I. J. SCHOENBERG
ANNE WHITNEY

A theorem on polygons in n dimensions with applications to variation-diminishing and cyclic variation-diminishing linear transformations

Compositio Mathematica, tome 9 (1951), p. 141-160
<http://www.numdam.org/item?id=CM_1951__9__141_0>

A theorem on polygons in n dimensions with applications to variation-diminishing and cyclic variation-diminishing linear transformations

by

I. J. Schoenberg and Anne Whitney

Philadelphia, Pa

Introduction and statement of results.

1. Let

$$(1)\qquad \begin{array}{l} y_1 \;= a_{11}x_1 + \ldots + a_{1n}x_n \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y_m = a_{m1}x_1 + \ldots + a_{mn}x_n \end{array}$$

be a general real linear transformation which, in terms of its matrix A, may also be written as $(y) = A(x)$. Throughout this paper we denote the rank of A by r. Let $v(y)$ denote the number of variations of sign in the sequence of the numbers $y_1, y_2, \ldots, y_m$. If we pick a set of r linearly independent y's, and call them $y_1', y_2', \ldots, y_r'$, it is clear that we may assign to them arbitrary values alternating in sign; for such a set of x's and corresponding y's we now have that $v(y) \geqq v(y') = r - 1$. This proves the inequality

$$(2)\qquad \sup_{(x)} v(y) \geqq r - 1.$$

We now inquire as to when the equality sign holds in (2). In other words: *When do we always have the inequality*

$$(3)\qquad v(y) \leqq r - 1,$$

for arbitrary x's?

The answer is conveniently described in terms of the following concepts. We say that a real matrix is *definite* if it has no two elements of opposite signs. We say that a matrix has *definite columns*, if each one of its columns is a definite one-column matrix. Let r be the rank of A; if $1 \leqq i \leqq r$, we denote by $A^{(i)}$ the matrix, of $\binom{m}{i}$ rows and $\binom{n}{i}$ columns, whose elements are the ith order minors of A, where all minors formed from the same set of i rows of A appear in the same row of $A^{(i)}$ and the same rule holds regarding columns. It is well known that $A^{(r)}$ has rank unity so that its columns are proportional. The answer to the question raised above may now be described as follows:

THEOREM 1. *The inequality* (3) *always holds if and only if the matrix* $A^{(r)}$ *has definite columns.*

A proof is given in § 1.

2. As mentioned in the title, our result may be interpreted geometrically. Let us interpret the rows of A as points $P_i = (a_{i1}, \ldots, a_{in})$, $(i = 1, \ldots, m)$, in the space E_n, the vectors $\overrightarrow{OP_i}$ spanning an r - flat. Now $v(y)$ represents the number of times the hyperplane $a_1x_1 + \ldots + a_nx_n = 0$ (of coefficients $x_1, \ldots, x_n$) is crossed by the polygonal line $\Pi = P_1P_2 \ldots P_m$. The inequality (2) shows that appropriate hyperplanes through the origin are crossed by Π at least $r - 1$ times, while (3) requires that there be no such hyperplanes which are crossed by Π more than $r - 1$ times. By Theorem 1 we have a configuration satisfying (3) if and only if the matrix $A^{(r)}$ has definite columns.

Let us apply our result to the following related situation. Let $Q_i = (b_{i1}, \ldots, b_{in})$, $(i = 1, \ldots, m)$, be m points in E_n and let r be the dimension of the least flat manifold containing them; the dimension r is determined by the rank of the matrix

$$(4) \qquad B = \left\| \begin{matrix} 1 & b_{i1} & \ldots & b_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ 1 & b_{m1} & \ldots & b_{mn} \end{matrix} \right\|,$$

which must be $r + 1$. As before we find that appropriate (unrestricted) hyperplanes will be crossed by the polygonal line $\Pi^* = Q_1Q_2 \ldots Q_m$ at least r times, since for appropriate values of the variables (x) the m linear forms

$$y_i = x_o + b_{i1}x_1 + \ldots + b_{in}x_n, \ (i = 1, \ldots, m),$$

will show a number of variations of signs $v(y) \geqq r$. Applying Theorem 1, we may state the following.

COROLLARY. *Let the polygonal line* $\Pi^* = Q_1Q_2 \ldots Q_m$ *in* E_n *be such that the matrix* (4) *is of rank* $r + 1$. *Then* Π^* *crosses no hyperplane of* E_n *more than* r *times if and only if the matrix* $B^{(r+1)}$ *has definite columns. In the particular case when* $r = n$, *the requirement reduces to the condition that all non-singular minors of* B, *of order* $n + 1$, *be of the same sign.*

Such polygonal lines Π^* are said to be of order r. If $r = 2$, Π^* is an ordinary plane convex polygonal line. This geometric interpretation of Theorem 1 will not be referred to in the sequel; it was mentioned explicitly because of its possible usefulness in the theory of arcs of order r in the sense of Juel, Haupt, and Scherk (See [4]).

3. Let us return to the transformation (1) and let us denote by $v(x)$, and $v(y)$, the number of variations of sign in the sequence of the variables (x), and (y), respectively. In a number of analytical investigations those transformations (1) are of importance which have the property that we always have the inequality

$$v(y) \leqq v(x), \tag{5}$$

for arbitrary values of the x's. Special instances of such transformations were found by Laguerre, A. Hurwitz, Polya and Fekete (See [2]); Polya called them *variation-diminishing transformations*. The problem of characterizing such transformations was attacked by Schoenberg in 1930 with only partial success (See [5]); it was solved in 1933 by Th. Motzkin (See [3]) as follows.

