Mathematics 331
Spring, 2015
Thomas Wieting

LINEAR SPACES

01 Linear Spaces

01° Let $\mathbf{F}$ be a field. For our purposes, we take $\mathbf{F}$ to be the real number field $\mathbf{R}$ or the complex number field $\mathbf{C}$. Of course, we employ familiar notation:

$$a + b, \quad a \cdot b = ab$$

where a and b are any elements of $\mathbf{F}$.

02° By a *linear space*, we mean a set $\mathbf{V}$ supplied with operations of addition and scalar multiplication:

$$X + Y, \quad c.Z = cZ$$

where X, Y, and Z are any members of $\mathbf{V}$ and where c is any member of $\mathbf{F}$. The operation of addition must satisfy the familiar conditions:

$$\begin{gathered} X + Y = Y + X \\ X + (Y + Z) = (X + Y) + Z \\ X + 0 = X \\ X + (-X) = 0 \end{gathered}$$

where X, Y, and Z are any members of $\mathbf{V}$. By the third condition, we mean to assert that there is particular member 0 of $\mathbf{V}$, necessarily unique, which serves as the *neutral* member for addition. By the fourth condition, we mean to assert that, for each member X of $\mathbf{V}$, there is a member Y of $\mathbf{V}$, in relation to X necessarily unique, which serves as the *additive inverse* of X. We denote Y by $-X$.

03° Moreover, the operations of addition and scalar multiplication must together satisfy the following conditions:

$$\begin{gathered} a.(X + Y) = a.X + a.Y \\ (a + b).X = a.X + b.X \\ (a \cdot b).X = a.(b.X) \\ 1.X = X \end{gathered}$$

where X and Y are any members of $\mathbf{V}$ and where a and b are any members of $\mathbf{F}$.

04° By a *linear subspace* of the linear space $\mathbf{V}$, we mean a nonempty subset $\mathbf{U}$ of $\mathbf{V}$ which is invariant under the operations of addition and scalar multiplication on $\mathbf{V}$. We mean to assert that, for any members X, Y, and Z of $\mathbf{U}$ and for any member c of $\mathbf{F}$:

$$X + Y \in \mathbf{U}, \quad c.Z \in \mathbf{U}$$

Obviously, under the restrictions to $\mathbf{U}$ of the operations on $\mathbf{V}$, $\mathbf{U}$ is itself a linear space.

05° The set $\mathbf{F}^3$ provides a serviceable example of a linear space. The members of $\mathbf{F}^3$ have the following form:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

The operations stand as follows:

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}, \quad c.\mathbf{z} = \begin{pmatrix} c\,z_1 \\ c\,z_2 \\ c\,z_3 \end{pmatrix}$$

while:

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad -\mathbf{x} = \begin{pmatrix} -x_1 \\ -x_2 \\ -x_3 \end{pmatrix}$$

06° Just as well, we may introduce the linear space $\mathbf{F}^n$, where n is any positive integer. We refer to $\mathbf{F}^n$ as a *cartesian* linear space.

02 Linear Mappings

01° By a *linear mapping*, we mean a mapping L for which the domain $\mathbf{V}'$ and the codomain $\mathbf{V}''$ are linear spaces:

$$L : \mathbf{V}' \longrightarrow \mathbf{V}''$$

and for which the following conditions hold:

$$\begin{gathered} L(X' + Y') = L(X') + L(Y') \\ L(c.Z') = c.L(Z') \end{gathered}$$

where X', Y', and Z' are any members of $\mathbf{V}'$ and where c is any member of $\mathbf{F}$.

02° For such a mapping L, we define the *kernel* and the *range*:

$$ker(L), \quad ran(L)$$

as follows. First, $ker(L)$ consists of all members X' of $\mathbf{V}'$ such that:

$$L(X') = 0''$$

Second, $ran(L)$ consists of all members X'' of $\mathbf{V}''$ for which there exists at least one member X' of $\mathbf{V}'$ such that:

$$L(X') = X''$$

Clearly, $ker(L)$ is a linear subspace of $\mathbf{V}'$ and $ran(L)$ is a linear subspace of $\mathbf{V}''$.

03$^\bullet$ Obviously, L is surjective iff $ran(L) = \mathbf{V}''$. Moreover, L is injective iff $ker(L) = \{0'\}$, though this fact requires a little thought.

04° Finally, by definition, L is bijective iff it is both injective and surjective. In such a case, we claim that L^{-1} is linear. Accordingly, we would refer to L as a *linear isomorphism*. Let us prove the claim. Let X'', Y'', and Z'' be any members of $\mathbf{V}''$ and let c be any member of $\mathbf{F}$. Let X', Y', and Z' be the (uniquely determined) members of $\mathbf{V}'$ for which $L(X') = X''$, $L(Y') = Y''$, and $L(Z') = Z''$. We find that:

$$\begin{aligned} L^{-1}(X'' + Y'') &= L^{-1}(L(X') + L(Y')) \\ &= L^{-1}(L(X' + Y')) \\ &= X' + Y' \\ &= L^{-1}(X'') + L^{-1}(Y'') \end{aligned}$$

and:

$$\begin{aligned} L^{-1}(cZ'') &= L^{-1}(cL(Z')) \\ &= L^{-1}(L(cZ')) \\ &= cZ' \\ &= cL^{-1}(Z'') \end{aligned}$$

The proof of the claim is complete.

05° Let L_1 and L_2 be mappings for which the codomain of L_1 and the domain of L_2 coincide:

$$L_1 : \mathbf{V}' \longrightarrow \mathbf{V}'', \quad L_2 : \mathbf{V}'' \longrightarrow \mathbf{V}'''$$

Of course, we may form the composition L of L_1 and L_2:

$$L : \mathbf{V}' \longrightarrow \mathbf{V}'''$$

for which the domain is $\mathbf{V}'$ and the codomain is $\mathbf{V}'''$. By definition:

$$L(X) = L_2(L_1(X))$$

where X is any member of $\mathbf{V}'$. One can easily show that if L_1 and L_2 are linear then L is also linear. The proof takes the following form:

$$\begin{aligned} L(X+Y) &= L_2(L_1(X+Y)) \\ &= L_2(L_1(X) + L_1(Y)) \\ &= L_2(L_1(X)) + L_2(L_1(Y)) \\ &= L(X) + L(Y) \end{aligned}$$

and:

$$\begin{aligned} L(cZ) &= L_2(L_1(cZ)) \\ &= L_2(cL_1(Z)) \\ &= cL_2(L_1(Z)) \\ &= cL(Z) \end{aligned}$$

where X, Y, and Z are any members of $\mathbf{V}'$ and where c is any member of $\mathbf{F}$.

03 Bases

01° Let $\mathbf{V}$ be a linear space and let n be a positive integer. Let $\mathcal{B}$ be a finite list of (nonzero) members of $\mathbf{V}$ having length n:

$$\mathcal{B}: \quad B_1, B_2, \dots, B_n$$

Let K be the mapping carrying $\mathbf{F}^n$ to $\mathbf{V}$, defined as follows:

$$K(\mathbf{x}) = \sum_{j=1}^{n} x_j B_j$$

where:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

is any member of $\mathbf{F}^n$. Obviously, K is linear. We refer to the members of $ran(K)$ as *combinations* of $\mathcal{B}$ in $\mathbf{V}$.

02° It may happen that K is surjective, which is to say that $ran(K) = \mathbf{V}$. In such a case, every member of $\mathbf{V}$ is a combination of $\mathcal{B}$. We say that $\mathcal{B}$ *generates* $\mathbf{V}$. It may happen that K is injective, which is equivalent to the condition that $ker(K) = \{0\}$. It is the same to say that, for each member $\mathbf{x}$ of $\mathbf{F}^n$:

$$\sum_{j=1}^{n} x_j B_j = 0 \iff x_1 = 0\,, x_2 = 0, \dots, x_n = 0$$

We say that $\mathcal{B}$ is *independent*.

03° It may happen that K is both injective and surjective, hence, bijective, so that it is a linear isomorphism. Now, for each member Z of $\mathbf{V}$, there is precisely one member $\mathbf{x}$ of $\mathbf{F}^n$ such that:

$$Z = \sum_{j=1}^{n} x_j B_j$$

We refer to the numbers:

$$x_1, x_2, \dots, x_n$$

as the *coordinates* of Z relative to $\mathcal{B}$. We refer to $\mathcal{B}$ itself as a *basis* for $\mathbf{V}$.

04° For the cartesian linear space $\mathbf{F}^3$, we introduce the list $\mathcal{E}$ in $\mathbf{F}^3$ having length 3:

$$\mathcal{E}: \quad E_1, E_2, E_3$$

where:

$$E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ E_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \ E_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Clearly, $\mathcal{E}$ is a basis for $\mathbf{F}^3$. It is the *standard* basis. In fact, for each member $\mathbf{z}$ of $\mathbf{F}^3$:

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

we have:

$$\mathbf{z} = z_1 E_1 + z_2 E_2 + z_3 E_3$$

Consequently, the coordinates of $\mathbf{z}$ relative to $\mathcal{E}$ are the entries in $\mathbf{z}$:

$$z_1, \ z_2, \ z_3$$

05° The same design applies to the cartesian linear space $\mathbf{F}^n$, where n is any positive integer.

06° For efficient development of the concept of basis, we require three operations: Reduction, Expansion, and Exchange. Before describing them, however, let us describe a convenient maneuver: Renumbering. Let $\mathcal{C}$ be a finite list of (nonzero) members of $\mathbf{V}$ having length r:

$$\mathcal{C} : \quad C_1, C_2, \ \ldots \ , C_r$$

Very often, we find it convenient to permute the members of the foregoing list, then to renumber them in natural order. For instance:

$$\begin{gathered} C_1, C_2, C_3, C_4, C_5, C_6 \\ C_4, C_5, C_1, C_3, C_6, C_2 \\ C_1', C_2', C_3', C_4', C_5', C_6' \end{gathered}$$

where:

$$\begin{gathered} C_1' = C_4 \\ C_2' = C_5 \\ C_3' = C_1 \\ C_4' = C_3 \\ C_5' = C_6 \\ C_6' = C_2 \end{gathered}$$

Then we drop the primes. We refer to this process as Renumbering.

07° Now let us assume that $\mathcal{C}$ generates $\mathbf{V}$. It may happen that there is a member of $\mathcal{C}$, let it be C_j, which is a combination of the other members of $\mathcal{C}$. By Renumbering, we may assume that C_j is in fact C_r. Excising C_r, we obtain the abbreviated list:

$$\mathcal{C}' : \ C_1, C_2, \ \ldots \ , C_{r-1}$$

By elementary argument, we find that $\mathcal{C}'$ generates $\mathbf{V}$. Again, it may happen that there is a member of $\mathcal{C}'$, let it be C_k, which is a combination of the other members of $\mathcal{C}'$. By Renumbering, we may assume that C_k is in fact C_{r-1}. Excising C_{r-1}, we obtain the abbreviated list:

$$\mathcal{C}'' : \ C_1, C_2, \ \ldots \ , C_{r-2}$$

Again, we find that $\mathcal{C}''$ generates $\mathbf{V}$. Continuing in this way, we are led to a "terminal" list:

$$\bar{\mathcal{C}} : \ C_1, C_2, \ \ldots \ , C_q$$

such that $\bar{\mathcal{C}}$ generates $\mathbf{V}$ but no member of $\bar{\mathcal{C}}$ is a combination of the other members of $\bar{\mathcal{C}}$. At this point, the operation of Reduction stops. Now $\bar{\mathcal{C}}$ generates $\mathbf{V}$ and it is also independent. So $\bar{\mathcal{C}}$ is a basis for $\mathbf{V}$.

08° Again let us assume that $\mathcal{C}$ generates $\mathbf{V}$. Let p be a positive integer for which $1 \leq p \leq r$. It may happen that, among the r members of $\mathcal{C}$, there are p members which are independent. For instance, p might be 1, in which case any member of $\mathcal{C}$ would serve our purpose. By Renumbering, we may assume that these members lie at the beginning of $\mathcal{C}$:

$$\mathcal{C} : \ \mathcal{C}_\circ, \mathcal{C}_\bullet$$

where $\mathcal{C}_\circ$ and $\mathcal{C}_\bullet$ are the lists:

$$\mathcal{C}_\circ : \ C_1, C_2, \ \ldots \ , C_p; \quad \mathcal{C}_\bullet : \ C_{p+1}, \ \ldots \ , C_r$$

It may happen that there is a member of $\mathcal{C}_\bullet$, let it be C_j, which is not a combination of $\mathcal{C}_\circ$. By Renumbering, we may assume that C_j is in fact C_{p+1}. Now we can reform the list $\mathcal{C}$ as follows:

$$\mathcal{C} : \ \mathcal{C}'_\circ, \mathcal{C}'_\bullet$$

where $\mathcal{C}'_\circ$ and $\mathcal{C}'_\bullet$ are the lists:

$$\mathcal{C}'_\circ : \ C_1, C_2, \ \ldots \ , C_p, C_{p+1}; \quad \mathcal{C}'_\bullet : \ C_{p+2}, \ \ldots \ , C_r$$

By design, $\mathcal{C}'_\circ$ is independent. Again, it may happen that there is a member of $\mathcal{C}'_\bullet$, let it be C_k, which is not a combination of $\mathcal{C}'_\circ$. By Renumbering, we may assume that C_k is in fact C_{p+2}. Again, we can reform the list $\mathcal{C}$ as follows:

$$\mathcal{C}: \quad \mathcal{C}''_\circ, \mathcal{C}''_\bullet$$

where $\mathcal{C}''_\circ$ and $\mathcal{C}''_\bullet$ are the lists:

$$\mathcal{C}''_\circ: \quad C_1, C_2, \ \dots\ , C_p, C_{p+1}, C_{p+2}; \qquad \mathcal{C}''_\bullet: \quad C_{p+3}, \ \dots\ , C_r$$

Again, $\mathcal{C}''_\circ$ is independent. Continuing in this way, we are led to a "terminal" reformation:

$$\mathcal{C}: \quad \bar{\mathcal{C}}_\circ, \bar{\mathcal{C}}_\bullet$$

where $\bar{\mathcal{C}}_\circ$ and $\bar{\mathcal{C}}_\bullet$ are the lists:

$$\bar{\mathcal{C}}_\circ: \quad C_1, C_2, \ \dots\ , C_p, \ \dots\ , C_q; \qquad \bar{\mathcal{C}}_\bullet: \quad C_{q+1}, \ \dots\ , C_r$$

where $\bar{\mathcal{C}}_\circ$ is independent, and where every member of $\bar{\mathcal{C}}_\bullet$ is a combination of $\bar{\mathcal{C}}_\circ$. At this point, the operation of Expansion stops. Now $\bar{\mathcal{C}}_\circ$ is independent and it generates $\mathbf{V}$. So $\bar{\mathcal{C}}_\circ$ is a basis for $\mathbf{V}$.

09° Now let us introduce a pair of lists of (nonzero) members of $\mathbf{V}$ having lengths p and r, respectively:

$$\begin{aligned} &\mathcal{B}: \ B_1, B_2, \ \dots\ , B_p \\ &\mathcal{C}: \ C_1, C_2, \ \dots\dots\ , C_r \end{aligned}$$

Let us assume that $\mathcal{B}$ is independent and that $\mathcal{C}$ generates $\mathbf{V}$. Very soon, we will find that $p \leq r$, which justifies the seemingly biased display of $\mathcal{B}$ and $\mathcal{C}$. Of course, B_1 must be a combination of $\mathcal{C}$:

$$B_1 = x_1C_1 + x_2C_2 + x_3C_3 + \ \cdots\ + x_rC_r$$

Moreover, $B_1 \neq 0$, so that at least one of the displayed coefficients:

$$x_1, x_2, x_3, \dots, x_r$$

must be nonzero. By Renumbering, we may assume that $x_1 \neq 0$, so that:

$$x_1C_1 = B_1 - x_2C_2 - x_3C_3 - \ \cdots\ - x_rC_r \tag{1}$$

Now we exchange B_1 and C_1 to form the list:

$$\mathcal{C}': \quad B_1, C_2, C_3 \ \dots\ , C_r$$

The displayed relation (1) shows that $\mathcal{C}'$ generates $\mathbf{V}$. In turn, B_2 must be a combination of $\mathcal{C}'$:

$$B_2 = y_1 B_1 + y_2 C_2 + y_3 C_3 + \ \cdots \ + y_r C_r$$

Moreover, $\mathcal{B}$ is independent, so that at least one of the displayed coefficients:

$$y_2, y_3, \ldots, y_r$$

must be nonzero. By Renumbering, we may assume that $y_2 \neq 0$, so that:

(2) $$y_2 C_2 = -y_1 B_1 + B_2 - y_3 C_3 - \ \cdots \ - y_r C_r$$

Now we exchange B_2 and C_2 to form the list:

$$\mathcal{C}'' : \ \ B_1, B_2, C_3, \ \ldots \ , C_r$$

The displayed relation (2) shows that $\mathcal{C}''$ generates $\mathbf{V}$.

10° Continuing the foregoing Exchanges, we must eventually exhaust the list $\mathcal{B}$, to obtain the "terminal" list:

$$\bar{\mathcal{C}} : \ \ B_1, B_2, \ldots, B_p, C_{p+1}, \ldots, C_r$$

By design, $\bar{\mathcal{C}}$ generates $\mathbf{V}$. It follows that $p \leq r$.

