

PROOF OF THE VAN DER WAERDEN CONJECTURE REGARDING THE PERMANENT
OF A DOUBLY STOCHASTIC MATRIX

D. I. Falikman

In 1926 van der Waerden formulated the following conjecture [1-3]: the permanent of every double stochastic $(n \times n)$ matrix is not less than $n!/n^n$.

Since then various results have been obtained regarding the van der Waerden conjecture. In particular, it has been proved for $n \leq 5$. In [2] one has proved the following: if on the set of all doubly stochastic $(n \times n)$ matrices, the permanent attains its least value at a matrix without zero elements, then its permanent is equal to $n!/n^n$.

The field of real numbers will be denoted by \mathbf{R} . Let $X = (x_{ij})$ be a matrix of dimension $(n \times n)$ with elements from \mathbf{R} . If the conditions

$$\begin{aligned} x_{ij} &\geq 0 \quad (1 \leq i, j \leq n), \\ x_{i1} + \dots + x_{in} &= 1 \quad (1 \leq i \leq n), \\ x_{1j} + \dots + x_{nj} &= 1 \quad (1 \leq j \leq n), \end{aligned}$$

hold, then the matrix X is said to be doubly stochastic. The set of all doubly stochastic matrices of order n will be denoted by Ω_n . This is a closed subset of \mathbf{R}^n . Moreover, it is bounded: if $X = (x_{ij}) \in \Omega_n$, then $0 \leq x_{ij} \leq 1$ for $1 \leq i, j \leq n$. Consequently, Ω_n is compact. We consider now the set

$$\Omega_n^* = \{X = (x_{ij}) \in \Omega_n \mid x_{ij} \neq 0 \text{ for } 1 \leq i, j \leq n\}.$$

If $X = (x_{ij}) \in \Omega_n$ and $\delta > 0$, then

$$X_\delta = \left(\frac{x_{ij} + \delta}{1 + n\delta} \right) \in \Omega_n^*.$$

Since $X_\delta \rightarrow X$ for $\delta \rightarrow 0$, it follows that Ω_n^* is dense in Ω_n .

On the set of all $(n \times n)$ matrices with elements from \mathbf{R} we define two numerical functions: the product and the permanent. Namely, let $X = (x_{ij})$ be an $(n \times n)$ matrix with elements from \mathbf{R} . Then

$$\Pi(X) = \prod_{1 \leq i, j \leq n} x_{ij} \text{ and } \text{per}(X) = \sum_{\sigma \in S_n} \prod_{i=1}^n x_{i\sigma(i)},$$

where S_n is the set of all $n!$ permutations of the set $\{1, \dots, n\}$. One can verify that the permanent is a symmetric semilinear function of the rows (columns).

THEOREM 1. If $X \in \Omega_n$, then $\text{per}(X) \geq n!/n^n$.

In other words, the van der Waerden conjecture is true. The proof is based on a series of lemmas. First we define a certain family of functions.

We take $\varepsilon \in \mathbf{R}$. Let $X = (x_{ij})$ be an $(n \times n)$ matrix with elements from \mathbf{R} such that $x_{ij} \neq 0$ for $1 \leq i, j \leq n$. We set

$$F_\varepsilon(X) = \text{per}(X) + \varepsilon/\Pi(X).$$

Thus, the function F_ε is defined on the open set $(\mathbf{R} - 0)^n \subset \mathbf{R}^n$, and on this set it is continuous and differentiable. Since $(1/x)' = -1/x^2$ for $x \neq 0$, we have

$$\frac{\partial F_\varepsilon}{\partial x_{ij}} = \text{per}(X_{ij}) - \frac{\varepsilon}{x_{ij} \Pi(X)}, \quad (1)$$

where $X_{i,j}$ is the matrix obtained from the matrix X by deleting the i -th row and the j -th column.

LEMMA 1. Let $\varepsilon > 0$. Then there exists a point of minimum of the function F_ε in the set Ω_n^* .