THEOREM 2 (Motzkin). *The transformation* (1) *is variation-diminishing if and only if its matrix A, of rank r, enjoys the following two properties*:

A. *The matrices A, $A^{(2)}$, . . ., $A^{(r-1)}$ are definite.*

B. *The matrix $A^{(r)}$ has definite columns.*

In § 2 we show as a first application of Theorem 1 how Motzkin's Theorem 2 may be readily derived from it.

4. As a second application of Theorem 1 we solve in § 3 the related problem of cyclic variation-diminishing transformations. By the number $v_c(x)$ of the cyclic variations of sign of the numbers $x_1, \ldots, x_n$, we mean the following: Place the numbers $x_1, \ldots, x_n$ in order, at the vertices of a regular n-gon and let $v_c(x)$ denote the number of variations of signs counted along the circumscribed circle. Clearly $v_c(x)$ is always an even number; more exactly

$$v_c(x) = \begin{cases} v(x) & \text{if } v(x) \text{ is even,} \\ v(x)+1 & \text{if } v(x) \text{ is odd.} \end{cases} \tag{6}$$

Thus $v_c(x)$ is the least even number $\geqq v(x)$. We call the transformation (1) *cyclic variation-diminishing* if and only if we always have the inequality

$$v_c(y) \leqq v_c(x), \tag{7}$$

for arbitrary x's.

Since (5) implies (7), in view of (6), it is seen that a variation-diminishing transformation is also cyclic variation-diminishing; the converse, however, is not true. A convenient formulation of the main result requires the following.

DEFINITION 1. *Let $r = 2s$ be an even positive integer. We say that the matrix $B = ||b_{\alpha\beta}||$, of $r + 2$ rows and r columns, is a separable matrix, or an S-matrix, or $B \in S$, if and only if its $r + 2$ consecutive cyclic block-minors of order r*

$$B_1 = \begin{vmatrix} b_{1\beta} \\ b_{2\beta} \\ \vdots \\ b_{r\beta} \end{vmatrix}, \quad B_2 = \begin{vmatrix} b_{2\beta} \\ b_{3\beta} \\ \vdots \\ b_{r+1,\beta} \end{vmatrix}, \ldots, B_{r+1} = \begin{vmatrix} b_{r+1,\beta} \\ b_{r+2,\beta} \\ b_{1\beta} \\ \vdots \\ b_{r-2,\beta} \end{vmatrix}, \quad B_{r+2} = \begin{vmatrix} b_{r+2,\beta} \\ b_{1\beta} \\ \vdots \\ b_{r-1,\beta} \end{vmatrix},$$

alternate strictly in sign.

Cyclic variation-diminishing transformations are described by the following theorem.

THEOREM 3. *The transformation* (1) *is cyclic variation-diminishing, i. e. the inequality* (7) *always holds, if and only if its matrix A, of rank r, enjoys the following three properties*:

a. *For every odd i, $1 \leqq i < r$, the matrix $A^{(i)}$ is definite.*

b. *If r is odd, then $A^{(r)}$ has definite columns.*

c. *If r is even, then A should have no submatrix B, of $r + 2$ rows and r columns, which is an S-matrix* (See Definition 1).

The conditions of Theorem 3 are better understood if we realize that they must be invariant with respect to cyclic permutations of rows, or of columns, performed on A. This remark explains the role played by the parity of the orders of minors of A: Notice the obvious fact that an odd order determinant never changes sign if we permute its rows cyclically and that this is no longer true for even orders.

The conditions of Theorem 3 simplify considerably in the special case when $m = r$; then only condition (a) remains to be required since conditions (b) and (c) are vacuously satisfied.

A final remark concerns the purpose of a study of cyclic variation-diminishing transformations. One of the authors has studied the convolution transformation

$$(8) \qquad g(x) = \int_{\infty}^{\infty} \Lambda(x - t) f(t) dt,$$

where $\Lambda(x)$ is a given summable function, and has found the necessary and sufficient conditions which $\Lambda(x)$ is to satisfy in order that the transformation (8) should be variation-diminishing. By this we mean that for every bounded $f(t)$ the transformed function $g(x)$ should have no more variations in sign than the original function $f(t)$ (See [6]). Motzkin's Theorem 2 was one of the tools used in the solution of this problem. Let now $\Phi(x)$ be a given function of periode 2π which is summable over a period. If $f(t)$ is a bounded function of period 2π then

$$g(x) = \int_0^{2\pi} \Phi(x - t) f(t) dt \tag{9}$$

is a continuous function of period 2π. The following periodic analogue of the previous problem was proposed orally by H. Cartan: For which functions $\Phi(x)$ will (9) be cyclic variation-diminishing in the sense that we shall always have the inequality

$$v_c(g) \leqq v_c(f)?$$

For this problem, to which we hope to return at a future date, Theorem 3 should prove to be an important tool.

§ 1. A proof of Theorem 1.

5. PROOF OF NECESSITY. We assume that the system $(y) = A(x)$ always implies that

$$v(y) \leqq r - 1 \tag{1.1}$$

and we are to show that the columns of the matrix $A^{(r)}$ are definite. We first dispose of the special case when

$$r = n = m - 1,$$

in which case our transformation becomes

$$\begin{aligned} y_1 &= a_{11} x_1 + \dots + a_{1n} x_n \\ &\dots\dots\dots\dots\dots\dots \\ y_{n+1} &= a_{n+1,\,1}\, x_1 + \dots + a_{n+1,\,n}\, x_n, \end{aligned}$$

with a matrix of rank n. Since we always have that $v(y) \leqq n - 1$, we see that $v(y) = n$ is impossible so that the y's will never strictly alternate in sign. However, with an obvious notation for minors, the relation

$$\sum_1^{n+1} (-1)^{\nu-1} y_\nu A_\nu = 0 \tag{1.2}$$

is the only constraint among the y's. Hence no two among the minors of order n $A_1, A_2, \dots, A_{n+1}$ can be of opposite signs, for otherwise we could clearly satisfy the relation (1.2) with values such that $(-1)^{\nu-1}y_\nu > 0$, for all ν, which we know to be impossible.