11° Obviously, the contrary case, in which $r < p$, cannot occur.

12° The operations of Reduction and Expansion both yield bases for $\mathbf{V}$. In the first case, one can see that the list $\mathcal{C}$ plays a criticsl role. In the second case, however, one might be led to think that the list $\mathcal{C}$ is irrelevant. One might presume to start with a nonzero member B_1 of $\mathbf{V}$, then produce an ever longer list $\mathcal{B}_\circ$ by expansion:

$$\mathcal{B}_\circ : \ \ B_1, B_2, \ \ldots \ , B_j, \ \ldots$$

at each step adding a member B_j of $\mathbf{V}$ which is not a combination of the predecessors, until no such member exists. At the point of termination:

$$\mathcal{B} : \ \ B_1, B_2, \ \ldots \ , B_p$$

the list $\mathcal{B}$ would be a basis for $\mathbf{V}$. However, absent the context set by the list $\mathcal{C}$, one cannot be certain that the growing list $\mathcal{B}_\circ$ will terminate. By contrast, the lists produced by the operation of Expansion, as described in article 08°, at each step generate $\mathbf{V}$. Obviously, p cannot exceed r. Hence, the list $\mathcal{B}_\circ$ must terminate.

13° The following example illustrates the foregoing reservations. Let **P** be the linear space composed of the polynomial functions:

$$P(x) = \sum_{j=0}^{n} c_j x^j$$

where n is any nonnegative integer, where:

$$c_0, c_1, c_2, \ \ldots \ , c_n$$

are any numbers in **F**, and where x is a real variable. The operations of addition and scaler multiplication are defined as usual. Now one can easily see that the polynomials:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_2(x) &= x^2 \\ P_3(x) &= x^3 \\ &\vdots \end{aligned}$$

form an ever expanding chain of independent lists. Obviously, the naive process of expansion, just described, would fail to terminate.

14° Now let us apply the operation of Exchange to prove, in a few swift steps, several fundamental properties of bases. Let **V** be a linear space. Let us assume that there is a basis $\mathcal{B}$ for **V**:

$$\mathcal{B}: \ \ B_1, B_2, \ \ldots \ , B_n$$

as described in article 03°. In such a case, we say that **V** is *finite dimensional*. In turn, let $\mathcal{B}'$ and $\mathcal{B}''$ be bases for **V** having p members and q members, respectively: By the operation of Exchange, applied to $\mathcal{B}'$ and $\mathcal{B}''$ and to $\mathcal{B}''$ and $\mathcal{B}'$, respectively, we find that $p \leq q$ and $q \leq p$. Consequently, $p = q$. We conclude that, for a finite dimensional linear space, any two bases have the same number of members, let it be n. We refer to n as the *dimension* of **V**:

$$dim(\mathbf{V}) = n$$

15° Now let **V** be a linear space and let **U** be a linear subspace of **V**. We claim that if **V** is finite dimensional then **U** is finite dimensional and:

$$m = dim(\mathbf{U}) \leq dim(\mathbf{V}) = n$$

Moreover, $m = n$ iff $\mathbf{U} = \mathbf{V}$. To prove the claim, we introduce a basis $\mathcal{C}$ for **V**, having n members, and we consider independent lists $\mathcal{B}$ in **U**. From the operation of Exchange, we infer that the lengths of such lists in $\mathbf{U}$ cannot exceed n. Consequently, there must be independent lists $\mathcal{B}$ in $\mathbf{U}$ of maximum length, let it be m. Obviously, such a list would be a basis for $\mathbf{U}$. We declare the proof to be complete.

16° Finally, let us present the first of our fundamental theorems. Let $\mathbf{V}'$ and $\mathbf{V}''$ be finite dimensional linear spaces and let L be a linear mapping with domain $\mathbf{V}'$ and codomain $\mathbf{V}''$:

$$L : \mathbf{V}' \longrightarrow \mathbf{V}''$$

We contend that:

$$\text{(RT)} \qquad dim(\mathbf{U}) + dim(\mathbf{W}) = dim(\mathbf{V}')$$

where $\mathbf{U} = ker(L)$ and $\mathbf{W} = ran(L)$. One refers to this relation as:

THE RANK THEOREM

17° To prove the contention, we introduce a basis $\mathcal{B}$ for $\mathbf{U}$ and a basis $\mathcal{C}$ for $\mathbf{V}'$:

$$\mathcal{B} : \; B_1, B_2, \; \dots \; , B_p$$

$$\mathcal{C} : \; C_1, C_2, \; \dots \; , C_r$$

By articles 09° and 10°, $p \leq r$. By the operation of Exchange, we may reform $\mathcal{C}$ as follows:

$$\bar{\mathcal{C}} : \;\; B_1, B_2, \; \dots \; , B_p, \; C_{p+1}, C_{p+2}, \; \dots \; , C_r$$

We claim that the list:

$$\mathcal{D} : \;\; L(C_{p+1}), L(C_{p+2}), \; \dots \; , L(C_r)$$

of members of $\mathbf{V}''$ is a basis for $\mathbf{W}$. Having proved the claim, we will have proved the contention. Let Z'' be a member of $\mathbf{W}$. Let Z' be a member of $\mathbf{V}'$ such that $L(Z') = Z''$. Of course, there must be numbers:

$$x_1, x_2, \; \dots \; , x_p, y_{p+1}, y_{p+2}, \; \dots \; , y_r$$

in $\mathbf{F}$ such that:

$$Z' = x_1 B_1 + x_2 B_2 + \; \dots \; + x_p B_p + y_{p+1} C_{p+1} + y_{p+2} C_{p+2} + \; \dots \; + y_r C_r$$

Hence:

$$\begin{aligned} Z'' &= L(Z') \\ &= y_{p+1} L(C_{p+1}) + y_{p+2} L(C_{p+2}) + \; \dots \; + y_r L(C_r) \end{aligned}$$

Consequently, $\mathcal{D}$ generates $\mathbf{W}$. In turn, let:

$$y_{p+1}, y_{p+2}, \ \dots \ , y_r$$

be numbers in $\mathbf{F}$ such that:

$$y_{p+1}L(C_{p+1}) + y_{p+2}L(C_{p+2}) + \ \dots \ + y_rL(C_r) = 0$$

Hence:

$$y_{p+1}C_{p+1} + y_{p+2}C_{p+2} + \ \dots \ + y_rC_r \in \mathbf{U}$$

Since $\bar{\mathcal{C}}$ is independent:

$$y_{p+1} = 0, \, y_{p+2} = 0, \ \dots \ , y_r = 0$$

Consequently, $\mathcal{D}$ is independent. It follows that $\mathcal{D}$ is a basis for $\mathbf{W}$. The proof is complete.

04 Matrices

01° By a *matrix*, we mean a linear mapping M for which the domain and the codomain are cartesian spaces:

$$M : \mathbf{F}^p \longrightarrow \mathbf{F}^q$$

One may set the positive integers p and q at will. Let us show that M determines a rectangular array $\bar{M}$ having q rows and p columns, the entries for which are numbers in $\mathbf{F}$.

02° For precise expression, let us take p to be 3 and q to be 5. Let $\mathcal{E}'$ and $\mathcal{E}''$ be the standard bases for $\mathbf{F}^3$ and $\mathbf{F}^5$, respectively:

$$\mathcal{E}' : \quad E_1' = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, E_2' = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, E_3' = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\mathcal{E}' : \quad E_1'' = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, E_2'' = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, E_3'' = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, E_4'' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, E_5'' = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Now we introduce the array $\bar{M}$ as follows:

$$\bar{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \\ m_{41} & m_{42} & m_{43} \\ m_{51} & m_{52} & m_{53} \end{pmatrix}$$

where:

$$M(E_1') = m_{11}E_1'' + m_{21}E_2'' + m_{31}E_3'' + m_{41}E_4'' + m_{51}E_5'' = \begin{pmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{41} \\ m_{51} \end{pmatrix}$$

$$M(E_2') = m_{12}E_1'' + m_{22}E_2'' + m_{32}E_3'' + m_{42}E_4'' + m_{52}E_5'' = \begin{pmatrix} m_{12} \\ m_{22} \\ m_{32} \\ m_{42} \\ m_{52} \end{pmatrix}$$

$$M(E_3') = m_{13}E_1'' + m_{23}E_2'' + m_{33}E_3'' + m_{43}E_4'' + m_{53}E_5'' = \begin{pmatrix} m_{13} \\ m_{23} \\ m_{33} \\ m_{43} \\ m_{53} \end{pmatrix}$$

In this way, M determines $\bar{M}$. Just as well, $\bar{M}$ determines M. In fact, for any $\mathbf{x}$ in $\mathbf{F}^3$ and for any $\mathbf{y}$ in $\mathbf{F}^5$:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix}$$

we find that:

$$\begin{aligned} M(\mathbf{x}) &= M(x_1 E'_1 + x_2 E'_2 + x_3 E'_3) \\ &= x_1 M(E'_1) + x_2 M(E'_2) + x_3 M(E'_3) \end{aligned}$$

so that:

$$\mathbf{y} = M(\mathbf{x}) \iff \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = x_1 \begin{pmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{41} \\ m_{51} \end{pmatrix} + x_2 \begin{pmatrix} m_{12} \\ m_{22} \\ m_{32} \\ m_{42} \\ m_{52} \end{pmatrix} + x_3 \begin{pmatrix} m_{13} \\ m_{23} \\ m_{33} \\ m_{43} \\ m_{53} \end{pmatrix}$$

Clearly, the correspondence between linear mappings M carrying $\mathbf{F}^3$ to $\mathbf{F}^5$ and rectangular arrays $\bar{M}$ having 5 rows and 3 columns is bijective.

03° Of course, one may replace the positive integers 3 and 5 by any positive integers p and q.

04° Now let us lift the foregoing discussion to its proper level of generality. Let $\mathbf{V}'$ and $\mathbf{V}''$ be finite dimensional linear spaces. Let L be a linear mapping carrying $\mathbf{V}'$ to $\mathbf{V}''$. Let $\mathcal{B}'$ and $\mathcal{B}''$ be bases for $\mathbf{V}'$ and $\mathbf{V}''$, respectively. In this context, we will describe a matrix M for L. The corresponding rectangular array $\bar{M}$ of numbers in $\mathbf{F}$ will serve to define the *coordinates* of L relative to the bases $\mathcal{B}'$ and $\mathcal{B}''$.

05° We draw this connection between linear mappings in general and matrices in particular because, as a rule, theoretical developments proceed most smoothly in the general context while computational developments proceed most smoothly in the particular.

06° For explicit expression, let us set the dimensions of $\mathbf{V}'$ and $\mathbf{V}''$ to be 3 and 5, respectively. We display the bases $\mathcal{B}'$ and $\mathcal{B}''$ as follows:

$$\mathcal{B}' : \; B'_1, B'_2, B'_3; \qquad \mathcal{B}'' : \; B''_1, B''_2, B''_3, B''_4, B''_5$$

In turn, let us display the standard bases for $\mathbf{F}^3$ and $\mathbf{F}^5$:

$$\mathcal{E}' : \ E'_1, E'_2, E'_3; \qquad \mathcal{E}'' : \ E''_1, E''_2, E''_3, E''_4, E''_5$$

Of course, we obtain linear isomorphisms K' carrying $\mathbf{F}^p$ to $\mathbf{V}'$ and K'' carrying $\mathbf{F}^q$ to $\mathbf{V}''$, determined as follows:

$$\begin{array}{ll} & K''(E''_1) = B''_1 \\ K'(E'_1) = B'_1 & K''(E''_2) = B''_2 \\ K'(E'_2) = B'_2 & K''(E''_3) = B''_3 \\ K'(E'_3) = B'_3 & K''(E''_4) = B''_4 \\ & K''(E''_5) = B''_5 \end{array}$$

Now we declare that M shall be the linear mapping:

$$M = K''^{-1} \cdot L \cdot K'$$

carrying $\mathbf{F}^3$ to $\mathbf{F}^5$. It is the matrix for L relative to the bases $\mathcal{B}'$ and $\mathcal{B}''$. By a straightforward tour of the definitions, we find that:

$$\begin{aligned} L(B'_1) &= m_{11}B''_1 + m_{21}B''_2 + m_{31}B''_3 + m_{41}B''_4 + m_{51}B''_5 \\ L(B'_2) &= m_{12}B''_1 + m_{22}B''_2 + m_{32}B''_3 + m_{42}B''_4 + m_{52}B''_5 \\ L(B'_3) &= m_{13}B''_1 + m_{23}B''_2 + m_{33}B''_3 + m_{43}B''_4 + m_{53}B''_5 \end{aligned}$$

07° Obviously, L determines $\bar{M}$ while $\bar{M}$ determines L. The entries in $\bar{M}$ serve as coordinates for L relative to $\mathcal{B}'$ and $\mathcal{B}''$.

08° Now let $\mathbf{V}'$, $\mathbf{V}''$, and $\mathbf{V}'''$ be finite dimensional linear spaces. Let L° and $L^\bullet$ be linear mappings carrying $\mathbf{V}'$ to $\mathbf{V}''$ and $\mathbf{V}''$ to $\mathbf{V}'''$, respectively. Let $\mathcal{B}'$, $\mathcal{B}''$, and $\mathcal{B}'''$ be bases for $\mathbf{V}'$, $\mathbf{V}''$, and $\mathbf{V}'''$, respectively, and let K', K'', and K''' be the linear isomorphisms which they define. Let M° and $M^\bullet$ be the corresponding matrices:

$$M^\circ = K''^{-1} \cdot L^\circ \cdot K', \qquad M^\bullet = K'''^{-1} \cdot L^\bullet \cdot K''$$

In turn, let L be the linear mapping carrying $\mathbf{V}'$ to $\mathbf{V}'''$, defined by composition of L° and $L^\bullet$:

$$L = L^\bullet \cdot L^\circ$$

Let M be the corresponding matrix:

$$M = K'''^{-1} \cdot L \cdot K'$$

09° We can calculate M very easily:

$$\begin{aligned} M &= K'''^{-1} \cdot L \cdot K'^{-1} \\ &= K'''^{-1} \cdot L^{\bullet} \cdot L^{\circ} \cdot K'^{-1} \\ &= K'''^{-1} \cdot L^{\bullet} \cdot K'' \cdot K''^{-1} \cdot L^{\circ} \cdot K'^{-1} \\ &= M^{\bullet} \cdot M^{\circ} \end{aligned}$$

but we must work harder to calculate the rectangular array $\bar{M}$, corresponding to L, from the rectangular arrays $\bar{M}^{\circ}$ and $\bar{M}^{\bullet}$, corresponding to L° and $L^{\bullet}$, respectively.

10° Let us do so. Once again, for precise expression, we set the dimensions for $\mathbf{V}'$, $\mathbf{V}''$, and $\mathbf{V}'''$ at particular values, let them be 3, 4, and 2, respectively. We must find the means to calculate $\bar{M}$ from $\bar{M}^{\bullet}$ and $\bar{M}^{\circ}$:

$$\begin{pmatrix} m^{\bullet}_{11} & m^{\bullet}_{12} & m^{\bullet}_{13} & m^{\bullet}_{14} \\ m^{\bullet}_{21} & m^{\bullet}_{22} & m^{\bullet}_{23} & m^{\bullet}_{24} \end{pmatrix}, \begin{pmatrix} m^{\circ}_{11} & m^{\circ}_{12} & m^{\circ}_{13} \\ m^{\circ}_{21} & m^{\circ}_{22} & m^{\circ}_{23} \\ m^{\circ}_{31} & m^{\circ}_{32} & m^{\circ}_{33} \\ m^{\circ}_{41} & m^{\circ}_{42} & m^{\circ}_{43} \end{pmatrix} \Longrightarrow \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{pmatrix}$$

Of course, we require the standard bases $\mathcal{E}'$, $\mathcal{E}''$, and $\mathcal{E}'''$ for $\mathbf{F}^3$, $\mathbf{F}^4$, and $\mathbf{F}^2$, respectively. For any index ℓ $(1 \leq \ell \leq 3)$, we may adapt the computation in article 2° to show that:

$$\begin{aligned} \begin{pmatrix} m_{1\ell} \\ m_{2\ell} \end{pmatrix} &= M(E'_{\ell}) \\ &= (M^{\bullet} \cdot M^{\circ})(E'_{\ell}) \\ &= M^{\bullet}(\begin{pmatrix} m^{\circ}_{1\ell} \\ m^{\circ}_{2\ell} \\ m^{\circ}_{3\ell} \\ m^{\circ}_{4\ell} \end{pmatrix}) \\ &= m^{\circ}_{1\ell} \begin{pmatrix} m^{\bullet}_{11} \\ m^{\bullet}_{21} \end{pmatrix} + m^{\circ}_{2\ell} \begin{pmatrix} m^{\bullet}_{12} \\ m^{\bullet}_{22} \end{pmatrix} + m^{\circ}_{3\ell} \begin{pmatrix} m^{\bullet}_{13} \\ m^{\bullet}_{23} \end{pmatrix} + m^{\circ}_{4\ell} \begin{pmatrix} m^{\bullet}_{14} \\ m^{\bullet}_{24} \end{pmatrix} \end{aligned}$$

Hence:

$$\begin{aligned} m_{1\ell} &= m^{\bullet}_{11} m^{\circ}_{1\ell} + m^{\bullet}_{12} m^{\circ}_{2\ell} + m^{\bullet}_{13} m^{\circ}_{3\ell} + m^{\bullet}_{14} m^{\circ}_{4\ell} \\ m_{2\ell} &= m^{\bullet}_{21} m^{\circ}_{1\ell} + m^{\bullet}_{22} m^{\circ}_{2\ell} + m^{\bullet}_{23} m^{\circ}_{3\ell} + m^{\bullet}_{24} m^{\circ}_{4\ell} \end{aligned}$$

In the following more efficient notation:

$$m_{j\ell} = \sum_{k=1}^{4} m^{\bullet}_{jk} m^{\circ}_{k\ell} \qquad (1 \leq j \leq 2,\ 1 \leq \ell \leq 3)$$

we see how the computations would take form in general.