Proof. We set $F = F_\varepsilon$. Since $\varepsilon > 0$, for all $X \in \Omega_n^*$ we have $F(X) > 0$. Therefore, there exists the infimum

$$a = \inf_{X \in \Omega_n^*} F(X), \quad a \geq 0.$$

We select a sequence $(A_m)_{m \geq 1}$ of elements of the set Ω_n^* , for which $F(A_m) \rightarrow a$ for $m \rightarrow +\infty$. Since $A_m \in \Omega_n^* \subset \Omega_n$, and the set Ω_n is compact, it follows, switching to a subsequence, that one can assume that $A_m \rightarrow A \in \Omega_n$ for $m \rightarrow +\infty$. If $A \notin \Omega_n^*$, then $\Pi(A) = 0$ and $\Pi(A_m) \rightarrow 0$. Then from the inequality

$$F(A_m) > \varepsilon / \Pi(A_m)$$

and from the condition $\varepsilon > 0$ we obtain the contradiction $F(A_m) \rightarrow +\infty$. Consequently, $A \in \Omega_n^*$, $F(A_m) \rightarrow F(A)$, $a = F(A)$. The lemma is proved.

LEMMA 2. Let $\varepsilon > 0$ and let $A = (a_{ij}) \in \Omega_n^*$ be a point of minimum of the function F_ε in the set Ω_n^* . Then $A = (1/n)$.

Proof. We set $p = \text{per}(A)$, $c = \varepsilon / \Pi(A)$; $p_{ij} = \text{per}(A_{ij})$, $d_{ij} = \left. \frac{\partial F_\varepsilon}{\partial x_{ij}} \right|_{\lambda=A}$ ($1 \leq i, j \leq n$). We fix $1 < i, j \leq n$ and we consider the matrix

$$A_h = \begin{pmatrix} a_{11} + h & \dots & a_{1j} - h & \dots \\ \dots & \dots & \dots & \dots \\ a_{i1} - h & \dots & a_{ij} + h & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix},$$

where $h \in \mathbb{R}$, which differs from the matrix A only by the elements with the indices $(1, 1)$, $(1, j)$, $(i, 1)$, (i, j) . We take $\rho = \min(a_{11}, a_{1j}, a_{i1}, a_{ij}) > 0$. If $|h| < \rho$, then $A_h \in \Omega_n^*$, so that

$$F_\varepsilon(A_h) = F_\varepsilon(A) + (d_{11} - d_{1j} - d_{i1} + d_{ij})h + o(h) \geq F_\varepsilon(A).$$

Consequently, $d_{11} - d_{1j} - d_{i1} + d_{ij} = 0$ for $1 < i, j \leq n$. If $i = 1$ or $j = 1$, then the last equality is automatically satisfied so that it is valid for all $1 \leq i, j \leq n$. Thus,

$$d_{ij} = d_{i1} + d_{1j} - d_{11} \quad (1 \leq i, j \leq n).$$

We set $\lambda_i = d_{i1} - d_{11}$ for $1 \leq i \leq n$, $\mu_j = d_{1j}$ for $1 \leq j \leq n$. Then $d_{ij} = \lambda_i + \mu_j$ for $1 \leq i, j \leq n$. We have obtained the Lagrange condition of the relative minimum.

Then, from (1) it follows that

$$p_{ij} - c/a_{ij} = \lambda_i + \mu_j \quad (1 \leq i, j \leq n). \quad (2)$$

We prove that $\lambda_1 = \dots = \lambda_n$ and $\mu_1 = \dots = \mu_n$, i.e., the right-hand side in (2) is a constant. We multiply both sides of (2) by a_{ij} :

$$a_{ij}p_{ij} - c = \lambda_i a_{ij} + \mu_j a_{ij} \quad (1 \leq i, j \leq n). \quad (3)$$

We consider i ($1 \leq i \leq n$) fixed and we sum (3) with respect to j . Making use of the expansion of the permanent of the matrix A relative to the i -th row and taking into account the equality $\sum_{j=1}^n a_{ij} = 1$, we obtain

$$\sum_{j=1}^n a_{ij}p_{ij} - nc = p - nc = \lambda_i \sum_{j=1}^n a_{ij} + \sum_{j=1}^n \mu_j a_{ij} = \lambda_i + \sum_{j=1}^n \mu_j a_{ij}.$$