We pass now to the general case. Take a set of r linearly independent columns of A, say $C_1, C_2, \ldots C_r$. Setting

$$x_{r+1} = \ldots = x_n = 0,$$

we obtain the system

$$\begin{aligned} y_1 &= a_{11}x_1 + \ldots + a_{1r}x_r \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ y_m &= a_{m1}x_1 + \ldots + a_{mr}x_r, \end{aligned}$$

which inherits the property that $v(y) \leqq r - 1$, for arbitrary values of $x_1, \ldots, x_r$. We are to show that all minors of order r, of its matrix, are of the same sign. More precisely: All the non-singular among them are positive, or all are negative. Let us denote by A_r a generic non-singular minor of order r and let A_r' and A_r'' be any two such distinct minors. By a known theorem (See [3], page 64) these two minors may be made to be the first and last of a finite sequence of A_r''s such that two adjacent members of the sequence have precisely $r - 1$ rows in common. Let $A_r^{(1)}$, $A_r^{(2)}{}_r$ be adjacent members of the sequence. The system formed only of those $r + 1$ equations corresponding to rows of $A_r^{(1)}$ or $A_r^{(2)}$ also enjoys the property (1.1); by the previous case already settled we conclude that $A_r^{(1)} \,.\, A_r^{(2)} > 0$. Applying this to all pairs of adjacent elements of the sequence, we conclude that $A_r' \,.\, A_r'' > 0$ and the necessity proof is completed.

6. Proof of sufficiency. Let us assume that

(1.3) $\qquad A^{(r)}$ has definite columns,

and let us show that $(y) = A\ (x)$ always implies (1.1). We start by adding a number of additional assumptions without thereby losing generality.

A. *Without loss of generality we may assume that* $r = n$. Indeed, in view of the relation $(y) = A(x)$, the column of y's is a linear combination of the columns of A. Therefore the column of y's may also be expressed as a linear combination of a set of r linearly independent columns, $a_{i1}, \ldots, a_{ir}$, say. With appropriate $x_1', \ldots, x_r'$ we therefore have that

$$y_i = \sum_{k=1}^{r} a_{ik}x_k'. \tag{1.4}$$

Now (1.4) is a linear transformation with $r = n$, whose matrix has minors of order r of one sign only. Hence if our theorem is already established for this particular case, then (1.1) holds and everything is settled. *Let us therefore assume that A is of rank n and we are to prove that*

$$v(y) \leqq n - 1. \tag{1.5}$$

B. *Without loss of generality we may and do assume that none of the rows of A has only zero elements.* Indeed, any such rows, if present, may simply be struck out without impairing the assumptions on A.

C. *Without loss of generality we may and do assume that none of the y's vanish.* Indeed, if some of the y's vanish, we may so slightly alter the x's as to make those y's non-vanishing, without changing the signs of the non-vanishing y's. This operation can only increase $v(y)$ so that if (1.5) is true after the change, it certainly was true before.

7. At this point it is more convenient to invert the argument and to show that our assumptions are incompatible with the additional assumption that

$$v(y) \geqq n. \tag{1.6}$$

To this end we first prove the very simple

LEMMA 1. *The quantities $y_1, \ldots, y_m$ having fixed non-vanishing values we define the new quantities z_ν by*

$$\begin{aligned} z_1 &= \alpha_1 y_1 + \beta_1 y_2 \\ z_2 &= \qquad\quad \alpha_2 y_2 + \beta_2 y_3 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ z_{m-2} &= \qquad\qquad\qquad \alpha_{m-2} y_{m-2} + \beta_{m-2} y_{m-1} \\ z_{m-1} &= \qquad\qquad\qquad\qquad\qquad \alpha_{m-1} y_{m-1} + \beta_{m-1} y_m, \end{aligned} \tag{1.7}$$

concerning which we claim the following:

1. *If*

$$v(y) = m - 1, \tag{1.8}$$

we can always find positive coefficients α_ν, β_ν, $(\nu = 1, \ldots, m - 1)$, such that

$$v(z) = m - 2 = v(y) - 1 \tag{1.9}$$

2. *If*

$$v(y) < m - 1, \tag{1.10}$$

we can always find positive coefficients α_ν, β_ν, such that

$$v(z) \geqq v(y),\ z_\nu \neq 0 \text{ for all } \nu. \tag{1.11}$$

The first statement is clearly true, for (1.8) means that the y's alternate in sign; if we choose all $\beta_\nu = 1$ and all $\alpha_\nu = N$, then $\operatorname{sgn} z_\nu = \operatorname{sgn} y_\nu$ ($\nu = 1, \ldots, m-1$), provided N is large enough, and (1.9) follows.

The second statement is proved by induction with respect to m. Being trivially true for $m = 2$, let us assume the statement true for $m-1$ rather than m. Let us strike out the last equation (1.7). There are two cases:

α) $y_1, y_2, \ldots, y_{m-1}$ *alternate in sign.* By the first part of Lemma 1 we may choose α_ν, β_ν positive ($\nu = 1, \ldots, m-2$) such that $\operatorname{sgn} z_\nu = \operatorname{sgn} y_\nu$, ($\nu = 1, \ldots, m-2$) obtaining $v(z_1, \ldots, z_{m-2}) = v(y_1, \ldots, y_{m-1}) \dot{-} 1$. However, by (1.10) we must have $y_{m-1}\, y_m > 0$, hence also $\operatorname{sgn} z_{m-1} = \operatorname{sgn} y_{m-1} = -\operatorname{sgn} y_{m-2} = -\operatorname{sgn} z_{m-2}$. Hence in the end we have $v(z) = v(y)$ no matter how the positive α_{m-1}, β_{m-1} are chosen.