11° We summarize the foregoing computations by writing, very simply:

$$\bar{M} = \bar{M}^{\bullet}\bar{M}^{\circ}$$

In this way, we *multiply* rectangular arrays. Of course, we must insist that the number of columns of $\bar{M}^{\bullet}$ equals the number of rows of $\bar{M}^{\circ}$. From the properties of (linear) mappings, we infer that this operation is associative, though not in general commutative.

12° To be thorough, let us point to the operations of addition and scalar multiplication for rectangular arrays:

$$c\bar{M}, \quad \bar{M}^{\circ} + \bar{M}^{\bullet}$$

They would reflect the operations of addition and scalar multiplication for linear mappings:

$$cL, \quad L^{\circ} + L^{\bullet}$$

Of course, the domains and codomains for L° and $L^{\bullet}$ must coincide, while the numbers of rows for $\bar{M}^{\circ}$ and $\bar{M}^{\bullet}$ must be equal and the numbers of columns for $\bar{M}^{\circ}$ and $\bar{M}^{\bullet}$ must be equal as well.

13° Now we are led, one might say compelled, to introduce the linear spaces:

$$\mathbf{L}(\mathbf{V}', \mathbf{V}''), \quad \mathbf{M}(q, p)$$

where $\mathbf{V}'$ and $\mathbf{V}''$ are finite dimensional linear spaces and where p and q are positive integers. They are composed, in the first case, of all linear mappings L carrying $\mathbf{V}'$ to $\mathbf{V}''$, and, in the second case, of all rectangular arrays M having q rows and p columns.

14° By the foregoing discussion, it is plain that these new linear spaces are linearly isomorphic. In fact, the appropriate linear isomorphism $\mathcal{L}$ would be that which carries each member L of $\mathbf{L}(\mathbf{V}', \mathbf{V}'')$ to the corresponding rectangular array $\bar{M}$, as described in article 2°:

$$\mathcal{L}(L) = \bar{M}$$

Obviously:

$$dim(\mathbf{L}(\mathbf{V}', \mathbf{V}'')) = dim(\mathbf{M}(q, p)) = pq$$

15° It may happen that there is a finite dimensional linear space $\mathbf{V}$ such that $\mathbf{V}' = \mathbf{V} = \mathbf{V}''$. In such a case, we would write not $\mathbf{L}(\mathbf{V}', \mathbf{V}'')$ but $\mathbf{L}(\mathbf{V})$. Similarly, it may happen that there is a positive integer r such that $p = r = q$. We would write not $\mathbf{M}(q, p)$ but $\mathbf{M}(r)$. Now $\mathbf{L}(\mathbf{V})$ acquires the operation of composition and $\mathbf{M}(r)$ acquires the operation of multiplication. They are fully developed *algebras*.

05 Linear Functionals

01° Let $\mathbf{V}$ be a finite dimensional linear space. By the *dual space* for $\mathbf{V}$, we mean the linear space:

$$\mathbf{L}(\mathbf{V}, \mathbf{F})$$

consisting of all linear mappings for which the domain is $\mathbf{V}$ and the codomain is $\mathbf{F} = \mathbf{F}^1$. One refers to such mappings as *linear functionals*. Of course, the operations of addition and scalar multiple stand as follows:

$$\begin{aligned}(L' + L')(X) &= L'(X) + L''(X) \\ (c.L)(X) &= c.L(X)\end{aligned}$$

where L', L, and L'' are any members of $\mathbf{L}(\mathbf{V}, \mathbf{F})$, where c is any number in $\mathbf{F}$, and where X is any member of $\mathbf{V}$. In practice, we denote the dual space by the simpler symbol:

$$\mathbf{V}^*$$

02° Let us introduce a basis for $\mathbf{V}$:

$$\mathcal{B}: \quad B_1, B_2, \ \dots \ , B_n$$

Obviously, we intend that $dim(\mathbf{V}) = n$. In turn, let us design a basis for $\mathbf{V}^*$. For each index j $(1 \leq j \leq n)$, let Λ_j be the mapping carrying $\mathbf{V}$ to $\mathbf{F}$, defined as follows:

$$\Lambda_j(X) = x_j$$

where X is any member of $\mathbf{V}$:

$$X = x_1 B_1 + x_2 B_2 \ \cdots \ + x_n B_n$$

Clearly, Λ_j is a linear functional. We contend that:

$$\mathcal{L}: \quad \Lambda_1, \Lambda_2, \ \dots \ , \Lambda_n$$

is a basis for $\mathbf{V}^*$.

03° To prove the contention, we first display the following obvious but fundamental relations between $\mathcal{B}$ and $\mathcal{L}$:

$$\Lambda_j(B_k) = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}$$

where $1 \leq j \leq n$ and $1 \leq k \leq n$.

04° Let us consider an arbitrary combination $\bar{L}$ of $\mathcal{L}$:

$$\bar{L} = c_1\Lambda_1 + c_2\Lambda_2 + \cdots + c_n\Lambda_n$$

where c_1, c_2, ..., and c_n are any numbers in $\mathbf{F}$. If $\bar{L} = 0$ then:

$$0 = \bar{L}(B_j) = c_j$$

where j is any index $(1 \leq j \leq n)$. It follows that $\mathcal{L}$ is independent.

05° In turn, let L be any member of $\mathbf{V}^*$. Let us consider the combination $\bar{L}$ of $\mathcal{L}$:

$$\bar{L} = c_1\Lambda_1 + c_2\Lambda_2 + \cdots + c_n\Lambda_n$$

where c_1, c_2, ..., and c_n are the numbers in $\mathbf{F}$ defined as follows:

$$c_j = L(B_j)$$

Obviously:

$$\bar{L}(B_j) = c_j = L(B_j)$$

where j is any index $(1 \leq j \leq n)$. Hence, $L = \bar{L}$. It follows that $\mathcal{L}$ generates $\mathbf{V}^*$.

06° To be explicit, let us note that $\mathcal{B}$ and $\mathcal{L}$ have the same number of members, so that:

$$dim(\mathbf{V}^*) = dim(\mathbf{V})$$

07° Now let us describe a grand generalization of the design of $\mathbf{V}^*$, yielding a legion of new linear spaces:

$$\mathbf{L}^k(\mathbf{V})$$

where k is any positive integer. At this point, we do no more than define the spaces. In the following section, we will develop the means for analyzing their properties. We will focus attention upon certain linear subspaces:

$$\boldsymbol{\Sigma}^k(\mathbf{V}), \quad \boldsymbol{\Lambda}^k(\mathbf{V})$$

of $\mathbf{L}^k(\mathbf{V})$. They will play basic roles in our subsequent studies of determinants and of orthogonal and symplectic geometry.

08° We begin by introducing the product:

$$\mathbf{V}^k = \mathbf{V} \times \mathbf{V} \times \cdots \times \mathbf{V}$$

consisting of all ordered k-tuples:

$$(X_1, X_2, \ldots, X_k)$$

of members of $\mathbf{V}$. By a *k-linear functional*, we mean a mapping H for which the domain is $\mathbf{V}^k$, for which the codomain is $\mathbf{F}$:

$$H : \mathbf{V}^k \longrightarrow \mathbf{F}$$

and for which the following conditions hold:

$$\begin{aligned} &H(X_1, X_2, \ldots, X'_j + X''_j, \ldots, X_k) \\ &\qquad = H(X_1, X_2, \ldots, X'_j, \ldots, X_k) + H(X_1, X_2, \ldots, X''_j, \ldots, X_k) \\ &H(X_1, X_2, \ldots, cX_j, \ldots, X_k) \\ &\qquad = cH(X_1, X_2, \ldots, X_j, \ldots, X_k) \end{aligned}$$

where X_1, X_2, $\ldots$, X'_j, X_j, X''_j, $\ldots$, and X_k are any members of $\mathbf{V}$ and where c is any number in $\mathbf{F}$. Let:

$$\mathbf{L}^k(\mathbf{V})$$

stand for the set of all such functionals. Under the now familiar operations of addition and scalar multiplication, the set just described is a linear space.

09° Let us pause to describe a commonplace example of a k-linear functional. Let $\mathbf{V}$ be the linear space $\mathbf{F}^4$. Let:

$$\mathcal{E}: \quad E_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, E_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, E_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, E_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

be the standard basis for $\mathbf{F}^4$ and let:

$$\mathcal{F}: \quad \begin{aligned} F_1 &= (1 \quad 0 \quad 0 \quad 0) \\ F_2 &= (0 \quad 1 \quad 0 \quad 0) \\ F_3 &= (0 \quad 0 \quad 1 \quad 0) \\ F_4 &= (0 \quad 0 \quad 0 \quad 1) \end{aligned}$$

be the corresponding basis for $(\mathbf{F}^4)^*$. Of course, the members of $(\mathbf{F}^4)^*$ are matrices. We have represented the members of the basis $\mathcal{F}$ in terms of their corresponding rectangular arrays. Now, for illustration, we introduce the simple but useful 4-linear functional H in $\mathbf{L}^4(\mathbf{F}^4)$, as follows:

$$\begin{aligned} &H(X_1, X_2, X_3, X_4) \\ &\qquad = F_1(X_1)F_2(X_2)F_3(X_3)F_4(X_4) \end{aligned}$$

where X_1, X_2, X_3, and X_4 are any member of $\mathbf{F}^4$.

06 Determinants

01° Let k be a positive integer and let $\mathcal{K}$ be the set consisting of the first k positive integers:

$$\mathcal{K} = \{1, 2, 3, \ \dots\ , k\}$$

Let $\mathbf{S}_k$ be the set of all bijections carrying the set $\mathcal{K}$ to itself. We refer to the members of $\mathbf{S}_k$ as *permutations*. For any members σ and τ, the composition:

$$\tau \cdot \sigma$$

is itself a bijection carrying $\mathcal{K}$ to itself. Under this operation of composition, $\mathbf{S}_k$ is a *group*. The identity mapping ϵ carrying $\mathcal{K}$ to itself is the identity element for $\mathbf{S}_k$:

$$\epsilon \cdot \sigma = \sigma = \sigma \cdot \epsilon$$

Of course, the operation is associative:

$$\upsilon \cdot (\tau \cdot \sigma) = (\upsilon \cdot \tau) \cdot \sigma$$

It is not commutative. Moreover, for every member σ of $\mathbf{S}_k$, there is a member τ of $\mathbf{S}_k$ such that:

$$\sigma \cdot \tau = \epsilon = \tau \cdot \sigma$$

Of course, τ is the mapping inverse to σ: $\tau = \sigma^{-1}$. Now let p and q be (positive) integers in $\mathcal{K}$ for which $p < q$. Let π be the permutation in $\mathbf{S}_k$ defined as follows:

$$\pi(r) = \begin{cases} r & \text{if } r \neq p \text{ and } r \neq q \\ q & \text{if } r = p \\ p & \text{if } r = q \end{cases}$$

We refer to π as a *transposition* and we denote it by (pq). By a simple induction argument, one may prove that, for any σ in $\mathbf{S}_n$, there exist transpositions π_1, π_2, ... , and π_r such that:

$$\sigma = \pi_1 \cdot \pi_2 \cdot \ \cdots \ \cdot \pi_r$$

We claim that, for any two such presentations of σ:

$$\sigma = \pi'_1 \cdot \pi'_2 \cdot \ \cdots \ \cdot \pi'_{r'}, \qquad \sigma = \pi''_1 \cdot \pi''_2 \cdot \ \cdots \ \cdot \pi''_{r''}$$

the numbers r' and r'' must have the same parity, which is to say that both r' and r'' are even or both r' and r'' are odd. The proof of this claim lies just ahead.

02° In any case, we are justified, now, in referring to a permutation σ as *even* or *odd*, in accord with its presentation as a product of an even or an odd number of transpositions. Let us distinguish these cases by writing:

$$|\sigma| = -1 \text{ (odd)}, \qquad |\sigma| = +1 \text{ (even)}$$

03° Let k be any positive integer. Let Φ be any mapping carrying $\mathbf{V}^k$ to $\mathbf{F}$:

$$\Phi : \mathbf{V}^k \longrightarrow \mathbf{F}$$

Let σ be a permutation in $\mathbf{S}_k$. We define the *action* of σ on Φ as follows:

$$(\sigma \cdot \Phi)(X_1, X_2, \ \ldots \ , X_k) = \Phi(X_{\sigma(1)}, X_{\sigma(2)}, \ \ldots \ , X_{\sigma(k)})$$

where X_1, X_2, ... , and X_k are any members of $\mathbf{V}$.

04° Of course, $\sigma \cdot \Phi$ is a (new) mapping carrying $\mathbf{V}^k$ to $\mathbf{F}$.

05° Now let σ and τ be any permutations in $\mathbf{S}_k$. We contend that:

$$(\tau \cdot \sigma) \cdot \Phi = \tau \cdot (\sigma \cdot \Phi)$$

To prove the contention, we set k to be 6 and we interpret the foregoing definition in terms of the following notation:

$$\begin{aligned}
X_{\sigma(1)} &= \sigma^{-1}(X_1) \\
X_{\sigma(2)} &= \sigma^{-1}(X_2) \\
X_{\sigma(3)} &= \sigma^{-1}(X_3) \\
X_{\sigma(4)} &= \sigma^{-1}(X_4) \\
X_{\sigma(5)} &= \sigma^{-1}(X_5) \\
X_{\sigma(6)} &= \sigma^{-1}(X_6)
\end{aligned}$$

We note that σ does not change the given members of $\mathbf{V}$. It simply permutes them. In effect, the action of σ on $\mathcal{K} = \{1, 2, 3, 4, 5, 6\}$ has migrated to a corresponding action of σ^{-1} on the given members of $\mathbf{V}$. It is the same for τ and $\tau \cdot \sigma$. With this understanding, the proof of the contention is simple.

06° Now we consider the particular mapping Φ, defined as follows:

$$\Phi(X_1, X_2, \ \ldots \ , X_k) = \prod_{1 \leq p < q \leq n} (X_q - X_p)$$

where X_1, X_2, ... , and X_k are any members of $\mathbf{V}$. We contend that, for any transposition π in $\mathbf{S}_k$:

$$\pi \cdot \Phi = -\Phi \qquad (\pi = (rs))$$

Obviously, the relation just stated serves to prove the claim about parity at the end of article 01°. To prove the contention, we simply note that the transposition $\pi = (rs)$ will change the sign of Φ precisely $2b+1$ times, where b is the number of integers between r and s. Of course, b might be 0. In any case, $2b+1$ is odd.

07° Let us consider the interplay between $\mathbf{L}^k(\mathbf{V})$ and $\mathbf{S}_k$. The action of the group $\mathbf{S}_k$ on the linear space $\mathbf{L}^k(\mathbf{V})$ provides an elegant language with which to express the properties of symmetry and antisymmetry for k-linear functionals, yielding the basic linear subspaces:

$$\mathbf{\Sigma}^k(\mathbf{V}), \quad \mathbf{\Lambda}^k(\mathbf{V})$$

of $\mathbf{L}^k(\mathbf{V})$.

08° Let H be any member of $\mathbf{L}^k(\mathbf{V})$. We say that H is *symmetric* iff, for each σ in $\mathbf{S}_k$:

$$\sigma \cdot H = H$$

In turn, we say that H is *antisymmetric* iff, for each σ in $\mathbf{S}_k$:

$$\sigma \cdot H = |\sigma| H$$

These properties define the linear subspaces:

$$\mathbf{\Sigma}^k(\mathbf{V}), \quad \mathbf{\Lambda}^k(\mathbf{V})$$

of $\mathbf{L}^k(\mathbf{V})$.