Setting $b = p - nc$, we have

$$\lambda_i = b - \sum_{j=1}^n \mu_j a_{ij} \quad (1 \leq i \leq n). \quad (4)$$

Similarly, we consider j ($1 \leq j \leq n$) fixed and we sum (3) relative to i . Making use of the expansion of the permanent of the matrix A relative to the j -th column and taking into account the equality $\sum_{i=1}^n a_{ij} = 1$, we obtain

$$\sum_{i=1}^n a_{ij} p_{ij} - nc = p - nc = \sum_{i=1}^n \lambda_i a_{ij} + \mu_j \sum_{i=1}^n a_{ij} = \sum_{i=1}^n \lambda_i a_{ij} + \mu_j.$$

From here

$$\mu_j = b - \sum_{i=1}^n \lambda_i a_{ij} \quad (1 \leq j \leq n). \quad (5)$$

Now we express λ_i ($1 \leq i \leq n$) in terms of $\lambda_1, \dots, \lambda_n$. To this end, in the equality (4) we insert (5):

$$\begin{aligned} \lambda_i &= b - (a_{i1}\mu_1 + \dots + a_{in}\mu_n) \\ &= b - [a_{i1}(b - a_{11}\lambda_1 - \dots - a_{n1}\lambda_n) + \dots + \\ &\quad + a_{in}(b - a_{1n}\lambda_1 - \dots - a_{nn}\lambda_n)] = \\ &= (a_{i1}a_{11} + \dots + a_{in}a_{1n})\lambda_1 + \dots + (a_{i1}a_{n1} + \dots + a_{in}a_{nn})\lambda_n. \end{aligned}$$

Thus, $\lambda_i = b_{i1}\lambda_1 + \dots + b_{in}\lambda_n$, and $b_{i1}, \dots, b_{in} > 0$ and $b_{i1} + \dots + b_{in} = 1$. Then, taking $\lambda = \min(\lambda_1, \dots, \lambda_n)$, we have $\lambda_1 \geq \lambda, \dots, \lambda_n \geq \lambda$ and for some i we have $\lambda_i = \lambda$. For this i we write:

$$\lambda_i - \lambda = b_{i1}(\lambda_1 - \lambda) + \dots + b_{in}(\lambda_n - \lambda) = 0,$$

and therefore $\lambda_1 = \lambda, \dots, \lambda_n = \lambda$. Inserting these values into (5), we see that $\mu_j = b - \lambda$ for $1 \leq j \leq n$. Thus, (2) leads to the equality

$$p_{ij} = b + c/a_{ij} \quad (1 \leq i, j \leq n).$$

We have not used yet the condition $\varepsilon \geq 0$. From it we obtain that $c > 0$. Thus, in order to conclude the proof of Lemma 2, it is sufficient to show that we have

LEMMA 3. Assume that for the matrix $A = (a_{ij}) \in \Omega_n^*$ we have

$$\text{per}(A_{ij}) = b + c/a_{ij} \quad (1 \leq i, j \leq n), \quad (6)$$

where $b, c \in \mathbb{R}, c \geq 0$. Then $A = (1/n)$.

In turn, for the proof of Lemma 3 it is necessary to introduce and to investigate some symmetric bilinear forms on \mathbb{R}^n . This is done in the following two lemmas.

LEMMA 4. Let E be a vector space over \mathbb{R} and let $f: E \times E \rightarrow \mathbb{R}$ be a symmetric bilinear form. We assume that there exists a vector $a \in E$ for which $f(a, a) = 0$ and, if $x \in E, x \notin \mathbb{R}a$, $f(x, a) = 0$, then $f(x, x) < 0$. In this case, if $t, s \in E, t \neq 0, f(t, s) = 0, f(s, s) > 0$, then $f(t, t) < 0$.

Proof. We note that if $x \in E, f(x, a) = 0$, then $f(x, x) \leq 0$. Indeed, either $x \notin \mathbb{R}a$, when $f(x, x) < 0$, or $x \in \mathbb{R}a$, when $f(x, x) = 0$. In particular, $f(s, a) \neq 0$.