β) $y_1, y_2, \ldots, y_{m-1}$ *do not alternate in sign.* By our induction assumption we can choose $z_1, \ldots, z_{m-2}$ so that

$$v(z_1, \ldots, z_{m-2}) \geqq v(y_1, \ldots, y_{m-1}).$$

If we now consider y_m there are two possibilities:

If $y_{m-1}\, y_m < 0$, we can arrange to have z_{m-1} of either sign and we choose it so as to produce a variation of sign in the z's.

If $y_{m-1}\, y_m > 0$, we cannot help achieving the desired inequality (1.11).

8. We now return to the assumptions of Article 6 and we are to show that they are incompatible with the additional assumption (1.6). Clearly (1.6) implies that $m \geqq n + 1$. We proceed by induction in m.

If $m = n + 1$, then (1.6) implies

$$v(y) = n,$$

so that the y's alternate in sign. However, the y's again satisfy the relation (1.2), which again turns out to be impossible and for the same reason as in Article 5.

Let us assume now that $m > n + 1$ and that the impossibility of (1.6) has already been established for lesser values of m. In order to prove its impossibility for m, we determine the positive coefficients α_ν, β_ν, in (1.7) so as to satisfy the conditions (1.9), or (1.11), of Lemma 1. In terms of the matrix

$$(1.12) \qquad M = \left\| \begin{matrix} \alpha_1 & \beta_1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & \alpha_2 & \beta_2 & \ldots & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & & \alpha_{m-2} & \beta_{m-2} & 0 \\ 0 & 0 & 0 & & 0 & \alpha_{m-1} & \beta_{m-1} \end{matrix} \right\|$$

the transformation (1.7) may be written as $(z) = M(y)$ and combining it with $(y) = A(x)$ we obtain the linear transformation

$$(1.13) \qquad (z) = MA(x).$$

Let us assume for the moment that the matrix MA, of dimensions $(m - 1)$ by n, has all the properties enjoyed by A, namely:

1. *MA is of rank n.*
2. *All non-singular minors of order n of MA have the same sign.*

If these properties are granted for the moment, we may easily complete the induction argument by operation with the transformation (1.13) instead of $(y) = A(x)$. Indeed, if (1.8) holds, then also (1.9) is true. Hence

$$v(z) = m - 2 \geqq n,$$

against our induction assumption, for (1.13) has only $m - 1$ equations, However, if (1.10) holds, then (1.11) follows, so that by (1.6) we find

$$v(z) \geqq v(y) \geqq n, \quad \text{or} \quad v(z) \geqq n,$$

again in contradiction to our induction assumption.

Finally, a proof of the properties 1 and 2 of MA will follow from the following properties of the matrix M:

(i) *All minors of M, of all orders, are $\geqq 0$.*

(ii) *The columns of every set of n columns of M are linearly independent.*

The property (i) is easily shown by induction in m. As to the second, we remark that $n < m - 1$ and that the columns of every set of $m - 1$ columns of M are linearly independent, because the corresponding determinant of order $m - 1$ is positive (it either reduces to the product of its positive diagonal terms or else it splits into the product of two minors having this property).

The properties 1 and 2 of the matrix MA are now established as follows:

1. *MA is of rank n*; indeed let

$$A\begin{pmatrix} j_1, j_2, \ldots, j_n \\ 1, 2, \ldots, n \end{pmatrix} > 0$$

be a positive minor of A formed from the rows $j_1, \ldots, j_n$. Consider in M the set of columns $j_1, \ldots, j_n$; these being independent, by (ii), we can select in M rows $i_1, \ldots, i_n$, such that

$$M\begin{pmatrix} i_1, \ldots, i_n \\ j_1, \ldots, j_n \end{pmatrix} > 0.$$

We may now identify a *positive* minor, of order n, of MA, since

$$MA\begin{pmatrix} i_1, \ldots, i_n \\ 1, \ldots, n \end{pmatrix} = M\begin{pmatrix} i_1, \ldots, i_n \\ j_1, \ldots, j_n \end{pmatrix} \cdot A\begin{pmatrix} j_1, \ldots, j_n \\ 1, \ldots, n \end{pmatrix} + \text{(non-negative terms)} > 0.$$

2. *All non-singular minors of order n of MA have the same sign*, since both factors M and A have this property which is preserved by multiplication. This concludes a proof of Theorem 1.

§ 2. A proof of Theorem 2 concerning variation-diminishing transformations.

9. Proof of necessity. Let us assume that (1) $(y) = A(x)$ always implies that

$$v(y) \leqq v(x). \tag{2.1}$$

We show first that

$$A^{(i)} \quad (1 \leqq i \leqq r) \text{ has definite columns.} \tag{2.2}$$

Indeed, let $C_1, C_2, \ldots, C_i$, say, be a set of i linearly independent columns of A. Setting $x_{i+1} = \ldots = x_n = 0$ we obtain the transformation

$$y_\nu = \sum_{j=1}^{i} a_{\nu j} x_j, \quad (\nu = 1, \ldots, n), \tag{2.3}$$

of rank i, which is also variation-diminishing, so that we always have the inequalities

$$v(y) \leqq v(x) \leqq i - 1.$$

By Theorem 1 we learn that all non-singular minors of order i, of the matrix of (2.3), have the same sign; this proves (2.2).