09° For a first impression of these matters, we set k equal to 2. We introduce the linear mappings $\mathcal{S}$ and $\mathcal{A}$ carrying $\mathbf{L}^2(\mathbf{V})$ to itself, as follows:

$$\begin{aligned}\mathcal{S}(L)(X_1, X_2) &= \frac{1}{2}(L(X_1, X_2) + L(X_2, X_1)) \\ \mathcal{A}(L)(X_1, X_2) &= \frac{1}{2}(L(X_1, X_2) - L(X_2, X_1))\end{aligned}$$

where H is any member of $\mathbf{L}^2(\mathbf{V})$ and where X_1 and X_2 are any members of $\mathbf{V}$. Obviously, $\mathcal{S}(H)$ is symmetric and $\mathcal{A}(H)$ is antisymmetric. Moreover, if H is symmetric then $\mathcal{S}(H) = H$ while if H is antisymmetric then $\mathcal{A}(H) = H$. Finally,

$$H = \mathcal{S}(H) + \mathcal{A}(H)$$

10° The linear mappings $\mathcal{S}$ and $\mathcal{A}$ admit the following crisp expression:

$$\mathcal{S}(H) = \frac{1}{2}\sum_{\sigma} \sigma \cdot H, \quad \mathcal{A}(H) = \frac{1}{2}\sum_{\sigma} |\sigma|\, \sigma \cdot H$$

where H is any member of $\mathbf{L}^2(\mathbf{V})$. In fact:

$$\mathcal{S} = \frac{1}{2}\sum_{\sigma}\sigma, \quad \mathcal{A} = \frac{1}{2}\sum_{\sigma}|\sigma|\,\sigma$$

For the summations, we intend that σ run through the group $\mathbf{S}_2$. In this case, there are just two terms in each summation.

11° In due course, we will find that the symmetric and the antisymmetric 2-linear functionals are basic concepts underlying orthogonal and symplectic geometry.

12° Let us develop these matters in general. Let k be any positive integer. We introduce the linear mappings $\mathcal{S}$ and $\mathcal{A}$ carrying $\mathbf{L}^k(\mathbf{V})$ to itself:

$$\mathcal{S} = \frac{1}{k!}\sum_{\sigma}\sigma, \quad \mathcal{A} = \frac{1}{k!}\sum_{\sigma}|\sigma|\,\sigma$$

For the summations, we intend that σ run through the group $\mathbf{S}_k$. In this general setting, there are $k!$ terms in the two summations.

13° We contend that, for each H in $\mathbf{L}^k(\mathbf{V})$, $\mathcal{S}(H)$ is symmetric and $\mathcal{A}(H)$ is antisymmetric. Moreover, if H is symmetric then $\mathcal{S}(H) = H$ while if H is antisymmetric then $\mathcal{A}(H) = H$. Let us prove these contentions.

14° Let H be any member of $\mathbf{L}^k(\mathbf{V})$ and let $\bar{H} = \mathcal{S}(H)$. By definition:

$$\bar{H} = \frac{1}{k!}\sum_{\sigma}\sigma \cdot H$$

Let τ be any permutation in $\mathbf{S}_k$. We find that:

$$\tau \cdot \bar{H} = \frac{1}{k!}\sum_{\sigma}(\tau \cdot \sigma) \cdot H = \frac{1}{k!}\sum_{\tau\cdot\sigma}(\tau \cdot \sigma) \cdot H = \bar{H}$$

In the foregoing summations, $\tau \cdot \sigma$ runs through $\mathbf{S}_k$ as σ runs through $\mathbf{S}_k$. Hence, $\bar{H}$ is symmetric.

15° If at the outset H is symmetric then $\bar{H} = H$.

16° Again, let H be any member of $\mathbf{L}^k(\mathbf{V})$ but let $\bar{H} = \mathcal{A}(H)$. By definition:

$$\bar{H} = \frac{1}{k!}\sum_{\sigma}|\sigma|\,\sigma \cdot H$$

Let τ be any permutation in $\mathbf{S}_k$. We find that:

$$\tau \cdot \bar{H} = \frac{1}{k!} \sum_{\sigma} |\sigma| \, (\tau \cdot \sigma) \cdot H = |\tau| \, \frac{1}{k!} \sum_{\tau \cdot \sigma} |\tau \cdot \sigma| \, (\tau \cdot \sigma) \cdot H = |\tau| \, \bar{H}$$

In the foregoing summations, $\tau \cdot \sigma$ runs through $\mathbf{S}_k$ as σ runs through $\mathbf{S}_k$. Moreover:

$$|\tau \cdot \sigma| = |\tau||\sigma|$$

Hence, $\bar{H}$ is antisymmetric.

17° If at the outset H is antisymmetric then $\bar{H} = H$.

18° Now let n be any positive integer and let $\mathbf{V}$ be a linear space for which the dimension is n. Let us concentrate upon the linear space:

$$\mathbf{\Lambda}^n(\mathbf{V})$$

This space sets the context from which the theory of determinants springs. We contend that the dimension of $\mathbf{\Lambda}^n(\mathbf{V})$ is one:

$$dim(\mathbf{\Lambda}^n(\mathbf{V})) = 1$$

Let us prove this striking fact. We must produce a nonzero member Δ of $\mathbf{\Lambda}^n(\mathbf{V})$ and we must show that every member D of $\mathbf{\Lambda}^n(\mathbf{V})$ is a scalar multiple of Δ.

19° To that end, let us introduce bases:

$$\begin{aligned} \mathcal{B} &:= B_1, B_2, \ \ldots \ , B_n \\ \mathcal{L} &:= \Lambda_1, \Lambda_2, \ \ldots \ , \Lambda_n \end{aligned}$$

for $\mathbf{V}$ and $\mathbf{V}^*$ related in the following familiar manner:

$$\Lambda_j(B_k) = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}$$

Let H be the member of $\mathbf{L}^n(\mathbf{V})$ defined as follows:

$$H(X_1, X_2, \ \ldots \ , X_n) = n! \, \Lambda_1(X_1)\Lambda_2(X_2) \ \cdots \ \Lambda_n(X_n)$$

where X_1, X_2, ..., and X_n are any members of $\mathbf{V}$. Obviously, for any permutation σ in $\mathbf{S}_n$:

$$(\sigma \cdot H)(B_1, B_2, \ldots, B_n) = \begin{cases} 0 & \text{if } \sigma \neq \epsilon \\ n! & \text{if } \sigma = \epsilon \end{cases}$$

where ϵ is the identity permutation in $\mathbf{S}_n$. In turn, let Δ be the member of $\mathbf{\Lambda}^n(\mathbf{V})$ defined from H by anti-symmetrization:

$$\Delta = \mathcal{A}(H) = \frac{1}{n!}\sum_{\sigma} |\sigma|\, \sigma \cdot H$$

Obviously:

$$\Delta(B_1, B_2, \ldots, B_n) = 1$$

Hence, $\Delta \neq 0$.

20° Now let D be any member of $\mathbf{\Lambda}^n(\mathbf{V})$. We must show that D is a scalar multiple of Δ. Of course, if it were so:

$$D = d\Delta$$

then the number d would be, necessarily:

$$d = D(B_1, B_2, \ldots, B_n)$$

We are led to introduce the member E of $\mathbf{\Lambda}^n(\mathbf{V})$, as follows:

$$E = D - d\Delta$$

Obviously:

$$E(B_1, B_2, \ldots, B_n) = 0$$

On that ground alone, we claim that $E = 0$. Having proved the claim, we may declare the argument to be complete. The linear space $\mathbf{\Lambda}^n(\mathbf{V})$ is one dimensional and, in fact, the single member Δ is a basis for it.

21° Let us set aside the hubbub. Let us simply concentrate upon a member E of $\mathbf{\Lambda}^n(\mathbf{V})$ for which:

$$E(B_1, B_2, \ldots, B_n) = 0$$

We claim that $E = 0$. We will argue by induction. For the case in which $n = 1$, the claim holds true. Let m be any positive integer and let us assume that, for the case in which $n = m$, the claim holds true. We must show that the claim holds true for the case in which $n = m + 1$. Let Y be any member of $\mathbf{V}$:

$$Y = X + yB_{m+1}, \quad X = \sum_{j=1}^{m} x_j B_j$$

We find that:

$$\begin{aligned} &E(B_1, B_2, \ldots, B_m, Y) \\ &\qquad = E(B_1, B_2, \ldots, B_m, X) + yE(B_1, B_2, \ldots, B_m, B_{m+1}) \\ &\qquad = 0 \end{aligned}$$

By the induction hypothesis, the claim holds true for the case in question.

22° Now we stand at the end of the road. Let L be any linear mapping carrying $\mathbf{V}$ to itself. We contend that there is a precisely one number λ in $\mathbf{F}$ such that:

$$(\circ) \qquad D(L(X_1), L(X_2), \ldots, L(X_n)) = \lambda D(X_1, X_2, \ldots, X_n)$$

where D is any member of $\mathbf{\Lambda}^n(\mathbf{V})$ and where X_1, X_2, $\ldots$, and X_n are any members of $\mathbf{V}$. We refer to λ as the *determinant* of L:

$$(\star) \qquad \lambda = det(L)$$

23° Granted our elaborate preparation, the proof of our contention, fundamental and formidable, is very easy. For smooth expression, we represent the left hand side of relation $(\circ)$ in terms of the operation of *pullback*:

$$L^*(D)(X_1, X_2, \ldots, X_n) = D(L(X_1), L(X_2), \ldots, L(X_n))$$

In turn, we recover the basic member Δ for $\mathbf{\Lambda}^n(\mathbf{V})$, described in article 19°. Now we may introduce numbers d and λ in $\mathbf{F}$ such that:

$$D = d\Delta, \quad L^*(\Delta) = \lambda\Delta$$

We find that:

$$L^*(D) = L^*(d\Delta) = dL^*(\Delta) = d\lambda\Delta = \lambda D$$

The proof of our contention is complete.

24° Now let L_1 and L_2 be any linear mappings carrying $\mathbf{V}$ to itself. We claim that the following relation, fundamental and formidable, holds among the determinants of L_1, L_2, and $L_2 \cdot L_1$:

$$det(L_2 \cdot L_1) = det(L_2)\, det(L_1)$$

Again, the proof is simple. We recover the basic member Δ of $\mathbf{\Lambda}^n(\mathbf{V})$, we verify that:

$$(L_2 \cdot L_1)^*(\Delta) = L_1^*(L_2^*(\Delta))$$

and we compute:

$$\begin{aligned}
det(L_2 \cdot L_1)\Delta &= (L_2 \cdot L_1)^*(\Delta) \\
&= L_1^*(L_2^*(\Delta)) \\
&= L_1^*(det(L_2)\Delta) \\
&= det(L_2)L_1^*(\Delta) \\
&= det(L_2)\, det(L_1)\Delta
\end{aligned}$$

The proof of the claim is complete.

25° Let us call attention to a fundamental relation between linear mappings and determinants. We claim that L is invertible iff $det(L) \neq 0$. In fact, if L is invertible then:

$$1 = det(I) = det(L \cdot L^{-1}) = det(L)det(L^{-1})$$

It follows that $det(L) \neq 0$. Moreover, $det(L)$ and $det(L^{-1})$ are reciprocals of one another. Conversely, if L is not invertible then there must be a nonzero member B of $\mathbf{V}$ such that $L(B) = 0$. If it were not so then, by the Rank Theorem, we would find that:

$$dim(ker(L)) = 0 \quad \text{and} \quad dim(ran(L)) = n$$

so that L would be both injective and surjective, hence invertible. Of course, we may take B to be the first member of a basis $\mathcal{B}$ for $\mathbf{V}$:

$$\mathcal{B}: \quad B = B_1, B_2, \ \ldots \ , B_n$$

Borrowing the now familiar basic member Δ of $\mathbf{\Lambda}^n(\mathbf{V})$, we find that:

$$\begin{aligned}
0 &= \Delta(0, L(B_2), \ \ldots \ , L(B_n)) \\
&= L^*(\Delta)(B_1, B_2, \ \ldots \ , B_n) \\
&= det(L)\Delta(B_1, B_2, \ \ldots \ , B_n) \\
&= det(L)
\end{aligned}$$

26° Finally, let us engage in the computation of determinants. Of course, L is a linear mapping carrying $\mathbf{V}$ to itself. Let us compute $det(L)$. For now, let us assume that $n = 4$, so that $dim(\mathbf{V}) = 4$. Let us introduce a basis for $\mathbf{V}$:

$$\mathcal{B}: \quad B_1, B_2, B_3, B_4$$

As usual:

$$\begin{aligned}
L(B_1) &= m_{11}B_1 + m_{21}B_2 + m_{31}B_3 + m_{41}B_4 \\
L(B_2) &= m_{12}B_1 + m_{22}B_2 + m_{32}B_3 + m_{42}B_4 \\
L(B_3) &= m_{13}B_1 + m_{23}B_2 + m_{33}B_3 + m_{43}B_4 \\
L(B_4) &= m_{14}B_1 + m_{24}B_2 + m_{34}B_3 + m_{44}B_4
\end{aligned}$$

and:

$$\bar{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix}$$

is the rectangular (indeed, square) array defined by L relative to the bases $\mathcal{B}$ and $\mathcal{B}$ for $\mathbf{V}$ and $\mathbf{V}$. We contend that we can compute $det(L)$ directly from $\bar{M}$:

$$(\lambda) \qquad det(L) = \sum_{\sigma} |\sigma|\, m_{1\sigma(1)}m_{2\sigma(2)}m_{3\sigma(3)} = \sum_{\tau} |\tau|\, m_{\tau(1)1}m_{\tau(2)2}m_{\tau(3)3}$$

Both of the foregoing summations yield the sum of the following list of 24 numbers:

$$|\sigma|\, m_{1\sigma(1)}m_{2\sigma(2)}m_{3\sigma(3)}m_{4\sigma(4)} = |\tau|\, m_{\tau(1)1}m_{\tau(2)2}m_{\tau(3)3}m_{\tau(4)4}$$

where σ and τ run through $\mathbf{S}_4$, while $\sigma = \tau^{-1}$.

27° One may view the products in the foregoing list, in terms of the "neutral positions" of four rooks on a four by four chess board. In such a position, no one rook can attack any of the others. Consequently, no two of the rooks can occupy the same row or the same column. Let us label the rooks by the positive integers 1, 2, 3, and 4. Now every neutral position defines a permutation in $\mathbf{S}_4$, in two different ways, as follows:

$$\sigma(j) = k$$

iff the rook named j occupies the k-th column of the chess board. Just as well:

$$\tau(j) = k$$

iff the rook named j occupies the k-th row of the chess board. Obviously, σ and τ are inverse to one another. These permutations correspond to the following products in the foregoing list:

$$\begin{aligned}
\sigma &\longrightarrow m_{1\sigma(1)}m_{2\sigma(2)}m_{3\sigma(3)}m_{4\sigma(4)} \\
\tau &\longrightarrow m_{\tau(1)1}m_{\tau(2)2}m_{\tau(3)3}m_{\tau(4)4}
\end{aligned}$$

Of course, the plus and minus signs attached to the products reflect the signs of the corresponding permutations. One should note that σ and τ have the same sign, that is, both are even or both are odd.

28° Let us prove the relation (λ). We have:

$$\begin{aligned} det(L) &= det(L)\Delta(B_1, B_2, B_3, B_4) \\ &= L^*(\Delta)(B_1, B_2, B_3, B_4) \\ &= \Delta(L(B_1), L(B_2), L(B_3), L(B_4)) \end{aligned}$$

Since Δ is 4-linear, we find that $det(L)$ is the sum of the following $4^4 = 256$ products:

$$m_{a1}m_{b2}m_{c3}m_{d4}\,\Delta(B_a, B_b, B_c, B_d)$$
$$(1 \leq a \leq 4,\ 1 \leq b \leq 4,\ 1 \leq c \leq 4, 1 \leq d \leq 4)$$

However, if any two among a, b, c, and d are equal then the corresponding value of Δ is 0. Consequently, just 24 of the products survive:

$$m_{\tau(1)1}m_{\tau(2)2}m_{\tau(3)3}\tau_{4\tau(4)}\Delta(B_{\tau(1)}, B_{\tau(2)}, B_{\tau(3)}, B_{\tau(4)})$$

where τ runs through $\mathbf{S}_4$. Of course:

$$\Delta(B_{\tau(1)}, B_{\tau(2)}, B_{\tau(3)}, B_{\tau(4)}) = |\tau|$$

The proof of relation (λ) is complete.

29° The same pattern of argument will prove relation (λ) for the general case in which $dim(\mathbf{V}) = n$.

30° Let us show that the Determinant Mapping:

$$det : \mathbf{L}(\mathbf{V}) \longrightarrow \mathbf{F}$$

can be identified, through a choice of basis for $\mathbf{V}$, with a member δ of $\mathbf{\Lambda}^n(\mathbf{F}^n)$.

31° To be explicit, let us set $n = 4$. Let $\mathbf{V}$ be a linear space for which $dim(\mathbf{V}) = 4$. Let $\mathcal{B}$ be a basis for $\mathbf{V}$:

$$\mathcal{B}: \quad B_1, B_2, B_3, B_4$$

Let K be the linear isomorphism carrying $\mathbf{F}^4$ to $\mathbf{V}$, determined by $\mathcal{B}$:

$$K(E_1) = B_1,\ K(E_2) = B_2,\ K(E_3) = B_3,\ K(E_4) = B_4$$

Now let L be any linear mapping carrying $\mathbf{V}$ to $\mathbf{V}$:

$$L : \mathbf{V} \longrightarrow \mathbf{V}$$

let M be the matrix for L relative to the bases $\mathcal{B}$ and $\mathcal{B}$ for $\mathbf{V}$ and $\mathbf{V}$:

$$M = K^{-1} \cdot L \cdot K$$

and let $\bar{M}$ be the corresponding square array, having 4 rows and 4 columns, of numbers in $\mathbf{F}^4$:

$$\bar{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix}$$

In review:

$$L(B_k) = \sum_{j=1}^{4} m_{jk} B_j$$

where k is any relevant index $(1 \leq k \leq 4)$.