We select a number $\eta \in \mathbb{R}$ for which $f(t + \eta s, a) = f(t, a) + \eta f(s, a) = 0$. Therefore $f(t + \eta s, t + \eta s) = f(t, t) + \eta^2 f(s, s) \leq 0$. If $\eta \neq 0$, then we obtain $f(t, t) \leq -\eta^2 f(s, s) < 0$. Let $\eta = 0$; consequently, $f(t, a) = 0$. If $t \notin \mathbb{R}a$, then $f(t, t) < 0$. We show that the inclusion $t \in \mathbb{R}a$ is not possible. Indeed, let $t = \tau a, \tau \in \mathbb{R}$. Then $f(t, s) = \tau f(a, s) = 0$, whence $\tau = 0, t = 0$, which is a contradiction. Thus, $f(t, t) < 0$. Lemma 4 is proved.

LEMMA 5. Let $C = (c_{ij})$ be an $(n-2) \times n$ matrix with elements from \mathbb{R} and such that $c_{ij} > 0$ for $1 \leq i \leq n-2, 1 \leq j \leq n$. We define a symmetric bilinear form $f: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ in the following manner:

$$f(x, y) = \text{per} \begin{pmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \\ \dots & \dots & \dots \\ c_{i1} & \dots & c_{in} \\ \dots & \dots & \dots \end{pmatrix}$$

for all $x = (x_1, \dots, x_n) \in \mathbb{R}^n, y = (y_1, \dots, y_n) \in \mathbb{R}^n$. In this case, if $t, s \in \mathbb{R}^n, t \neq 0, f(t, s) = 0, f(s, s) > 0$, then $f(t, t) < 0$.

Proof. We use induction on $n \geq 2$. For $n = 2$ we write $t = (t_1, t_2), s = (s_1, s_2)$. We have $f(t, s) = t_1 s_2 + t_2 s_1 = 0, f(s, s) = 2s_1 s_2 > 0$. In particular, $s_1 \neq 0, s_2 \neq 0$. If $t_1 = 0$, then $t_2 s_1 = 0, t_2 = 0, t = 0$, which is a contradiction. Thus, $t_1 \neq 0$. We multiply both sides of the equality $t_2 s_1 = -t_1 s_2$ by $t_1 s_2$. Then $t_1 t_2 s_1 s_2 = -t_1^2 s_2^2 < 0$, and, consequently, $t_1 t_2 < 0$. Thus, $f(t, t) = 2t_1 t_2 < 0$, so that for $n = 2$ the lemma is proved.

Assume now that $n > 2$ and assume that for $n-1$ the lemma holds. We select the vector $e_n = (0, \dots, 0, 1) \in \mathbb{R}^n$. Then $f(e_n, e_n) = 0$. Let $x = (x_1, \dots, x_{n-1}, x_n) \in \mathbb{R}^n$ and also $x \notin \mathbb{R}e_n, f(x, e_n) = 0$. We consider the vector $x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$. Since $x \notin \mathbb{R}e_n$, we have $x' \neq 0$, and since $f(x, e_n) = 0$, we have

$$\text{per} \begin{pmatrix} x_1 & \dots & x_{n-1} & x_n \\ 0 & \dots & 0 & 1 \\ \dots & \dots & \dots & \dots \\ c_{i1} & \dots & c_{i,n-1} & c_{in} \\ \dots & \dots & \dots & \dots \end{pmatrix} = \text{per} \begin{pmatrix} x_1 & \dots & x_{n-1} \\ \dots & \dots & \dots \\ c_{i1} & \dots & c_{i,n-1} \\ \dots & \dots & \dots \end{pmatrix} = 0. \quad (7)$$