There remains to show that if $1 \leqq i < r$, then not only are the columns of $A^{(i)}$ definite, but that $A^{(i)}$ itself is definite. For this purpose we consider the special case when

$$m = n = r, \quad i = n - 1, \tag{2.4}$$

and show that $A^{(n-1)}$ is definite. It is clear that the y's may now by chosen arbitrarily; let them have alternating values such that

$$(-1)^\nu y_\nu > 0, \quad (\nu = 1, \ldots, n).$$

Our transformation being variation-diminishing, we have $v(y) = n - 1 \leqq v(x)$, hence $v(x) = n - 1$, so that the corresponding x's also alternate in sign. On solving for the x's, we find by Cramer's rule the relations

$$
\begin{aligned}
|A|\, x_1 &= -\sum_{\nu} A_{\nu 1} (-1)^{\nu} y_{\nu} \\
|A|\, x_2 &= +\sum_{\nu} A_{\nu 2} (-1)^{\nu} y_{\nu} \\
|A|\, x_3 &= -\sum_{\nu} A_{\nu 3} (-1)^{\nu} y_{\nu} \\
&\vdots
\end{aligned}
$$

Since $(-1)^{\nu} y_{\nu} > 0$, and because we know already that all $A_{\nu 1}$, have the same sign, that all $A_{\nu 2}$ also have the same sign and so forth, the alternation in sign of the x's implies that *all* non-vanishing among the $A_{\nu i}$ have the same sign, i. e. $A^{(n-1)}$ is definite.

We return now to the general case to show that $A^{(i)}$ is definite if $1 \leqq i < r$. Let Γ, Γ' be two sets of i independent columns of A such that they have $i-1$ columns in common, differing in the last columns, and such that the combined set of $i+1$ columns $(\Gamma + \Gamma')$ are linearly independent. Let A_{i+1} be a non-singular minor of $\Gamma + \Gamma'$ of order $i+1$. By setting the x's $= 0$ which correspond to columns other than those of $\Gamma + \Gamma'$ and by ignoring the equations corresponding to rows other than those of A_{i+1}, we obtain a variation-diminishing transformation

$$(y) = A_{i+1}\,(x).$$

Here we have the special case corresponding to (2.4). We conclude that the minors of Γ and those of Γ' must have the same sign, for A_{i+1} must contain non-singular minors of order i of Γ and also such of Γ'. If the two sets Γ, Γ' do not satisfy the additional conditions assumed above, then by a theorem of Motzkin we know that they can be joined by a chain of sets of i independent columns, two adjacent members of which do satisfy those conditions (See [3], page 64). The above argument, applied consecutively to adjacent members of the chain, shows that $A^{(i)}$ is indeed definite.

10. PROOF OF SUFFICIENCY. We need the following

LEMMA 2. *If the matrix A satisfies the conditions of Theorem 2, then so does the matrix B obtained from A by striking out a column, and also the matrix A^* obtained from A by replacing two adjacent columns of A by their linear combination with positive coefficients.*

PROOF: The statement concerning B is clear, for if s is the rank of B and $i \leqq s \leqq r$, then $B^{(i)}$ is a submatrix of $A^{(i)}$. To prove the second part let us assume that

$$A = \|C_1, \ldots, C_n\|, \quad A^* = \|p_1 C_1 + p_2 C_2, \; C_3, \ldots, C_n\|,$$
$$(p_1 > 0,\; p_2 > 0),$$

and let r and r^* $(\leqq r)$ be their ranks.

1. $r^* = r$. If $i < r^* = r$, it is clear that $A^{*(i)}$ is definite, for a minor of order i of A^* is either one of A or else a positive linear combination of two such, hence all are of the same sign. Also the minors of order r, in r independent columns of A^*, are of the same sign because they are either identical with similar minors of A, or else they are the elements of a linear combination of two columns of $A^{(r)}$ which are definite and proportional.

2. $r^* < r$. In this case $A^{*(i)}$ is obviously definite if $i \leqq r^*$, since $i < r$. This completes a proof.

Let us now assume that A satisfies the conditions of Theorem 2 and let us show that the inequality (2.1) must hold for an arbitrary but fixed set of x's. Without losing generality we may assume that not all x's are zero. We now perform on the matrix A the following operations:

1. We strike out those columns of A corresponding to vanishing x's.

2. We replace two adjacent columns, corresponding to x's of the same sign, by their positive linear combination according to the equation

$$\ldots + a_{\nu j} x_j + a_{\nu, j+1} x_{j+1} + \ldots$$
$$= \ldots + \frac{a_{\nu j} x_j + a_{\nu, j+1} x_{j+1}}{x_j + x_{j+1}} \cdot (x_j + x_{j+1}) + \ldots$$

By repeating the second operation sufficiently often, we reach a stage where the x's alternate in sign. By Lemma 2, the conditions of Theorem 2 are thereby preserved; thus without loss of generality we may assume that

$$v(x) = n - 1,$$

i. e. that the x's alternate in sign to start with. Since $A^{(r)}$ has definite columns, by Theorem 1 we now conclude that

$$v(y) \leqq r - 1 \leqq n - 1 = v(x), \quad \text{or} \quad v(y) \leqq v(x).$$

§ 3. A proof of Theorem 3 concerning cyclic variation-diminishing transformations.

11. The conditions of Theorem 3 will be more conveniently handled in terms of the following definitions.

Definitions 2. *We say that A of rank r, is a Γ-matrix, or $A \in \Gamma$, if and only if the following two conditions are satisfied:*

1. *For every odd i, $i < r$, the matrix $A^{(i)}$ is definite.*
2. *If r is odd, then $A^{(r)}$ has definite columns.*

Definition 3. *We say that A is a Δ-matrix, or $A \in \Delta$, if and only if the following two conditions are satisfied*:

1. $A \in \Gamma$ (*See Definition* 2).
2. *If r is even, we require that A should have no sub-matrix B, of $r + 2$ rows and r columns, which is an S-matrix* (*See Definition* 1, *Article* 4).

We may now restate Theorem 3 as

Theorem 3′. *The transformation $(y) = A(x)$ is cyclic variation-diminishing, i. e. we always have*

$$v_c(y) \leqq v_c(x), \tag{3.1}$$

if and only if its matrix A is a Δ-matrix.