32° Finally, let X_1, X_2, X_3, and X_4 be the columns of $\bar{M}$:

$$X_1 = \begin{pmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{41} \end{pmatrix}, \; X_2 = \begin{pmatrix} m_{12} \\ m_{22} \\ m_{32} \\ m_{42} \end{pmatrix}, \; X_3 = \begin{pmatrix} m_{13} \\ m_{23} \\ m_{33} \\ m_{43} \end{pmatrix}, \; X_4 = \begin{pmatrix} m_{14} \\ m_{24} \\ m_{34} \\ m_{44} \end{pmatrix}$$

In this way, we identify:

$$\mathbf{L}(\mathbf{V}) \quad \text{and} \quad \mathbf{F}^4 \times \mathbf{F}^4 \times \mathbf{F}^4 \times \mathbf{F}^4$$

33° Now we introduce the member δ of $\mathbf{\Lambda}^4(\mathbf{F}^4)$, as follows::

$$\delta(X_1, X_2, X_3, X_4) = \begin{cases} \sum_\sigma |\sigma| \, m_{1\sigma(1)} m_{2\sigma(2)} m_{3\sigma(3)} m_{4\sigma(4)} \\ \text{or} \\ \sum_\tau |\tau| \, m_{\tau(1)1} m_{\tau(2)2} m_{\tau(3)3} m_{\tau(4)4} \end{cases}$$

Through the foregoing torrent of notation, we see that:

$$det(L) = \delta(X_1, X_2, X_3, X_4)$$

34° At this point, one might imagine a practical computational route to the definition of determinants. One may run the procedure from L through $\mathcal{B}$ to M and $\bar{M}$ in reverse. The foregoing relation shows the first step. However, proofs of the uniqueness of the determinant of L and of the basic relations involving products and inverses would be complicated.

07 Characteristic Values

01° Let $\mathbf{V}$ be a finite dimensional linear space and let L be a linear mapping carrying $\mathbf{V}$ to itself:

$$L : \mathbf{V} \longrightarrow \mathbf{V}$$

Let a be a number in $\mathbf{F}$. It may happen that there is a nonzero member X of $\mathbf{V}$ such that:

$$(c) \qquad L(X) = aX$$

In such a case, we refer to a as a *characteristic value* for L. In turn, the members Y of $\mathbf{V}$ for which $L(Y) = aY$ form a linear subspace $\mathbf{U}$ of $\mathbf{V}$. We refer to $\mathbf{U}$ as the *characteristic subspace* for L and to the members of $\mathbf{U}$ as *characteristic members* for L, relative to a.

02° Obviously, a is a characteristic value for L iff $aI - L$ is not invertible, which is to say that:

$$det(aI - L) = 0$$

03° Now we are led to introduce the *characteristic function* for L:

$$p(\zeta) = det(\zeta I - L)$$

where ζ is any number in $\mathbf{F}$. We contend that p is a polynomial.

04° To prove the contention, we reprise our conventional notation: L, $\mathcal{B}$, Δ. For simplicity, we set $n = 4$. We find that:

$$\begin{aligned} det(\zeta I - L) &= det(\zeta I - L)\Delta(B_1, B_2, B_3, B_4) \\ &= ((\zeta I - L)^*(\Delta))(B_1, B_2, B_3, B_4) \\ &= \Delta(\zeta B_1 - L(B_1), \zeta B_2 - L(B_2), \zeta B_3 - L(B_3), \zeta B_4 - L(B_4)) \\ &= \zeta^4 + c_3\zeta^3 + c_2\zeta^2 + c_1\zeta + c_0 \end{aligned}$$

where c_3, c_2, c_1, and c_0 are the sums of terms, as follows:

$$c_3 = \begin{array}{l} -\Delta(B_1, B_2, B_3, L(B_4)) \\ -\Delta(B_1, B_2, L(B_3), B_4) \\ -\Delta(B_1, L(B_2), B_3, B_4) \\ -\Delta(L(B_1), B_2, B_3, B_4) \end{array}$$

$$c_2 = \begin{array}{l} +\,\Delta(B_1, B_2, L(B_3), L(B_4)) \\ +\,\Delta(B_1, L(B_2), B_3, L(B_4)) \\ +\,\Delta(L(B_1), B_2, B_3, L(B_4)) \\ +\,\Delta(B_1, L(B_2), L(B_3), B_4) \\ +\,\Delta(L(B_1), B_2, L(B_3), B_4) \\ +\,\Delta(L(B_1), L(B_2), B_3, B_4) \end{array}$$

$$c_1 = \begin{array}{l} -\,\Delta(B_1, L(B_2), L(B_3), L(B_4)) \\ -\,\Delta(L(B_1), B_2, L(B_3), L(B_4)) \\ -\,\Delta(L(B_1), L(B_2), B_3, L(B_4)) \\ -\,\Delta(L(B_1), L(B_2), L(B_3), B_4) \end{array}$$

and:

$$c_0 = \; +\Delta(L(B_1), L(B_2), L(B_3), L(B_4))$$

The degree of p is the dimension of $\mathbf{V}$ and the leading coefficient of p is 1. By design, the characteristic values of L are the roots of p.

05° Just as well, we may pass to the square array $\bar{M}$ for L relative to $\mathcal{B}$. We find that:

$$det(\zeta I - L) = det\,(\begin{pmatrix} \zeta - m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & \zeta - m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & \zeta - m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & \zeta - m_{44} \end{pmatrix})$$

The former computation requires the assembly of $2^4 - 1 = 15$ terms, while the latter requires $4! - 1 = 23$. In general, the former requires $2^n - 1$, the latter $n! - 1$.

06° For the case in which $\mathbf{F} = \mathbf{R}$, the characteristic polynomial p may have no roots, so that L admits no characteristic values. However, for the case in which $\mathbf{F} = \mathbf{C}$, we may apply the Fundamental Theorem of Algebra to obtain a completely satisfactory conclusion. In fact, the characteristic polynomial would stand as follows:

$$p(\zeta) = (\zeta - a_1)^{n_1}(\zeta - a_2)^{n_2} \;\cdots\; (\zeta - a_r)^{n_r}$$

where a_1, a_2, ... , and a_r are the distinct roots of p. The exponents count the multiplicities of the roots.

08 The Theorem of Jordan

01° Now let us develop one of the central theorems of our subject:

THE THEOREM OF JORDAN

We will assume that $\mathbf{F} = \mathbf{C}$.

02° For the statement of the theorem, we require certain terminology and certain basic types of linear mapping. Let $\mathbf{V}$ be a finite dimensional linear space, having dimension n, and let L be any linear mapping in $\mathbf{L}(\mathbf{V})$. We say that L is *diagonalizable* iff there exists a basis:

$$\mathcal{B}: \quad B_1, B_2, \ \ldots \ , B_n$$

for $\mathbf{V}$ and there exists a corresponding array:

$$\mathcal{T}: \quad t_1, t_2, \ \ldots \ , t_n$$

of numbers in $\mathbf{C}$ such that, for each index j $(1 \leq j \leq n)$:

$$L(B_j) = t_j B_j$$

Obviously, the displayed numbers are the characteristic values of L. They are not necessarily distinct.

03° In turn, we say that L is *nilpotent* iff there is a positive integer k such that:

$$L^k = 0$$

One refers to the smallest such positive integer, let it be ν, as the *nilpotent degree* for L.

04° We also require the algebra $\mathbf{P}$ of polynomials, with coefficients in $\mathbf{C}$. The members of $\mathbf{P}$ have the following familiar form:

$$f(\zeta) = \sum_{j=0}^{d} c_j \zeta^j$$

where ζ represents an arbitrary complex number. Presuming that $c_d \neq 0$, we declare that the *degree* of f is d. The operations of addition, scalar multiplication, and multiplication stand as follows:

$$\begin{aligned} (f+g)(\zeta) &= f(\zeta) + g(\zeta) \\ (ch)(\zeta) &= ch(\zeta) \\ (fg)(\zeta) &= f(\zeta)g(\zeta) \end{aligned}$$

Under these operations, **P** is a commutative algebra.

05° Let us note that **L**(**V**) is also an algebra, but it fails to be commutative. By article 14° in **Section 04**, we know that $dim(\mathbf{V}) = n^2$, where n is the dimension of **V**.

06° Now let f be any polynomial in **P** and let L be any linear mapping in **L**(**V**). Granted the foregoing form for f, we may apply f to L, obtaining a linear mapping $f(L)$ in **L**(**V**) as follows:

$$f(L) = \sum_{j=0}^{d} c_j L^j$$

We refer to it as a *polynomial in* L.

07° By patient computation, we find that:

$$\begin{aligned}(f+g)(L) &= f(L) + g(L)\\ (ch)(L) &= ch(L)\\ (fg)(L) &= f(L)g(L)\end{aligned}$$

08° Now let **V** be a finite dimensional linear space, having dimension n. The Theorem of Jordan asserts that, for any linear mapping L in **L**(**V**), there exist a diagonalizable linear mapping L° and a nilpotent linear mapping $L^\bullet$ in **L**(**V**) such that:

(JT) $$L = L^\circ + L^\bullet$$

Both L° and $L^\bullet$ are polynomials in L. Moreover, under the stated conditions, L° and $L^\bullet$ are unique.

09° For the proof of the theorem, we require three elements:

- (•) the Theorem of Cayley and Hamilton
- (•) the concept of Direct Sum Decomposition
- (•) the Zero Places Theorem of Hilbert

The sense of the first element is simple:

(CH) $$p(L) = 0$$

where p is the characteristic polynomial for L:

$$p(\zeta) = det(\zeta - L)$$

This remarkable, though elementary result is critical to the argument which follows. To prove it, we return to article 33° in **Section 06** and to article 05° in the preceding section. (For those articles, we had set $n = 4$.)

10° ......

11° Now let consider the second element. We mean the concept of Direct Sum Decomposition of $\mathbf{V}$. It is in fact a natural generalization of the concept of basis for $\mathbf{V}$. Let $\mathcal{U}$ be a list of linear subspaces of $\mathbf{V}$.

$$\mathcal{U}: \quad \mathbf{U}_1, \mathbf{U}_2, \ \ldots \ , \mathbf{U}_r$$

We say that $\mathcal{U}$ generates $\mathbf{V}$ iff, for each member X of $\mathbf{V}$, there are members:

$$X_1, X_2, \ \ldots \ , X_r$$

of $\mathbf{V}$ such that, for each index j $(1 \leq j \leq r)$, $X_j \in \mathbf{U}_j$ and such that:

$$(*) \qquad X = X_1 + X_2 + \ \cdots \ + X_r$$

We say that $\mathcal{U}$ is independent iff, for any members:

$$X_1, X_2, \ \ldots \ , X_4 4$$

of $\mathbf{V}$, if, for each index j $(1 \leq j \leq r)$, $X_j \in \mathbf{U}_j$ and if:

$$X_1 + X_2 + \ \cdots \ + X_r = 0$$

then, for each index j $(1 \leq j \leq r)$, $X_j = 0$.

12° Obviously, if $\mathcal{U}$ is independent and if $\mathcal{U}$ generates $\mathbf{V}$ then every member of $\mathbf{V}$ can be presented, uniquely, in the form displayed in relation $(*)$. In this context, we say that $\mathcal{U}$ defines a Direct Sum Decomposition of $\mathbf{V}$. We summarize this complex of relations as follows:

$$\mathbf{V} = \mathbf{U}_1 \oplus \mathbf{U}_2 \oplus \ \cdots \ \oplus \mathbf{U}_r$$

13° Finally, let us turn to the third element, a Theorem of Hilbert, often called "der NullStellenSatz." Let $\mathcal{H}$ be a finite list of polynomials in $\mathbf{P}$:

$$\mathcal{H}: \quad h_1, h_2, \ \ldots \ , h_r$$

Let us assume that, for each number ζ in $\mathbf{C}$, at least one of the numbers:

$$h_1(\zeta), h_2(\zeta), \ \ldots \ , h_r(\zeta)$$

is nonzero. We claim that there exists a corresponding list $\mathcal{F}$ of polynomials in $\mathbf{P}$:

$$\mathcal{F}: \quad f_1, f_2, \ \ldots \ , f_r$$

such that:

$$\text{(HN)} \qquad f_1 h_1 + f_2 h_2 + \ \cdots \ + f_r h_r = 1$$

We will sketch a proof of the claim at the end of this section.

14° Let us return to our original objective: to prove the Theorem of Jordan. Let $\mathbf{V}$ be a finite dimensional linear space, having dimension n, and let L be any linear mapping in $\mathbf{L}(\mathbf{V})$. We must show that there exist polynomials $L^\circ = f^\circ(L)$ and $L^\bullet = f^\bullet(L)$ in L, the first diagonalizable and the second nilpotent, such that:

$$L = L^\circ + L^\bullet$$

and we must show that, subject to the stated conditions, L° and $L^\bullet$ are unique.

15° Let us prove existence. To that end, we introduce the polynomials:

$$\begin{array}{ll} p_1(\zeta) = (\zeta - a_1)^{n_1}, & \hat{p}_1(\zeta) \\ p_2(\zeta) = (\zeta - a_2)^{n_2}, & \hat{p}_2(\zeta) \\ \vdots & \vdots \\ p_r(\zeta) = (\zeta - a_r)^{n_r}, & \hat{p}_r(\zeta) \end{array}$$

related as follows:

$$\begin{array}{c} p_1(\zeta)\hat{p}_1(\zeta) = p(\zeta) \\ p_2(\zeta)\hat{p}_2(\zeta) = p(\zeta) \\ \vdots \\ p_r(\zeta)\hat{p}_r(\zeta) = p(\zeta) \end{array}$$

For the case in which $r = 4$, we would have:

$$\begin{aligned}
\hat{p}_1(\zeta) &= p_2(\zeta)p_3(\zeta)p_4(\zeta)\\
\hat{p}_2(\zeta) &= p_1(\zeta)p_3(\zeta)p_4(\zeta)\\
\hat{p}_3(\zeta) &= p_1(\zeta)p_2(\zeta)p_4(\zeta)\\
\hat{p}_4(\zeta) &= p_1(\zeta)p_2(\zeta)p_3(\zeta)
\end{aligned}$$

16° Obviously, the polynomials in the list:

$$\hat{\mathcal{P}}: \quad \hat{p}_1, \hat{p}_2, \ \ldots \ , \hat{p}_r$$

have no zeros in common. By the Theorem of Hilbert, we may introduce a list:

$$\mathcal{F}: \ f_1, f_2, \ \ldots \ , f_r$$

of polynomials such that:

$$f_1\hat{p}_1 + f_2\hat{p}_2 + \ \cdots \ + f_r\hat{p}_r = 1$$

Finally, we introduce the following linear mappings in $\mathbf{L}(\mathbf{V})$:

$$\begin{array}{lll}
P_1 = a_1 I, & Q_1 = L - a_1 I, & \Pi_1 = f_1(L)\hat{p}_1(L)\\
P_2 = a_2 I, & Q_2 = L - a_2 I, & \Pi_2 = f_2(L)\hat{p}_2(L)\\
\quad\vdots & \quad\vdots & \quad\vdots\\
P_r = a_s I, & Q_r = L - a_s I, & \Pi_s = f_s(L)\hat{p}_r(L)
\end{array}$$

These mappings will provide a clear proof of the Theorem of Jordan.

17° In fact, we contend that the linear mappings:

$$(\star) \qquad \begin{aligned}
L^\circ &= P_1\Pi_1 + P_2\Pi_2 + \ \cdots \ + P_r\Pi_r\\
L^\bullet &= Q_1\Pi_1 + Q_2\Pi_2 + \ \cdots \ + Q_r\Pi_r
\end{aligned}$$

serve the purposes of the theorem.

18° Let us emphasize that all the linear mappings now under consideration are polynomials in L. As a result, any two of them commute:

$$L'L'' = L''L'$$

19° By elementary observations, we find that:

$$\Pi_1 + \Pi_2 + \cdots + \Pi_r = I$$

Moreover, for any indices j and k ($1 \leq j \leq r$, $1 \leq k \leq r$), if $j \neq k$ then p divides $\hat{p}_j\hat{p}_k$. By the Theorem of Cayley and Hamilton, we have $p(L) = 0$. Hence:

$$\Pi_j\Pi_k = 0$$

It follows that, for any index ℓ ($1 \leq \ell \leq s$):

$$\Pi_\ell^2 = \Pi_\ell\Pi_\ell = \Pi_\ell$$

Consequently, the list $\mathcal{P}$ consisting of the ranges of $\Pi_1, \Pi_2, \ldots$, and Π_r defines a direct sum decomposition of $\mathbf{V}$:

$$\mathbf{V} = ran(\Pi_1) \oplus ran(\Pi_2) \oplus \cdots \oplus ran(\Pi_r)$$

For each X in $\mathbf{V}$, we find that:

$$X = X_1 + X_2 + \cdots + X_r$$

where, for each index j ($1 \leq j \leq r$):

$$X_j = \Pi_j(X)$$

The linear mapping Π_j serves as a *projection*, which assigns to each member X of $\mathbf{V}$ its (unique) "representative" in $ran(\Pi_j)$.