We write

$$f(x, x) = \text{per} \begin{pmatrix} x_1 & \dots & x_{n-1} & x_n \\ x_1 & \dots & x_{n-1} & x_n \\ \dots & \dots & \dots & \dots \\ c_{i1} & \dots & c_{i,n-1} & c_{in} \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

and we expand the permanent with respect to the last column. Taking into account (7), we have

$$f(x, x) = \sum_{i=1}^{n-2} c_{in} f_i(x', x'), \quad (8)$$

where $f_i: \mathbf{R}^{n-1} \times \mathbf{R}^{n-1} \rightarrow \mathbf{R}$ is a form constructed over the matrix C_{in} obtained from the matrix C by the deletion of the i -th row and the n -th column. We consider the vector $c'_i = (c_{i1}, \dots, c_{i,n-1}) \in \mathbf{R}^{n-1}$. Equality (7) can be written as: $f_i(x', c'_i) = 0$. In addition, $f_i(c'_i, c'_i) > 0$, $x' \neq 0$. Therefore, by the inductive hypothesis, $f_i(x', x') < 0$ for $1 \leq i \leq n-2$. Now, from expansion (8) it follows that $f(x, x) < 0$.

Thus, the form $f: \mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}$ satisfies the conditions of Lemma 4 (for $\alpha = e_n$). Consequently, if $t, s \in \mathbf{R}^n$, $t \neq 0$, $f(t, s) = 0$, $f(s, s) > 0$, then $f(t, t) < 0$. Lemma 5 is proved.

Proof of Lemma 3. We select in the matrix A two rows $u = (u_1, \dots, u_n) \in \mathbf{R}^n$, $v = (v_1, \dots, v_n) \in \mathbf{R}^n$:

$$A = \begin{pmatrix} \dots & \dots & \dots \\ u_1 & \dots & u_n \\ \dots & \dots & \dots \\ v_1 & \dots & v_n \\ \dots & \dots & \dots \end{pmatrix}.$$

Since $A \in \mathcal{D}_n^*$, we have $u_i > 0$, $v_i > 0$ for $1 \leq i \leq n$, $\sum_{i=1}^n u_i = 1$, $\sum_{i=1}^n v_i = 1$. We show that $u = v$.

In matrix A we consider all the remaining rows fixed and we define a symmetric bilinear form $f: \mathbf{R}^n \times \mathbf{R}^n \rightarrow \mathbf{R}$ in the following manner:

$$f(x, y) = \text{per} \begin{pmatrix} \dots & \dots & \dots \\ x_1 & \dots & x_n \\ \dots & \dots & \dots \\ y_1 & \dots & y_n \\ \dots & \dots & \dots \end{pmatrix}$$

for all $x = (x_1, \dots, x_n) \in \mathbf{R}^n$, $y = (y_1, \dots, y_n) \in \mathbf{R}^n$.

Thus, in the last matrix and in matrix A , the selected rows occupy the same places and all the remaining rows coincide.

Assume, as usual, that e_1, \dots, e_n is the standard basis of the space \mathbf{R}^n so that for every vector $x = (x_1, \dots, x_n) \in \mathbf{R}^n$ we have $x = \sum_{i=1}^n x_i e_i$.

From conditions (6) it follows that

$$f(e_i, v) = b + c/u_i, \quad f(u, e_i) = b + c/v_i \quad (1 \leq i \leq n).$$

We consider the vector $t = (t_1, \dots, t_n) \in \mathbf{R}^n$, where $t_i = u_i - v_i$ for $1 \leq i \leq n$. Then $t = u - v$ and

$$f(t, e_i) = f(u, e_i) - f(v, e_i) = c \left(\frac{1}{v_i} - \frac{1}{u_i} \right) = \frac{c}{u_i v_i} t_i \quad (1 \leq i \leq n).$$

Therefore

$$f(t, t) = \sum_{i=1}^n f(t, e_i) t_i = c \sum_{i=1}^n \frac{t_i^2}{u_i v_i} \geq 0.$$

We consider the vector $s = (s_1, \dots, s_n) \in \mathbb{R}^n$, where $s_i = u_i v_i > 0$ for $1 \leq i \leq n$. We have

$$f(t, s) = \sum_{i=1}^n f(t, e_i) s_i = c \sum_{i=1}^n t_i = 0.$$

In addition, $f(s, s) > 0$. Since Lemma 5 can be applied to the form f , from the conditions $f(t, t) \geq 0$, $f(t, s) = 0$, $f(s, s) > 0$ there follows that $t = 0$, $u = v$.