12. We need a number of auxiliary theorems. The role played by S-matrices is revealed by the following

Lemma 3. *Let r be a positive even number and let $A = ||a_{ij}||$ be an S-matrix of $r + 2$ rows and r columns. The linear transformation $(y) = A(x)$ has the property*

$$\sup_{(x)} v_c(y) = r + 2 \tag{3.2}$$

In other words: For appropriate values of $x_1, \ldots, x_r$, the quantities $y_1, \ldots, y_{r+2}$ will alternate in sign.

The statement (3.2) has a simple geometric interpretation: If we consider in the space E_r the $r + 2$ points

$$P_i = (a_{i1}, a_{i2}, \ldots, a_{ir}), \quad (i = 1, \ldots, r + 2),$$

and if we denote by Π the closed polygon $P_1P_2 \ldots P_{r+2}P_1$, then (3.2) means that some appropriate hyperplane through the origin 0, of E_r, will be crossed (strictly) by Π $r + 2$ times. Such a hyperplane will separate strictly the points, $P_1, P_3, \ldots, P_{r+1}$, from the points $P_2, P_4, \ldots, P_{r+2}$; this remark explains why A is called a „separable matrix" (Definition 1). The proof will be geometric and free use will be made of fundamental notions concerning convex polyhedral domains. We shall denote by $K(V_1, V_2, \ldots, V_s)$ the convex extension of the points $V_1, \ldots, V_s$.

Proof of Lemma 3: 1. *The assumptions of Lemma* 3 *imply that the origin* 0 *is not a point of the polyhedron*

$$K = K(P_1, -P_2, P_3, -P_4, \ldots, P_{r+1}, -P_{r+2}),$$

where $-P_1 = (-a_{i1}, \ldots, -a_{ir})$.

Indeed, suppose that $0 \in K$; then it follows by a known theorem (See [1], page 607) that 0 is already a point of the convex extension of some $r + 1$ among the $r + 2$ points spanning K. This will be shown to be impossible by considering simultanously the matrices

$$A = ||P_1, P_2, \ldots, P_{r+2}||, \quad B = ||P_1, -P_2, P_3, \ldots, -P_{r+2}||,$$

where the symbols indicate rows. Notice now the following facts:

a. *Since A is an S-matrix, also B is an S-matrix,* because all cyclic block-minors of A either agree in value, or agree with changed sign, with the corresponding cyclic block-minors of B.

Let $Q_1 = P_1, Q_2 = -P_2, \ldots, Q_{r+2} = -P_{r+2}$, so that we may write

$$B = ||Q_1, Q_2, \ldots, Q_{r+2}||.$$

There now remains to prove the following statement.

b. *No matter what group of $r + 1$ among the points Q_ν we choose, their convex extension never contains the origin.* Indeed, we get such a group by striking out one of the $r + 2$ points Q_ν. The property of B of being an S-matrix is clearly invariant by cyclic permutations of the points Q_ν and so is the property (b) which we wish to prove. We therefore lose no generality if we only prove that

$$O \mathrel{\overline{\in}} K(Q_1, Q_2, \ldots, Q_{r+1}). \tag{3.3}$$

This statement depends on the signs of the solutions of a system of r homogeneous equations in $r + 1$ unknowns:

$$O = p_1 Q_1 + p_2 Q_2 + \ldots + p_{r+1} Q_{r+1}.$$

Hence there is, up to a factor, only one solution proportional to the minors of the matrix $||Q_1, Q_2, \ldots, Q_{r+1}||$ with successively changed signs. However, since $B \in S$, we know that the two minors D_1 and D_{r+1}, obtained by striking out the *first* and *last* column respectively, are of opposite signs. Since r is even, the solutions are given by the relations

$$\frac{p_1}{D_1} = \frac{p_2}{-D_2} = \ldots = \frac{p_r}{-D_r} = \frac{p_{r+1}}{D_{r+1}},$$

which show that p_1 and p_{r+1} are indeed of opposite signs. This establishes (3.3) and concludes a proof of our statement 1.

2. *Let K_1 be the convex cone spanned by the vectors $\overrightarrow{OP}_1$, $\overrightarrow{OP}_3, \ldots, \overrightarrow{OP}_{r+1}$, and let K_2 be the convex cone spanned by $\overrightarrow{OP}_2$, $\overrightarrow{OP}_4, \ldots, \overrightarrow{OP}_{r+2}$. These two cones do not have a ray in common.*

For if they had a ray in common, they would have a common point $R \neq 0$ such that

$$
\begin{aligned}
R &= p_1 P_1 + p_3 P_3 + \ldots + p_{r+1} P_{r+1},\\
p_\mu &\geqq 0,\ p_1 + p_3 + \ldots + p_{r+1} > 0,\\
R &= p_2 P_2 + p_4 P_4 + \ldots + p_{r+2} P_{r+2},\\
p_\nu &\geqq 0,\ p_2 + p_4 + \ldots + p_{r+2} > 0.
\end{aligned}
$$

However, these relations imply by substraction, that

$$\frac{\sum\limits_{\mu \text{ odd}} p_\mu P_\mu + \sum\limits_{\nu \text{ even}} p_\nu(-P_\nu)}{\sum\limits_{i=1}^{r+2} p_i} = O,$$

a relation which contradicts our statement 1.

A proof of Lemma 3 is now readily concluded. Indeed, the cones K_1 and K_2, having no common ray, may be strictly separated by a certain hyperplane

$$a_1 x_1 + \ldots + a_r x_r = 0,$$

where $a_1, \ldots, a_r$ are the running coordinates. However, this means that for this choice of the x's, in the system $(y) = A(x)$, the y's will alternate in sign.

13. We pass to a couple of theorems which show the invariance of the Δ-property with respect to certain simple operations.