20° For each index ℓ ($1 \leq \ell \leq r$):

$$(Q_\ell\Pi_\ell)^{n_\ell} = Q_\ell^{n_\ell}\Pi_\ell = p_\ell(L)f_\ell(L)\hat{p}_\ell(L) = f_\ell(L)p(L) = 0$$

Hence, $Q_\ell\Pi_\ell$ is nilpotent. It follows that $L^\bullet$ is nilpotent.

21° In turn, it is obvious that L° is diagonalizable. Moreover, the characteristic values of L° are the roots of p.

22° Let us prove uniqueness. ....

23° Let us sketch, very quickly, a proof of the Theorem of Hilbert. Let $\mathbf{P}$ be the algebra of polynomials with coefficients in $\mathbf{C}$. Let $\mathcal{H}$ be a finite list of (nonzero) polynomials in $\mathbf{P}$:

$$\mathcal{H}: \ \ h_1, h_2, \ \dots \ , h_r$$

and let $\mathbf{J}$ be the subset of $\mathbf{P}$ consisting of all polynomials of the form:

$$f_1 h_1 + f_2 h_2 \ \cdots \ + f_r h_r$$

where $\mathcal{F}$ is any (finite) list of polynomials in $\mathbf{P}$:

$$\mathcal{F}: \ \ f_1, f_2, \ \dots \ , f_r$$

Obviously, for any polynomial f in $\mathbf{P}$ and for any polynomials g, g_1, and g_2 in $\mathbf{J}$, fg and $g_1 + g_2$ are in $\mathbf{J}$. Now let q be a polynomial in $\mathbf{J}\backslash\{0\}$ for which the degree, let it be ν, is the smallest among the degrees of all polynomials in $\mathbf{J}\backslash\{0\}$. In turn, let g be any polynomial in $\mathbf{J}$. By division in $\mathbf{P}$, we may introduce polynomials δ and ρ in $\mathbf{P}$ such that:

$$g = \delta q + \rho$$

where the degree of ρ is less than ν. Clearly, $g - \delta q$ must be in $\mathbf{J}$. Hence, ρ must be 0. We infer that $\mathbf{J}$ consists of all and only multiples of q::

$$\mathbf{J} = \mathbf{P}\, q$$

Now it is plain that, if the polynomials in the list $\mathcal{H}$ have no common zeros in $\mathbf{C}$, then q is constant. In such a case, there would be a list $\mathcal{F}$ for which:

$$f_1 h_1 + f_2 h_2 \ \cdots \ + f_s h_s = 1$$

This result figured in our proof of the Theorem of Jordan.

09 Positive Definite Orthogonal Geometries

01° We plan to study the relation between Linear Algebra and Geometry. With reference to article 09° in **Section 06**, we distinguish, by symmetry and by antisymmetry, two basic cases:

ORTHOGONAL GEOMETRY and SYMPLECTIC GEOMETRY

Mindful of scope and of practical applications, we will concentrate upon the special case of Positive Definite Orthogonal Geometry.

02° Hereafter, we will assume that $\mathbf{F} = \mathbf{R}$. Let $\mathbf{V}$ be a linear space, having finite dimension n, and let Γ be a 2-linear functional in $\mathbf{\Lambda}^2(\mathbf{V})$. We will refer to Γ as a *bilinear form*. For the various members X and Y in $\mathbf{V}$, we prefer to write:

$$\text{not} \quad \Gamma(X, Y) \quad \text{but} \quad \langle\!\langle\, X, Y \,\rangle\!\rangle$$

At the outset, we require that Γ be *nondegenerate*. We mean to say that:

for each X in $\mathbf{V}$,
if $X \neq 0$ then there exists Y in $\mathbf{V}$ such that $\langle\!\langle\, X, Y \,\rangle\!\rangle \neq 0$

for each Y in $\mathbf{V}$,
if $Y \neq 0$ then there exists X in $\mathbf{V}$ such that $\langle\!\langle\, X, Y \,\rangle\!\rangle \neq 0$

Subject to the requirement just stated, we refer to the ordered pair $(\mathbf{V}, \Gamma)$ as a *geometry*.

03° For the case in which Γ is symmetric, we refer to the geometry $(\mathbf{V}, \Gamma)$ as an *orthogonal* geometry. For the case in which Γ is antisymmetric, we refer to the geometry $(\mathbf{V}, \Gamma)$ as a *symplectic* geometry.

04° For a given orthogonal geometry $(\mathbf{V}, \Gamma)$, it may happen that, for each X in $\mathbf{V}$, if $X \neq 0$ then:

$$(*) \qquad 0 < \langle\!\langle\, X, X \,\rangle\!\rangle$$

In such a case, we refer to the geometry $(\mathbf{V}, \Gamma)$ as a *positive definite* orthogonal geometry. Hereafter, we focus our attention exclusively upon such geometries. We will refer to them simply by mentioning the underlying linear space $\mathbf{V}$, taking for granted that $\mathbf{V}$ has been supplied, in some manner, with a positive definite symmetric bilinear form and, in turn, presenting the values of the form as follows:

$$\langle\!\langle\, X, Y \,\rangle\!\rangle$$

05° For a simple example, we need only cite the familiar case of $\mathbf{R}^3$, supplied with the standard *inner product*:

$$\langle\!\langle\, \mathbf{x}, \mathbf{y} \,\rangle\!\rangle = x_1y_1 + x_2y_2 + x_3y_3$$

where:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

06° Let $\mathbf{V}$ be a positive definite orthogonal (pdo) geometry. Let n be the dimension of $\mathbf{V}$. For each Z in $\mathbf{V}$, we introduce, as usual, the *norm* of Z:

$$\|Z\| = \sqrt{\langle\!\langle\, Z, Z \,\rangle\!\rangle}$$

and the *normalization* of Z:

$$\hat{Z} = \frac{1}{\|Z\|} Z \qquad (\|\hat{Z}\| = 1)$$

Of course, for the latter, we require that $Z \neq 0$.

07° By simple (very familiar) arguments, we find that, for any X and Y in $\mathbf{V}$:

$$|\langle\!\langle\, X, Y \,\rangle\!\rangle| \leq \|X\|\|Y\| \quad \text{and} \quad \|X + Y\| \leq \|X\| + \|Y\|$$

08° For our first step forward, we concentrate upon the special character of certain bases for $\mathbf{V}$, which figure in the fundamental theorems to follow. Let $\mathcal{C}$ be a basis for $\mathbf{V}$:

$$\mathcal{C}: \quad C_1, C_2, \ \ldots\ , C_n$$

We say that $\mathcal{C}$ is an *orthonormal* basis iff:

$$\langle\!\langle\, C_j, C_k \,\rangle\!\rangle = \begin{cases} 0 & \text{if } j \neq k \\ 1 & \text{if } j = k \end{cases}$$

09° Relative to such a basis, we can calculate coordinates very easily. For each X in $\mathbf{V}$:

$$X = x_1C_1 + x_2C_2 + \ \cdots\ + x_nC_n$$

we find that, for each index j $(1 \leq j \leq n)$:

$$x_j = \langle\!\langle\, X, C_j \,\rangle\!\rangle$$

10° Orthonormal bases are in one sense very special, in another sense rather common. In any case, we contend that, for any basis $\mathcal{B}$ for $\mathbf{V}$:

$$\mathcal{B}: \quad B_1, B_2, \ldots, B_n$$

we may "convert" $\mathcal{B}$ to an orthonormal basis $\mathcal{C}$ for $\mathbf{V}$ while causing "minimal disturbance". The latter (vaguely phrased) condition means that, for any index k ($1 \leq k \leq n$):

$$B_k \quad \text{depends on} \quad C_1, C_2, \ldots, C_k$$

and:

$$C_k \quad \text{depends on} \quad B_1, B_2, \ldots, B_k$$

Just as well, the condition means that, for any index k ($1 \leq k \leq n$), the linear subspaces of $\mathbf{V}$ generated by:

$$B_1, B_2, \ldots, B_k \quad \text{and} \quad C_1, C_2, \ldots, C_k$$

respectively, are the same.

11° Let us prove the contention. In fact, let us describe an algorithm for converting an arbitrary basis $\mathcal{B}$ to an orthonormal basis $\mathcal{C}$. We set $n = 4$. The pattern of computation runs as follows:

$$\begin{aligned}
&C_1 = B_1 \\
&C_2 = B_2 - \langle\!\langle B_2, \hat{C}_1 \rangle\!\rangle \hat{C}_1 \\
&C_3 = B_3 - \langle\!\langle B_3, \hat{C}_2 \rangle\!\rangle \hat{C}_2 - \langle\!\langle B_3, \hat{C}_1 \rangle\!\rangle \hat{C}_1 \\
&C_4 = B_4 - \langle\!\langle B_4, \hat{C}_3 \rangle\!\rangle \hat{C}_3 - \langle\!\langle B_4, \hat{C}_2 \rangle\!\rangle \hat{C}_2 - \langle\!\langle B_4, \hat{C}_1 \rangle\!\rangle \hat{C}_1
\end{aligned}$$

By this pattern of computation, we obtain the orthonormal basis:

$$\mathcal{C}: \quad \hat{C}_1, \hat{C}_2, \hat{C}_3, \hat{C}_4$$

Obviously, the condition of "minimum disturbance" is satisfied.

10 The Spectral Theorem

01° Let $\mathbf{V}'$ and $\mathbf{V}''$ be pdo geometries. Let L' and L'' be linear mappings, the first carrying $\mathbf{V}'$ to $\mathbf{V}''$ and the second carrying $\mathbf{V}''$ to $\mathbf{V}'$:

$$L' \in \mathbf{L}(\mathbf{V}', \mathbf{V}''), \quad L'' \in \mathbf{L}(\mathbf{V}'', \mathbf{V}')$$

We say that L' and L'' are *compatible* iff, for any members X of $\mathbf{V}'$ and Y of $\mathbf{V}''$:

$$\langle\!\langle\, L'(X), Y \,\rangle\!\rangle = \langle\!\langle\, X, L''(Y) \,\rangle\!\rangle$$

To express this relation, we will write:

$$L' \approx L''$$

02° We contend that, for each L', there is precisely one L'' such that $L' \approx L''$. Of course, it would follow that, for each L'', there is precisely one L' such that $L' \approx L''$.

03° The argument for uniqueness is simple. We imagine linear mappings L', L''_1, and L''_2 such that $L' \approx L''_1$ and $L' \approx L''_2$. Clearly, for any X in $\mathbf{V}'$ and Y in $\mathbf{V}''$:

$$\langle\!\langle\, X, L''_1(Y) - L''_2(Y) \,\rangle\!\rangle = \langle\!\langle\, L'(X), Y \,\rangle\!\rangle - \langle\!\langle\, L'(X), Y \,\rangle\!\rangle = 0$$

Since the bilinear form on $\mathbf{V}'$ is nondegenerate, we infer that $L''_1(Y) = L''_2(Y)$. Hence, $L''_1 = L''_2$.

04° The argument for existence requires further developments. We define linear mappings Ω' carrying $\mathbf{V}'$ to $\mathbf{V}'^*$ and Ω'' carrying $\mathbf{V}''$ to $\mathbf{V}''^*$, as follows:

$$\Omega'(X'')(X') = \langle\!\langle\, X', X'' \,\rangle\!\rangle, \qquad \Omega''(Y'')(Y') = \langle\!\langle\, Y', Y'' \,\rangle\!\rangle$$

where X'' and X' are any members of $\mathbf{V}'$ and Y'' and Y' are any members of $\mathbf{V}''$. Since the bilinear forms on $\mathbf{V}'$ and $\mathbf{V}''$ are nondegenerate, we find that Ω' and Ω'' are injective. Since the dimensions of $\mathbf{V}'$ and $\mathbf{V}'^*$ are the same and the dimensions of $\mathbf{V}''$ and $\mathbf{V}''^*$ are the same, we infer that Ω' and Ω'' are in fact bijective.

05° In practical terms, we infer that, for each Λ' in $\mathbf{V}'^*$, there is precisely one X'' in $\mathbf{V}'$ such that $\Omega'(X'') = \Lambda'$, while for each Λ'' in $\mathbf{V}''^*$, there is precisely one Y'' in $\mathbf{V}''$ such that $\Omega''(Y'') = \Lambda''$. These assertions mean that, for each X' in $\mathbf{V}'$ and for each Y' in $\mathbf{V}''$:

$$(*) \qquad \Lambda'(X') = \langle\!\langle\, X', X'' \,\rangle\!\rangle \quad \text{and} \quad \Lambda''(Y') = \langle\!\langle\, Y', Y'' \,\rangle\!\rangle$$

06° Now let L' be any linear mapping in $\mathbf{L}(\mathbf{V}', \mathbf{V}'')$. For any Y'' in $\mathbf{V}''$, we find that $\Omega''(Y'')$ lies in $\mathbf{V}''^*$ and $\Omega''(Y'') \cdot L'$ lies in $\mathbf{V}'^*$. Hence, there is precisely one X'' in $\mathbf{V}'$ such that:

$$\Omega'(X'') = \Omega''(Y'') \cdot L'$$

Denoting X'' by $L''(Y'')$ and unwinding the notation, we infer that, for any Y'' in $\mathbf{V}''$ and any X' in $\mathbf{V}'$:

$$\langle\!\langle\, X', L''(Y'') \,\rangle\!\rangle = \langle\!\langle\, L'(X'), Y'' \,\rangle\!\rangle$$

In this way we have succeeded in defining a linear mapping L'' in $\mathbf{L}(\mathbf{V}'', \mathbf{V}')$ for which $L' \approx L''$. The argument is complete.

07° Let us introduce the conventional notation and terminology:

$$L' \approx L'' \quad \longleftrightarrow \quad L'' = L'^* \text{ and } L' = L''^*$$

We say that L' and L'' are *adjoints* of one another.

08° To understand the relation:

$$L' \approx L''$$

let us consider corresponding matrices for L' and L''. Let the dimensions of $\mathbf{V}'$ and $\mathbf{V}''$ be n' and n'', respectively. Let:

$$\mathcal{C}' : \quad C'_1, C'_2, \,\ldots\, , C'_{n'}\,; \qquad \mathcal{C}'' : \quad C''_1, C''_2, \,\ldots\, , C''_{n''}$$

be orthonormal bases for $\mathbf{V}'$ and $\mathbf{V}''$, respectively. Clearly, the entries in the corresponding rectangular arrays stand as follows:

$$m'_{kj} = \langle\!\langle\, L'(C'_j), C''_k \,\rangle\!\rangle, \qquad m''_{jk} = \langle\!\langle\, L''(C''_k), C'_j \,\rangle\!\rangle$$

where j and k are any indices for which $1 \leq j \leq n'$ and $1 \leq k \leq n''$. Now the relation $L' \approx L''$ means that:

$$m'_{kj} = m''_{jk}$$

09° Now let us assume that $\mathbf{V}'$ and $\mathbf{V}''$ are the same. Accordingly, we may introduce a pdo geometry $\mathbf{V}$ such that $\mathbf{V}' = \mathbf{V} = \mathbf{V}''$. Let S and T be linear mappings in $\mathbf{L}(\mathbf{V})$ for which $S \approx T$. So S and T are adjoints of one another. We say that S is *self adjoint* iff $S = T$, that is:

$$S = S^*$$

The condition means that, for any X and Y in $\mathbf{V}$:

$$\langle\!\langle\, S(X), Y \,\rangle\!\rangle = \langle\!\langle\, X, S(Y) \,\rangle\!\rangle$$

It is the same to say that $T = T^*$. Now the condition on the corresponding quadratic array takes the following form:

$$s_{kj} = s_{jk}$$

10° Let us turn to describe another of the central theorems of our subject:

THE SPECTRAL THEOREM

It states that, in effect, every self adjoint linear mapping is diagonalizable. But we can describe the matter much more clearly.

11° For precision of expression, we present a refinement of the concept of direct sum decomposition. Let $\mathbf{V}$ be a pdo geometry. Let $\mathcal{U}$ be a list:

$$\mathcal{U}: \quad \mathbf{U}_1, \mathbf{U}_2, \ \ldots\ , \mathbf{U}_r$$

of linear subspaces of $\mathbf{V}$ which defines a direct sum decomposition of $\mathbf{V}$, but which, in addition, meets the condition that, for any indices j and k ($1 \leq j \leq s$, $1 \leq k \leq s$), if $j \neq k$ then $\mathbf{U}_j$ and $\mathbf{U}_k$ are orthogonal to one another:

$$\mathbf{U}_j \perp \mathbf{U}_k$$

We mean to say that, for any X in $\mathbf{U}_j$ and Y in $\mathbf{U}_k$, $\langle\!\langle\, X, Y \,\rangle\!\rangle = 0$. For such a case, we declare that $\mathcal{U}$ defines an Orthogonal Direct Sum Decomposition of $\mathbf{V}$. We summarize this complex of relations as follows:

$$\mathbf{V} = \mathbf{U}_1 \perp \mathbf{U}_2 \perp \ \cdots\ \perp \mathbf{U}_r$$

12° Here is a simple example of such a decomposition. Let $\mathbf{U}$ be a linear subspace of $\mathbf{V}$. Let the dimensions of $\mathbf{U}$ and $\mathbf{V}$ be d and n, respectively. Let $\mathbf{U}^\perp$ be the linear subspace of $\mathbf{V}$ composed of all Y in $\mathbf{V}$ such that, for all X in $\mathbf{U}$, $\langle\!\langle\, X, Y \,\rangle\!\rangle = 0$. We refer to $\mathbf{U}^\perp$ as the *orthogonal complement* of $\mathbf{U}$ in $\mathbf{V}$. We claim that:

$$\mathbf{V} = \mathbf{U} \perp \mathbf{U}^\perp$$

To prove the claim, we introduce a basis $\mathcal{B}'$ for $\mathbf{U}$, extend $\mathcal{B}'$ to a basis $\mathcal{B}$ for $\mathbf{V}$:

$$\mathcal{B}: \quad B'_1, B'_2, \ \ldots\ , B'_d; \quad B''_1, B''_2, \ \ldots\ , B''_{n-d}$$

then convert $\mathcal{B}$ to an orthonormal basis $\mathcal{C}$ for $\mathbf{V}$:

$$\mathcal{C}: \quad C'_1, C'_2, \dots, C'_d; \quad C''_1, C''_2, \dots, C''_{n-d}$$

while meeting the condition of minimum disturbance. Obviously:

$$\mathcal{C}': \quad C'_1, C'_2, \dots, C'_d \quad \text{and} \quad \mathcal{C}'': \quad C''_1, C''_2, \dots, C''_{n-d}$$

are orthonormal bases for $\mathbf{U}$ and $\mathbf{U}^\perp$, respectively. Now one may complete the proof by simple observations.