Thus, all the rows of matrix A are mutually equal. Since their sum is the row $(1, \dots, 1)$, we have $A = (1/n)$. Lemma 3 and, simultaneously, Lemma 2 are proved.

Proof of Theorem 1. We select $\varepsilon > 0$. From Lemmas 1 and 2 it follows that the matrix $(1/n) \in \Omega_n^*$ is a point of minimum of the function F_ε in the set Ω_n^* . Thus, if $X \in \Omega_n^*$, then

$$\text{per}(X) + \varepsilon / \Pi(X) \geq n! / n^n + \varepsilon n^n$$

for any $\varepsilon > 0$. From here for $\varepsilon \rightarrow 0$ we obtain that

$$\text{per}(X) \geq n! / n^n$$

for all $X \in \Omega_n^*$. Since Ω_n^* is dense in Ω_n , the continuity of the permanent implies that the last inequality holds for all $X \in \Omega_n$. Theorem 1 is proved.

THEOREM 2. Let $A \in \Omega_n^*$ and $\text{per}(A) = n! / n^n$. Then $A = (1/n)$.

Proof. By virtue of Theorem 1, the point $A \in \Omega_n^*$ is a point of minimum of per in Ω_n^* . Therefore, from Lemma 2 for $\varepsilon = 0$ it follows that $A = (1/n)$. Theorem 2 is proved.

LITERATURE CITED

1. B. L. van der Waerden, "Aufgabe 45," Jber. Deutsch. Math. Verein., 35, 117 (1926).
2. M. Marcus and M. Newman, "On the minimum of the permanent of a doubly stochastic matrix," Duke Math. J., 26, 61-72 (1959).
3. H. J. Ryser, Combinatorial Mathematics, Math. Assoc. Am. (1963).

OBSTRUCTIONS TO LOCAL EQUIVALENCE OF DISTRIBUTIONS

A. N. Varchenko

An n -dimensional distribution on the space \mathbb{R}^{n+k} is a smooth field σ of n -dimensional tangential directions, i.e., a function that associates with each point $x \in \mathbb{R}^{n+k}$ an n -dimensional linear subspace σ_x of the tangent space $T_x \mathbb{R}^{n+k}$ [1]. Two n -dimensional distributions on \mathbb{R}^{n+k} are said to be equivalent if there exists a diffeomorphism of the space \mathbb{R}^{n+k} that transforms one of the distributions into the other one. In this note we indicate a natural local invariant of a distribution. It is proved that this invariant takes different values at different points for a quite general germ of an eight-dimensional distribution on \mathbb{R}^{11} .

1. Definitions. Let there be given an n -dimensional distribution σ on \mathbb{R}^{n+k} . Define a skew-symmetric bilinear function $\varphi_{\sigma, x}$ at each point $x \in \mathbb{R}^{n+k}$ on the linear space σ_x with values in $T_x \mathbb{R}^{n+k} / \sigma_x$. Let $u, v \in \sigma_x$ and suppose that U and V are arbitrary vector fields with the following properties: The values of these fields at each point of a certain neighborhood of x belong to a plane of the distribution σ , $U(x) = u$, $V(x) = v$. Let $\pi_x: T_x \mathbb{R}^{n+k} \rightarrow T_x \mathbb{R}^{n+k} / \sigma_x$ denote the quotient mapping. Set $\varphi_{\sigma, x}(u, v) = \pi_x([U, V](x))$, where $[\cdot, \cdot]$ is the commutator of vector fields. The independence of this expression from the choice of the fields U and V is proved by the following obvious lemma.

LEMMA 1. Let σ be an n -dimensional distribution on \mathbb{R}^{n+k} , and U and V be vector fields on \mathbb{R}^{n+k} whose values belong to the planes of the distribution σ . Suppose that $U(x) = 0$ at a certain point $x \in \mathbb{R}^{n+k}$. Then $[U, V](x) \in \sigma_x$.

Proof. We call a collection of n vector fields, whose values at each point in a certain neighborhood of a point x form a basis of a plane of the distribution σ , the basic