Lemma 4. *If $A \epsilon \Delta$ and B is a submatrix of A, then also $B \epsilon \Delta$.*

Proof: Let r and s be the ranks of A and B, respectively, $s \leqq r$.

If $s = r$ there is nothing to prove, for s independent columns of B remain a fortiori independent if continued into A, and then $B \epsilon \Delta$ is clear if $s = r$ is odd and equally so if it is even.

If $s < r$, there are two cases. If s is odd, the matter is again clear. *Assume s to be even.* Let $C_1', \ldots, C_s'$ be independent columns of B and $C_1, \ldots, C_s$ the same columns continued into A, where they are also independent. Finally, let C be a new column of A, so selected that $C_1, \ldots, C_s, C$ are $s + 1$ independent columns of A. The number $s + 1$ being odd, the matrix

$$B^* = ||C_1, \ldots, C_s, C|| \epsilon \Gamma,$$

provided the columns of B^* appear in the same order as in A and not in the artificial order of this defining equation. By Theorem 1, the corresponding system

$$(y) = B^*(x) \text{ has the property } v(y) \leqq (s+1) - 1 = s.$$

If we set $= 0$ the variable x, corresponding to the column C, and drop all equations not formed with rows of B, we see that also the system

$$(y) = B(x) \text{ has the property } v(y) \leqq s, \text{ or } v_c(y) \leqq s.$$

By Lemma 3 we conclude that $B \epsilon \Delta$: Indeed, the presence of an S-submatrix of B would imply that $\sup v_c(y) \geqq s + 2$, against our last statement that $v_c(y) \leqq s$, always.

LEMMA 5. *If A is a Δ-matrix, then so is a matrix A^* obtained from A by replacing two adjacent columns of A by their linear combination with positive coefficients.*

PROOF: To fix the ideas let

$$A = ||C_1, \ldots, C_n||, \; A^* = ||p_1C_1 + p_2C_2, \; C_3, \ldots, C_n||,$$
$$(p_1 > 0, \; p_2 > 0),$$

and let r and r^* be their respective ranks. We have that

$$0 \leqq r - r^* \leqq 1.$$

Comparing with the corresponding proof of Lemma 2 (Article 10) we see that the only case requiring discussion is when *r is odd and $r^* = r - 1$ is even*. Suppose then, that A^* were not a Δ-matrix, i. e. that A^* does have a submatrix, of $r^* + 2$ rows and r^* columns, which is an S-matrix. But then, by Lemma 3, we conclude that the system $(y) = A^*(x)$ has the property

$$\sup v_c(y) \geqq r^* + 2,$$

and the same conclusion remains valid for the system $(y) = A(x)$. But $v_c(y) \geqq r^* + 2 = r + 1$ implies that $v(y) \geqq r$, in contradiction to Theorem 1.

14. Our last auxiliary theorem is an analogue of Theorem 1 which we state as

THEOREM 4. *Let A be a Γ-matrix* (Definition 2, Article 11) *of rank r. The corresponding system $(y) = A(x)$ has the property*

$$v_c(y) \leqq r, \textit{ for all } x\text{'s}, \tag{3.4}$$

if and only if A is a Δ-matrix (Definition 3, Article 11).

PROOF OF NECESSITY: The necessity of the condition $A \epsilon \Delta$ is clear, for if A were not a Δ-matrix, then Lemma 3 would imply that $\sup v_c(y) \geqq r + 2$, in contradiction to the assumption (3.4).

PROOF OF SUFFICIENCY. Let us now assume that $A \epsilon \Delta$ and let us prove that (3.4) holds. This result is clear if r is odd, because then $A^{(r)}$ has definite columns so that Theorem 1 implies that $v(y) \leqq r - 1$, hence $v_c(y) \leqq r - 1$ (because $r - 1$ is even) and a fortiori (3.4). We may therefore limit ourselves exclusively to *even* values of r. The statement (3.4) being trivial if $r = 0$, we use induction and assume (3.4) to be true for all *smaller even* values of r. Assume now for the remainder of the proof that (3.4) is wrong and that, on the contrary, for a certain set of x's we have

$$v_c(y) \geqq r + 2. \tag{3.5}$$

By expressing, if necessary, all columns as linear combinations of a set of r linearly independent columns (as in Article 6, A) we see that we may assume that $n = r$ without loss of generality. Select $r + 2$ y's having alternating signs, which is possible in view of (3.5). In this way we get a system which, on changing notations, may be written as

$$\begin{aligned} y_1 &= b_{11}x_1 + \dots + b_{1r}x_r \\ &\dots\dots\dots\dots\dots\dots\dots \\ y_{r+2} &= b_{r+2,\,1}\, x_1 + \dots + b_{r+2,\,r}\, x_r, \end{aligned} \tag{3.6}$$

whose matrix B is a Δ-matrix (by Lemma 4), and such that

$$v(y) = r + 1. \tag{3.7}$$

Let r^* be the rank of B. We distinguish two cases:

1. $r^* < r$. If r^* is odd, Theorem 1 implies that

$$v(y) \leqq r^* - 1 < r - 1,$$

in contradiction to (3.7). Let r^* be even; by our induction assumption concerning (3.4), we conclude that $v_c(y) \leqq r^*$, hence $v(y) \leqq r^* < r$, again contradicting (3.7).

2. $r^* = r$. We claim:

I. *None of the* $r + 2$ *cyclic block-minors of order* r, *of* B, *can vanish.* Indeed, suppose that one of these were zero. By cyclic permutations of rows, which still preserve the property $B \in \Delta$, we may assume that

$$\begin{vmatrix} b_{11} \dots b_{1r} \\ \dots\dots\dots \\ b_{r1} \dots b_{rr} \end{vmatrix} = 0.$$

But then, for the partial system

$$\begin{aligned} y_1 &= b_{11}x_1 + \dots + b_{1r}x_r \\ &\dots\dots\dots\dots\dots\dots \\ y_r &= b_{r1}x_1 + \dots + b_{rr}x_r \end{aligned}$$

of rank r', we have $r' < r$, $v(y) = r - 1$, $v_c(y) = r$. If r' is odd we have a contradiction with Theorem 1; if r' is even, the induction assumption shows that $v_c(y) \leqq r' < r$, in contradiction to $v_c(y) = r$.