13° Let $\mathbf{V}$ be a pdo geometry, having dimension n. Let S be a self adjoint linear mapping in $\mathbf{L}(\mathbf{V})$. In precise terms, the Spectral Theorem states that there exist an orthonormal basis $\mathcal{C}$ for $\mathbf{V}$:

$$\mathcal{C}: \; C_1, C_2, \dots, C_n$$

and a list $\mathcal{T}$ of real numbers:

$$\mathcal{T}: \; t_1, t_2, \dots, t_n$$

such that, for each index k ($1 \leq k \leq n$):

$$\text{(ST)} \qquad S(C_k) = t_j C_k$$

14° Obviously, the numbers in the list $\mathcal{T}$ are characteristic values for S. However, as displayed, they may not be distinct. In any case, we claim that there are no others. To prove the claim, we imagine a real number a and a nonzero member X of $\mathbf{V}$ such that $S(X) = aX$. Of course, there must be an index j ($1 \leq k \leq n$) such that $\langle\!\langle X, C_k \rangle\!\rangle \neq 0$. It would follow that:

$$a\langle\!\langle X, C_k \rangle\!\rangle = \langle\!\langle S(X), C_k \rangle\!\rangle = \langle\!\langle X, S(C_k) \rangle\!\rangle = t_k \langle\!\langle X, C_k \rangle\!\rangle$$

Consequently, $a = t_k$.

15° Let us compress the list $\mathcal{T}$, so that the entries are mutually distinct:

$$\mathcal{T}: \; \tau_1, \tau_2, \dots, \tau_r$$

For each index j ($1 \leq j \leq r$), let $\mathbf{U}_j$ be the characteristic subspace of $\mathbf{V}$ corresponding to τ_j. By definition, $\mathbf{U}_j$ consists of all members X of $\mathbf{V}$ for which $S(X) = \tau_j X$. Obviously, for any indices j and k ($1 \leq j \leq r$, $1 \leq k \leq r$), if $j \neq k$ then:

$$\mathbf{U}_j \perp \mathbf{U}_k$$

because, for any X in $\mathbf{U}_j$ and Y in $\mathbf{U}_k$:

$$\tau_j \langle\!\langle X, Y \rangle\!\rangle = \langle\!\langle S(X), Y \rangle\!\rangle = \langle\!\langle X, S(Y) \rangle\!\rangle = \tau_k \langle\!\langle X, Y \rangle\!\rangle$$

16° The foregoing developments show that the list:

$$\mathcal{U}: \ \ \mathbf{U}_1, \mathbf{U}_2, \ \dots \ , \mathbf{U}_r$$

of subspaces of $\mathbf{V}$ defines an orthogonal direct sum decomposition of $\mathbf{V}$:

$$\mathbf{V} = \mathbf{U}_1 \perp \mathbf{U}_2 \perp \ \cdots \ \perp \mathbf{U}_r$$

Reflecting upon the presentation of the Theorem of Jordan in **Section 07**, we introduce the following linear mappings in $\mathbf{L}(\mathbf{V})$:

$$\begin{array}{cc} P_1 = \tau_1 I, & \Pi_1 \\ P_2 = \tau_2 I, & \Pi_2 \\ \vdots & \vdots \\ P_r = \tau_r I, & \Pi_r \end{array}$$

The various linear mappings Π_j $(1 \leq j \leq r)$ are defined, by *projection*, as follows:

$$X = \Pi_1(X) + \Pi_2(X) + \ \cdots \ + \Pi_r(X)$$

where X is any member of $\mathbf{V}$. We mean that, for each index j $(1 \leq j \leq r)$, $\Pi_j(X)$ is the representative of X in $\mathbf{U}_j$. Obviously:

$$\Pi + \Pi + \ \cdots \ \Pi_s = I$$

One can easily verify that:

$$\Pi_k \Pi_k = \Pi_k, \ \ \Pi_k^* = \Pi_k, \ \ ran(\Pi_k) = \mathbf{U}_k, \ \ ker(\Pi_k) = \mathbf{U}_k^{\perp}$$

We say that the list $\mathcal{P}$ of self adjoint projections:

$$\mathcal{P}: \ \ \Pi_1, \Pi_2, \ \dots \ , \Pi_s$$

forms a *resolution of the identity* in $\mathbf{V}$. We obtain the following elegant presentation of S:

$$S = P_1 \Pi_1 + P_2 \Pi_2 + \ \cdots \ + P_s \Pi_s$$

17° Let us prove the Spectral Theorem. To that end, we will state a basic fact, then proceed to prove the theorem by induction. In the last section, **Section 12**, we will prove that basic fact, by borrowing a well known theorem from Multivariable Calculus.

18° Let $\mathbf{V}$ be a pdo geometry, having dimension n. Let S be a self adjoint linear mapping in $\mathbf{L}(\mathbf{V})$. We must design an orthonormal basis $\mathcal{C}$ for $\mathbf{V}$:

$$\mathcal{C}: \quad C_1, C_2, \ \dots \ , C_n$$

and a list $\mathcal{T}$ of real numbers:

$$\mathcal{T}: \quad t_1, t_2, \ \dots \ , t_n$$

such that, for each index k $(1 \leq k \leq n)$:

$$S(C_k) = t_j C_k$$

We will argue by induction. However, in process, we will assume a basic fact: that S admits at least one characteristic value. We will prove that fact, independently, in the last section. For the case in which $n = 1$, we declare the theorem to be obvious. Let m be any positive integer. Let us assume that the theorem holds true for the case in which $n = m$. Let us prove that it holds true for the case in which $n = m + 1$. To begin, we apply the foregoing basic fact. We introduce a real number t and a member C of $\mathbf{V}$ such that:

$$S(C) = t\,C, \qquad \langle\!\langle\, C, C \,\rangle\!\rangle = 1$$

In turn, we introduce the linear subspaces $\mathbf{U} = \mathbf{R}\,C$ and $\mathbf{U}^{\perp}$ of $\mathbf{V}$. The former consists of all scalar multiples of C while the latter consists of all members Y in $\mathbf{V}$ for which $\langle\!\langle\, C, Y \,\rangle\!\rangle = 0$. Of course, the dimension of $\mathbf{U}^{\perp}$ is m. The corresponding list defines an orthogonal direct sum decomposition of $\mathbf{V}$:

$$\mathbf{V} = \mathbf{R}\,C \perp \mathbf{U}^{\perp}$$

We find that, for each Y in $\mathbf{V}$, if Y is in $\mathbf{U}^{\perp}$ then $S(Y)$ is in $\mathbf{U}^{\perp}$ as well, because:

$$\langle\!\langle\, C, S(Y) \,\rangle\!\rangle = \langle\!\langle\, S(C), Y \,\rangle\!\rangle = a \langle\!\langle\, C, Y \,\rangle\!\rangle = 0$$

Now we may apply the induction hypothesis, to introduce an orthonormal basis $\mathcal{C}^{\perp}$ for $\mathbf{U}^{\perp}$:

$$\mathcal{C}^{\perp}: \quad C_2, C_3, \ \dots \ , C_n$$

and a list $\mathcal{T}$ of real numbers:

$$\mathcal{T}: \quad t_2, t_3, \ \dots \ , t_n$$

such that, for each index k $(2 \leq k \leq n)$:

$$S(C_k) = t_k C_k$$

The proof is complete.

11 The Singular Value Decomposition

01° Let $\mathbf{V}'$ and $\mathbf{V}''$ be pdo geometries. Let p and q be the dimensions of $\mathbf{V}'$ and $\mathbf{V}''$, respectively. Let L be any linear mapping in $\mathbf{L}(\mathbf{V}', \mathbf{V}'')$. Let r be the dimension of the range of L. Obviously, $r \leq min\{p, q\}$. We contend that there exist orthonormal bases:

$$\mathcal{C}' : \; C'_1, C'_2, \, \dots \, , C'_p \, ; \qquad \mathcal{C}'' : \; C''_1, C''_2, \, \dots \, , C''_q$$

for $\mathbf{V}'$ and $\mathbf{V}''$, respectively, and a list:

$$\Sigma : \; \sigma_1, \sigma_2, \, \dots \, , \sigma_r$$

of positive real numbers such that:

$$(\mathrm{SV}) \qquad \begin{array}{ll} L(C'_1) = \sigma_1 C''_1 & L^*(C''_1) = \sigma_1 C'_1 \\ L(C'_2) = \sigma_2 C''_2 & L^*(C''_2) = \sigma_2 C'_2 \\ \vdots & \vdots \\ L(C'_r) = \sigma_r C''_r & L^*(C''_r) = \sigma_r C'_r \\ & \\ L(C'_s) = 0 & L^*(C''_s) = 0 \\ \vdots & \vdots \\ L(C'_p) = 0 & L^*(C''_q) = 0 \end{array}$$

where $s = r + 1$. This foregoing contention presents:

THE SINGULAR VALUE DECOMPOSITION

The entries in the list Σ are the *singular values* for L.

02° In the last section, we will describe and apply the Singular Value Decomposition in terms of matrices.

03° As a preamble to the proof, we observe a very neat relation between $ker(L)$ and $ran(L^*)$. For each X in $\mathbf{V}'$:

$$\begin{aligned} X \in ker(L) &\Longleftrightarrow L(X) = 0 \\ &\Longleftrightarrow (\forall Y \in \mathbf{V}'') \langle\!\langle \, L(X), Y \, \rangle\!\rangle = 0 \\ &\Longleftrightarrow (\forall Y \in \mathbf{V}') \langle\!\langle \, X, L^*(Y) \, \rangle\!\rangle = 0 \\ &\Longleftrightarrow X \in ran(L^*)^{\perp} \end{aligned}$$

Hence:

$$ker(L) = ran(L^*)^{\perp}$$

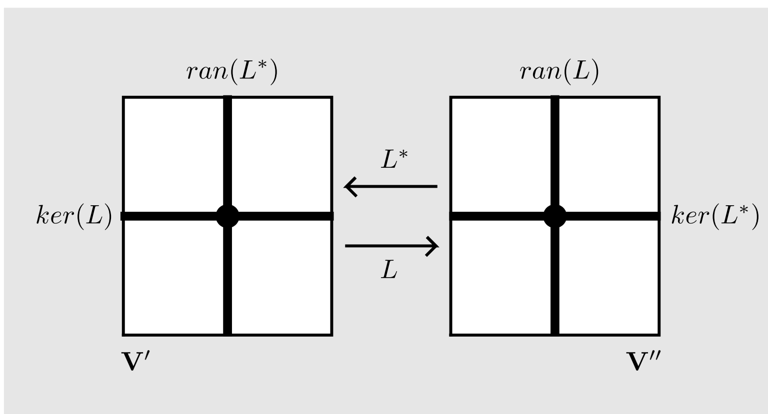
Similarly:

$$ker(L^*) = ran(L)^{\perp}$$

Consequently, we obtain the following orthogonal direct sum decompositions of $\mathbf{V}'$ and $\mathbf{V}''$:

$$\mathbf{V}' = ker(L) \perp ran(L^*), \qquad \mathbf{V}'' = ker(L^*) \perp ran(L)$$

We infer that L carries $ran(L^*)$ bijectively to $ran(L)$ and L^* carries $ran(L)$ bijectively to $ran(L^*)$. Hence, the dimensions of $ran(L)$ and $ran(L^*)$ are the same, namely, r.



SVD Format

04° Now we introduce the linear mapping S in $\mathbf{L}(\mathbf{V}')$, as follows:

$$S = L^* L,$$

Obviously, S carries $ran(L^*)$ bijectively to itself, while S carries $ker(L)$ to $\{0\}$. We find that, for any X and Y in $\mathbf{V}'$:

$$\begin{aligned} \langle\!\langle\, S(X), Y \,\rangle\!\rangle &= \langle\!\langle\, L^*(L(X)), Y \,\rangle\!\rangle \\ &= \langle\!\langle\, L(X), L(Y) \,\rangle\!\rangle \\ &= \langle\!\langle\, X, L^*(L(Y)) \,\rangle\!\rangle \\ &= \langle\!\langle\, X, S(Y) \,\rangle\!\rangle \end{aligned}$$

Hence, S is self adjoint. In fact, S is *nonnegative*. That is, for any X in $\mathbf{V}'$:

$$0 \leq \langle\!\langle\, S(X), X \,\rangle\!\rangle$$

because:

$$0 \leq \langle\!\langle\, L(X), L(X \,\rangle\!\rangle = \langle\!\langle\, L^*(L(X)), X \,\rangle\!\rangle = \langle\!\langle\, S(X), X \,\rangle\!\rangle$$

It follows that the characteristic values for S are nonnegative real numbers:

$$0 \leq \langle\!\langle\, S(X), X \,\rangle\!\rangle = \langle\!\langle\, aX, X \,\rangle\!\rangle = a\langle\!\langle\, X, X \,\rangle\!\rangle \qquad (X \neq 0)$$

05° By the Spectral Theorem, we may introduce an orthonormal basis $\bar{\mathcal{C}}'$ for $ran(L^*)$:

$$\bar{\mathcal{C}}' : \;\; C'_1, C'_2, \; \dots \; , C'_r$$

and a list $\mathcal{T}$ of positive real numbers:

$$\mathcal{T} : \;\; \tau_1, \tau_2, \; \dots \; , \tau_r$$

such that, for each index j $(1 \leq j \leq r)$:

$$S(C'_j) = \tau_j C'_j$$

06° Of course, we may extend $\bar{\mathcal{C}}'$ to an orthonormal basis for $\mathbf{V}'$:

$$\mathcal{C}' : \;\; \bar{\mathcal{C}}', \, C'_{r+1} \; \dots \; , C'_p$$

The latter members in the list form an orthonormal basis for $ker(L)$.

07° In turn, for any indices k and ℓ $(1 \leq k \leq r, \; 1 \leq \ell \leq r)$:

$$(*) \qquad \langle\!\langle\, L(C'_k), L(C'_\ell) \,\rangle\!\rangle = \langle\!\langle\, S(C'_k), C'_\ell) \,\rangle\!\rangle = \begin{cases} 0 & \text{if } k \neq \ell \\ \tau_j & \text{if } k = \ell = j \end{cases}$$

We are led to introduce the list Σ of (positive) square roots:

$$\Sigma : \;\; \sigma_1 = \sqrt{\tau_1}, \, \sigma_2 = \sqrt{\tau_2}, \; \dots \; , \, \sigma_r = \sqrt{\tau_r}$$

Now we can introduce an orthonormal basis $\bar{\mathcal{C}}''$ for $ran(L)$, as follows:

$$\bar{\mathcal{C}}'' : \;\; C''_1 = \frac{1}{\sigma_1} L(C'_1), C''_2 = \frac{1}{\sigma_2} L(C'_2), \; \dots \; , C''_r = \frac{1}{\sigma_r} L(C'_r)$$

For each index j $(1 \leq j \leq r)$, we observe that:

$$L(C'_j) = \sigma_j C''_j$$

Moreover:

$$L^*(C''_j) = \sigma_l C'_j$$

because:

$$L^*(\frac{1}{\sigma_j}L(C'_j)) = \frac{1}{\sigma_j}L^*(L(C'_j)) = \frac{1}{\sigma_j}S(C'_j) = \frac{1}{\sigma_j}\tau_j C'_j$$

08° Of course, we may extend $\bar{\mathcal{C}}''$ to an orthonormal basis for $\mathbf{V}''$:

$$\mathcal{C}'' : \quad \bar{\mathcal{C}}'', C''_{r+1} \ \dots \ , C''_q$$

The latter members in the list form an orthonormal basis for $ker(L^*)$.

09° One should note that the members C'_1, C'_2, …, and C'_r of $\mathbf{V}'$ and the members C''_1, C''_2, …, and C''_r of $\mathbf{V}''$ are subject to meaningful restrictions while the members C'_{r+1}, …, and C'_p of $\mathbf{V}'$ and the members C''_{r+1}, …, and C''_q of $\mathbf{V}''$ are not.

10° The proof of the theorem is complete.

