. Merzbacher, *Quantum Mechanics*, 3rd Ed. (Wiley, 1998).