II. *The consecutive cyclic block-minors of B alternate in sign* (hence $B \in S$, in contradiction to $A \in \Delta$). Indeed, in (3.6) we must have $x_1 \neq 0$, because $r^* = r$. We now write down the $r + 2$ systems of r equations, in r unknowns, from among the equations (3.6), whose matrices are respectively the $r + 2$ cyclic block-minors of B. Solving each of these systems for x_1 by Cramer's rule, we obtain the relations

$$\begin{aligned} D_{1,2,\ldots,r} \cdot x_1 &= y_1 B'_{11} - y_2 B'_{21} + \ldots \pm y_r B'_{r1} \\ D_{2,3,\ldots,r+1} \cdot x_1 &= y_2 B''_{11} - y_3 B''_{21} + \ldots \pm y_{r+1} B''_{r1} \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ D_{r+2,1,\ldots,r-1} \cdot x_1 &= y_{r+2} B^{(r+2)}_{11} - y_1 B^{(r+2)}_{21} + \ldots \pm y_{r-1} B^{(r+2)}_{r1}, \end{aligned}$$

where $(-1)^{\nu-1} y_\nu > 0$, say, while all the minors $B^{(i)}_{j1}$, being all of the same *odd* order $r - 1$, are necessarily of the same sign (or zero). Since none of the D's vanishes, by our previous statement I, and $x_1 \neq 0$, we conclude that the D's must indeed alternate in sign. As already remarked, this result contradicts our assumption that $A \in \Delta$ and the proof is completed.

We can finally turn to a proof of Theorem 3′ (Article 11).

15. Proof that the condition of Theorem 3′ is necessary. Let us assume that the inequality (3.1) always holds and let us prove that

$$A \in \Delta \tag{3.8}$$

We prove first that

$$A \in \Gamma. \tag{3.9}$$

The proof runs along the lines of the arguments of Article 9. Let i be *odd*, $1 \leqq i \leqq r$, and let $C_1, \ldots, C_2$, say, be i linearly independent columns of A. The system (2.3) is also cyclic variation-diminishing so that

$$v_c(y) \leqq v_c(x) = v_c(x_1, \ldots, x_i).$$

However, i being odd, we have $v_c(x_1, \ldots, x_i) \leqq i - 1$, hence $v_c(y) \leqq i - 1$, and a fortiori $v(y) \leqq i - 1$. By Theorem 1 we conclude that $A^{(i)}$ has definite columns. Again i being odd, let $1 \leqq i < r$, and let us show that $A^{(i)}$ is definite. Again we consider the special case (2.4) and make $(-1)^\nu y_\nu > 0$, $(\nu = 1, \ldots, n)$. Since $n = i + 1$ is even, we find for these particular values that

$$n = v_c(y) \leqq v_c(x),$$

hence $v(x) = n$, i. e. the x's alternate in sign. From this point on the proof is identical with the similar proof of Article 9. Thus (3.9) is established. If the rank r of A is odd, then (3.9) implies (3.8).

Let r be *even* and let us show that $A \in \Delta$. By expressing all columns of A as linear combinations of a set of r linearly independent columns we know that (3.1) also implies that

$$v_c(y) \leqq r$$

always holds. Now Theorem 4 implies the desired conclusion that A is a Δ-matrix.

16. Proof that the condition of Theorem 3′ is sufficient. Let $A \in \Delta$, $(y) = A(x)$, where the x's have fixed values, not all zero, and let us prove that

$$v_c(y) \leqq v_c(x). \tag{3.10}$$

We proceed as in the sufficiency proof of Theorem 2. We perform on A the same two operations 1. and 2. described in the last two paragraphs of Article 10, with the result that we reduce the argument to the case when

$$v(x) = n - 1. \tag{3.11}$$

Indeed, by Lemma 4 and Lemma 5, the matrix of the final system is also a Δ-matrix.

On the other hand, by Theorem 4 we conclude that $v_c(y) \leqq r$ hence

$$v_c(y) \leqq n \tag{3.12}$$

If n is even, (3.11) gives $v_c(x) = n$, which together with (3.12) implies (3.10). If n is odd, then (3.12) and (3.11) imply the relations $v_c(y) < n$, $v_c(x) = n - 1$, which again imply (3.10). This concludes a proof of Theorem 3′.

This work was done under the sponsorship of the Office of Naval Research.

REFERENCES.

P. Alexandroff and H. Hopf,
[1] Topologie I, Berlin 1935.

M. Fekete and G. Polya,
[2] Über ein Problem von Laguerre, Rendiconti del Circolo Matematico di Palermo, 34 (1912), 1—32.

Th. Motzkin,
[3] Beiträge zur Theorie der linearen Ungleichungen, Doctoral dissertation, Basel 1933, published in Jerusalem 1936, especially Chapter IV.

PETER SCHERK,

[4] Über differenzierbare Kurven und Bögen, Casopis pro pěstování Matematiky a Fysiky, **66** (1937), 165—191.

I. J. SCHOENBERG,

[5] Über variationsvermindernde lineare Transformationen, Mathematische Zeitschrift, **32** (1930), 321—328.

I. J. SCHOENBERG,

[6] On Polya frequency functions II: Variation-diminishing integral operators of the convolution type. Acta Szeged, **12** (1950), 97—106.

(Oblatum 31-10-50)