11° For a first impression of the utility of the Singular Value Decomposition (SVD), let us consider the following computation. Let us reprise the foregoing context. We imagine the following Problem:

$$L(X) = Y$$

where X is a member of $\mathbf{V}'$ and Y is a member of $\mathbf{V}''$. Given Y, we search for X. Of course, Y need not be in $ran(L)$, so the search requires flexibity. Precisely, we search for a member $\hat{X}$ of $ran(L^*)$ such that, among all such members, the error:

$$\|L(\hat{X}) - Y\|^2$$

is least. The solution stands as follows:

$$\hat{X} = (S^{-1}L^*)(Y)$$

because $\hat{Y} = L(\hat{X})$ is the orthogonal projection of Y on $ran(L)$:

$$(LS^{-1}L^*)^* = LS^{-1}L^*, \quad (LS^{-1}L^*)(LS^{-1}L^*) = LS^{-1}L^*$$

We have interpreted S^{-1} to be the inverse of S, regarded as a (self adjoint) linear isomorphism carrying $ran(L^*)$ to itself.

12° In context of matrices, the foregoing procedure supports many important applications, notably, in Mathematical Statistics. In the last section, we will describe the procedure in detail, in terms of coordinates.

12 Matrices Redux

01° Let $\mathbf{V}$ be a pdo geometry, having dimension n. We begin this (our last) section by noting that, for our practical purposes, $\mathbf{V}$ and $\mathbf{R}^n$ are indistinguishable. We mean to say that there is a linear isomorphism K carrying R^n to $\mathbf{V}$ such that, for any members X and Y of $\mathbf{R}^n$:

$$\langle\!\langle\, K(X), K(Y)\,\rangle\!\rangle = \langle\!\langle\, X, Y\,\rangle\!\rangle$$

In the foregoing relation, we have invoked one common symbol to represent the two bilinear forms, one for $\mathbf{R}^n$ and one for $\mathbf{V}$. In any case, it is plain that, whatever we can prove for $\mathbf{R}^n$, we can prove for $\mathbf{V}$, and conversely.

02° To design such a linear isomorphism, we need only introduce the standard basis $\mathcal{E}$ for $\mathbf{R}^n$:

$$\mathcal{E}: \quad E_1, E_2, \ \ldots\ , E_n$$

together with an orthonormal basis $\mathcal{C}$ for $\mathbf{V}$:

$$\mathcal{C}: \quad C_1, C_2, \ \ldots\ , C_n$$

Now we may characterize K as follows:

$$K(E_1) = C_1,\ K(E_2) = C_2,\ \ \ldots\ ,\ K(E_n) = C_n$$

03° Let us commit our exposition, now, to computation. From here forward we will consider nothing other than the specific pdo geometries:

$$\mathbf{R}^n$$

together with the various linear mappings M in $\mathbf{L}(\mathbf{R}^p, \mathbf{R}^q)$, called matrices:

$$M : \mathbf{R}^p \longrightarrow \mathbf{R}^q$$

and the corresponding rectangular arrays M having q rows and p columns:

$$M = \begin{pmatrix} m_{11} & m_{12} & \cdots & m_{1p} \\ m_{21} & m_{22} & \cdots & m_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ m_{q1} & m_{q2} & \cdots & m_{qp} \end{pmatrix}$$

Let us recall that:

$$m_{kj} = \langle\!\langle\, ME_j, E_k\,\rangle\!\rangle$$

where j and k are any indices $(1 \le j \le p,\ 1 \le k \le q)$. Of course:

$$E_1, E_2, \ \ldots\ , E_p\,; \quad E_1, E_2, \ \ldots\ , E_q$$

are the standard bases in $\mathbf{R}^p$ and $\mathbf{R}^q$, respectively. Mindful of the relations just stated, we will simply identify mapping and array.

04° Of course, we will make use of the operations on arrays, reflecting the operations on linear mappings:

$$M' + M'',\ cM,\ M''M'$$

05° We will also make use of the operation of *adjunction*:

$$M^*$$

As one should anticipate, this operation reflects the relation between the linear mappings M and M^* in $\mathbf{L}(\mathbf{R}^p, \mathbf{R}^q)$ and $\mathbf{L}(\mathbf{R}^q, \mathbf{R}^p)$, respectively. For instance, if $p = 5$ and $q = 3$:

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{15} \\ m_{21} & m_{22} & m_{23} & m_{24} & m_{25} \\ m_{31} & m_{32} & m_{33} & m_{34} & m_{35} \end{pmatrix}, \quad M^* = \begin{pmatrix} m_{11} & m_{21} & m_{31} \\ m_{12} & m_{22} & m_{32} \\ m_{13} & m_{23} & m_{33} \\ m_{14} & m_{24} & m_{34} \\ m_{15} & m_{25} & m_{35} \end{pmatrix}$$

06° Let n be a positive integer. We must describe the relation between the standard basis:

$$\mathcal{E}: \quad E_1, E_2, \ \ldots\ , E_n$$

and the various orthonormal bases for $\mathbf{R}^n$:

$$\mathcal{C}: \quad C_1, C_2, \ \ldots\ , C_n$$

To that end, let U be the linear isomorphism in $\mathbf{L}(\mathbf{R}^n)$, determined as follows:

$$UE_1 = C_1, UE_2 = C_2, \ \ldots\ , UE_n = C_n$$

Such a mapping is a matrix for which the columns form an orthonormal basis for $\mathbf{R}^n$. We will refer to U as an *orthogonal* matrix. One may characterize it very neatly, by the condition:

$$U^*U = I$$

07° Obviously, U is invertible. It follows that:

$$U^* = U^*UU^{-1} = U^{-1}$$

Hence, $UU^* = I$. Moreover, since $det(U^*) = det(U)$, we find that:

$$det(U) = \pm 1$$

08° Now let us present the Spectral Theorem and the Singular Value Decomposition, in the current context of cartesian spaces. Let n be a positive integer. Let S be a self adjoint matrix in $\mathbf{L}(\mathbf{R}^n)$. We may introduce an orthonormal basis $\mathcal{C}$ for $\mathbf{R}^n$:

$$\mathcal{C}: \quad C_1, C_2, \ \ldots \ , C_n$$

and a list $\mathcal{T}$ of real numbers:

$$\mathcal{T}: \quad t_1, t_2, \ \ldots \ , t_n$$

such that, for each index k $(1 \leq k \leq n)$:

$$SC_k = t_kC_k$$

Let U be the corresponding change of basis matrix. In turn, let $\bar{S} = U^*SU$. We find that, for each index k $(1 \leq k \leq n)$:

$$\bar{S}E_k = t_kE_k$$

because:

$$U^*SUE_k = U^*S\,C_k = U^*t_kC_k = t_kU^*C_k$$

We may say that U converts S to *diagonal form*.

09° Here is a simple example. We set $n = 3$ and we set S as follows:

$$S = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

By tinkering, we find the list $\mathcal{T}$:

$$\mathcal{T}: \quad -2, 0, 4$$

and we find the corresponding orthonormal basis $\mathcal{C}$:

$$\mathcal{C}: \quad \frac{1}{\sqrt{6}}\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}, \ \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \ \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

In this way, we obtain:

$$\bar{S} = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

10° Let us look back to **Section 10**. In that context, we stated a basic fact, then applied that fact to prove the Spectral Theorem. Let us complete the proof of the theorem by proving the fact. Let n be a positive integer. Let S be a self adjoint linear mapping carrying $\mathbf{R}^n$ to itself. Relative to the standard basis for $\mathbf{R}^n$, the matrix S would stand as follows:

$$S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1n} \\ s_{21} & s_{22} & \cdots & s_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{nn} \end{pmatrix} \qquad (s_{jk} = s_{kj})$$

We contend that there are a real number t and a member C of $\mathbf{R}^n$ for which:

$$\langle\!\langle\, C, C \,\rangle\!\rangle = 1 \quad \text{and} \quad S(C) = t\,C$$

To prove the contention, we introduce a particular Constrained Extremum Problem:

$$f(X) = \langle\!\langle\, S(X), X \,\rangle\!\rangle, \qquad g(X) = \langle\!\langle\, X, X \,\rangle\!\rangle = 1$$

where X runs through $\mathbf{R}^n$:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

We find that:

$$\begin{aligned} f(x_1, x_2, \ \ldots\ , x_n) &= \sum_{j=1}^{n}\sum_{k=1}^{n} x_j s_{jk} x_k \\ g(x_1, x_2, \ \ldots\ , x_n) &= x_1^2 + x_2^2 + \ \cdots\ + x_n^2 \end{aligned}$$

11° The constraint:

$$g(X) = 1$$

defines a compact subset of $\mathbf{R}^n$, namely, the unit sphere $\mathbf{S}^n$. The Extreme Value Theorem guarantees that there is a member C of $\mathbf{S}^n$ such that the value $f(C)$ is the largest among all values:

$$f(X)$$

where X is any member of $\mathbf{S}^n$. In turn, the Theorem of Lagrange guarantees that there is a real number t such that:

$$(\nabla f)(C) = t\,(\nabla g)(C)$$

By straightforward computation, we find that:

$$(\nabla f)(C) = 2S(C) \quad \text{and} \quad (\nabla g)(C) = 2C$$

We conclude that:

$$\langle\!\langle\, C, C \,\rangle\!\rangle = 1 \qquad S(C) = t\,C$$

The proof is complete.

12° Now let us describe the Singular Value Decomposition in terms of matrices. We return to the context of article 01° in **Section 11**. Let us identify $\mathbf{V}'$ with $\mathbf{R}^p$ and $\mathbf{V}''$ with $\mathbf{R}^q$. Let U and V be the change of basis matrices corresponding to $\mathcal{C}'$ and $\mathcal{C}''$, respectively:

$$UE'_j = C'_j \quad (1 \le j \le p), \qquad VE''_k = C''_k \quad (1 \le k \le q)$$

We have employed the markers $'$ and $''$ to distinguish the standard bases in $\mathbf{R}^p$ and $\mathbf{R}^q$, respectively. In turn, let $\bar{L} = V^* L U$. We find that, for each index ℓ $(1 \le \ell \le r)$:

$$\bar{L}E'_\ell = \sigma_\ell E''_\ell$$

because:

$$V^* L U E'_\ell = V^* L\, C'_\ell = V^* \sigma_\ell C''_\ell = \sigma_\ell V^* C''_\ell$$

The remaining columns of $\bar{L}$ equal 0. We may say that U and V convert L to *pseudo diagonal* form.

13° For $p = 4$, $q = 7$, and $r = 3$, we have:

$$\bar{L} = \begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Granted this simple context, let us return to article 11° in **Section 11**. We begin with a member Y of $\mathbf{R}^7$:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix}$$

We search for a member X of $\mathbf{R}^4$:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

such that, among all such members, the error:

$$\|LX - Y\|^2$$

is least. In fact, we need only consider members X of $ran(L^*)$, discounting the members of $ker(L)$.

14° Let us recall that $L = V\bar{L}U^*$ and let us introduce $\bar{X} = U^*X$ and $\bar{Y} = V^*Y$. We find that:

$$\|LX - Y\|^2 = \|\bar{L}\bar{X} - \bar{Y}\|^2$$

because:

$$\|V\bar{L}U^*X - Y\|^2 = \|\bar{L}U^*X - V^*Y\|^2$$

By comparing:

$$\begin{pmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} \quad \text{with} \quad \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \\ \bar{y}_4 \\ \bar{y}_5 \\ \bar{y}_6 \\ \bar{y}_7 \end{pmatrix}$$

we find the best choice of $\bar{X}$:

$$\begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \\ \bar{x}_4 \end{pmatrix} = \begin{pmatrix} \bar{\sigma}_1\bar{y}_1 \\ \bar{\sigma}_2\bar{y}_2 \\ \bar{\sigma}_3\bar{y}_3 \\ 0 \end{pmatrix}$$

where $\bar\sigma_1 = \sigma_1^{-1}$, $\bar\sigma_2 = \sigma_2^{-1}$, and $\bar\sigma_3 = \sigma_3^{-1}$. It is the same to say that:

$$\bar X = \bar S^{-1}\bar L^*\bar Y$$

because:

$$\begin{pmatrix}\bar x_1\\ \bar x_2\\ \bar x_3\\ \bar x_4\end{pmatrix} = \begin{pmatrix}\bar\tau_1 & 0 & 0 & 0\\ 0 & \bar\tau_2 & 0 & 0\\ 0 & 0 & \bar\tau_3 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\sigma_1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & \sigma_2 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & \sigma_3 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\bar y_1\\ \bar y_2\\ \bar y_3\\ \bar y_4\\ \bar y_5\\ \bar y_6\\ \bar y_7\end{pmatrix}$$

$$= \begin{pmatrix}\bar\sigma_1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & \bar\sigma_2 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & \bar\sigma_3 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}\bar y_1\\ \bar y_2\\ \bar y_3\\ \bar y_4\\ \bar y_5\\ \bar y_6\\ \bar y_7\end{pmatrix}$$

where $\bar\tau_1 = \tau_1^{-1}$, $\bar\tau_2 = \tau_2^{-1}$, and $\bar\tau_3 = \tau_3^{-1}$. One should note that $\bar L\,\bar S^{-1}\bar L^*$ is the projection on the range of $\bar L$:

$$\begin{pmatrix}* & 0 & 0 & 0\\ 0 & * & 0 & 0\\ 0 & 0 & * & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}* & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & * & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & * & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\end{pmatrix} = \begin{pmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\end{pmatrix}$$

15° Recalling the foregoing relations and definitions, we recover the result put forward in article 11° in **Section 11**:

$$X = S^{-1}L^*Y$$

because:

$$X = U\bar X = U\bar S^{-1}U^*U\bar L^*V^*V\bar Y = S^{-1}L^*Y$$

16° Here is an example which motivates application of the Singular Value Decomposition. We imagine three variables:

$$X',\ X'',\ Y$$

We set the values of X' and X'' prior to observation, then proceed to observe the value of Y. Following eight repetitions, we obtain the following matrix:

$$M = \begin{pmatrix} x'_1 & x''_1 & y_1 \\ x'_2 & x''_2 & y_2 \\ x'_3 & x''_3 & y_3 \\ x'_4 & x''_4 & y_4 \\ x'_5 & x''_5 & y_5 \\ x'_6 & x''_6 & y_6 \\ x'_7 & x''_7 & y_7 \\ x'_8 & x''_8 & y_8 \end{pmatrix}$$

We assume that the variables satisfy a linear relation

$$Y = a + b'X' + b''X'' + c'X'^2 + c^\circ X'X'' + c''X''^2$$

We require to find the "best estimate" of the coefficients:

$$W = \begin{pmatrix} a \\ b' \\ b'' \\ c' \\ c^\circ \\ c'' \end{pmatrix}$$

consistent with the observations:

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{pmatrix}$$

To that end, we form the matrix :

$$T = \begin{pmatrix} 1 & x'_1 & x''_1 & {x'_1}^2 & x'_1x''_1 & {x''_1}^2 \\ 1 & x'_2 & x''_2 & {x'_2}^2 & x'_2x''_2 & {x''_2}^2 \\ 1 & x'_3 & x''_3 & {x'_3}^2 & x'_3x''_3 & {x''_3}^2 \\ 1 & x'_4 & x''_4 & {x'_4}^2 & x'_4x''_4 & {x''_4}^2 \\ 1 & x'_5 & x''_5 & {x'_5}^2 & x'_5x''_5 & {x''_5}^2 \\ 1 & x'_6 & x''_6 & {x'_6}^2 & x'_6x''_6 & {x''_6}^2 \\ 1 & x'_7 & x''_7 & {x'_7}^2 & x'_7x''_7 & {x''_7}^2 \\ 1 & x'_8 & x''_8 & {x'_8}^2 & x'_8x''_8 & {x''_8}^2 \end{pmatrix},$$

and we seek to minimize the value:

$$\|TW - Y\|^2$$

where:

$$TW = \begin{pmatrix} a + b'x'_1 + b''x''_1 + c'{x'_1}^2 + c^\circ x'_1x''_1 + c''{x''_1}^2 \\ a + b'x'_2 + b''x''_2 + c'{x'_2}^2 + c^\circ x'_2x''_2 + c''{x''_2}^2 \\ a + b'x'_3 + b''x''_3 + c'{x'_3}^2 + c^\circ x'_3x''_3 + c''{x''_3}^2 \\ a + b'x'_4 + b''x''_4 + c'{x'_4}^2 + c^\circ x'_4x''_4 + c''{x''_4}^2 \\ a + b'x'_5 + b''x''_5 + c'{x'_5}^2 + c^\circ x'_5x''_5 + c''{x''_5}^2 \\ a + b'x'_6 + b''x''_6 + c'{x'_6}^2 + c^\circ x'_6x''_6 + c''{x''_6}^2 \\ a + b'x'_7 + b''x''_7 + c'{x'_7}^2 + c^\circ x'_7x''_7 + c''{x''_7}^2 \\ a + b'x'_8 + b''x''_8 + c'{x'_8}^2 + c^\circ x'_8x''_8 + c''{x''_8}^2 \end{pmatrix}$$

The Singular Value Decomposition yields the answer:

$$W = S^{-1}T^*Y \qquad (S = T^*T)$$

For completeness, let us recall that:

$$TS^{-1}T^*$$

is the projection carrying $\mathbf{V}''$ to the range of T.