Let v be the midpoint of the geodesic $yz_1|[u, x]$, and let $\lambda \in I(v)$ be a rotation which takes u into x. Then $\lambda(L_1)$ and $\lambda(L_2)$ are two geodesics which join x with $\lambda(w_2) \in U$. However, $\lambda(L_1)$ does not coincide to $\lambda(L_2)$, because L_1 is not the same as L_2: by Proposition 1 of Sec. 3, $\varphi(y) \neq y$, while $y \in L_1$, and $\varphi(y) \in L_2$. But the existence of distinct geodesics which join x with $\lambda(w_2)$ contradicts Theorem 3, and this proves (c).

(d) Consequently, M is a Buseman G-space. Now we see that Theorem 4 is a result of Theorem B [2]. This completes the proof of Theorem 4.

Proof of the Corollary. The two-dimensional locally Euclidean G-spaces are described in [2]. They are: Euclidean plane, cylinder, Möbius band, torus and Klein bottle. The only two-dimensional locally spherical G-spaces are the sphere and the projective plane [2]. The description of the two-dimensional locally hyperbolic spaces is given in [7] (see also [2]).

LITERATURE CITED

1. H. Freudenthal, "Fassungen des Riemann–Helmholtz–Lieschen Raumproblems," Math. Z., 63, 374-405 (1956).
2. H. Buseman, Geometry of Geodesics, Academic Press (1955).
3. D. Hilbert, Grundlagen der Geometrie, B. G. Teubner, Leipzig (1909).
4. A. Kolmogoroff, Zur topologisch-gruppentheoretischen Begründung der Geometrie, Nachr. Akad. Wiss. Göttingen Math. Phys., 208-210 (1930).
5. Wang Hsien-Chung, "Two-point homogeneous spaces," Ann. Math., 55, 177-191 (1955).
6. J. Tits, "Etude de certain espaces metriques," Bull. Soc. Math. Belgique, 5, 44-52 (1952).
7. F. Klein, Non-Euclidean Geometry [in German], Chelsea Publ.
8. A. K. Guts, "Remarks on the Helmholtz–Lie problem," Dokl. Akad. Nauk SSSR, 249, 780-783 (1979).
9. K. Kuratowski, Topology, Vol. 2, Academic Press (1969).
10. S. Helgason, Differential Geometry and Symmetric Spaces, Academic Press (1962).
11. L. S. Pontryagin, Topological Groups, Gordan and Breach (1966).
12. D. Montgomery and L. Zippin, "Topological transformation groups. I," Ann. Math., 41, No. 2, 778-791 (1940).
13. A. D. Aleksandrov and V. A. Zalgaller, "Two-dimensional manifolds with bounded curvature," Tr. Steklov Mat. Inst. Akad. Nauk SSSR, 63 (1962).

PROOF OF THE VAN DER WAERDEN CONJECTURE FOR PERMANENTS

G. P. Erorychev

UDC 519.10+3.918.3

1°. We prove (Theorem 1) the validity of the van der Waerden conjecture, formulated by him in 1926 ([1; 2, p. 155, Conjecture 1]), regarding the minimum of the permanent of a double stochastic matrix. In the course of the proof one answers positively (Theorem 2) the Marcus–Newman conjecture on the permanent of a doubly stochastic matrix ([2], p. 156, Conjecture 11; [3], Conjecture 11). The proof of Theorem 2, and with it also that of Theorem 1, is based on the representation of the permanent in terms of mixed discriminants and on the subsequent use of a geometric inequality for the permanents (Lemma), which follows directly from Aleksandrov's known inequalities for mixed discriminants [4]. The reduction from Theorem 2 to Theorem 1 is known and is based on the results of [5, 6]. As a consequence of Theorem 1 we obtain lower estimates, formulated previously by other authors (see [2], Sec. 8.2; 7, 8) under the assumption of the validity of the van der Waerden conjecture and improving in an essential manner the known estimates for the number of Latin rectangles and squares, the number of nonisomorphic Steiner triple systems and for the key constant λ_d in the d-dimensional dimer problem. We indicate some other applications of the results obtained in the paper.

2°. By the permanent of an $n \times n$ matrix $A = (a_{ij})$ over the field of complex numbers we mean the expression (see, for example, [9])

L. V. Kirenskii Institute of Physics, Siberian Branch, Academy of Sciences of the SSSR, Krasnoyarsk.

Translated from Sibirskii Matematicheskii Zhurnal, Vol. 22, No. 6, pp. 65-71, November-December, 1981. Original article submitted November 24, 1980.

0037-4466/82/2206-0854$07.50

$$\operatorname{per} A = \sum_{\sigma\in S_n} a_{1\sigma(1)} \dots a_{n\sigma(n)},$$

where S_n is the permutation group of order n. A nonnegative matrix A is said to be doubly stochastic if the sum of its elements in each row and in each column is equal to 1; Ω_n is the set of all n × n doubly stochastic matrices; a matrix $A \in \Omega_n$ is said to be minimizing if $\operatorname{per} A = \min_{X\in\Omega_n} \operatorname{per} X$; $J_n \in \Omega_n$ is the matrix with all elements equal to 1/n; A(i/j) is the (n − 1) × (n − 1) matrix obtained from the n × n matrix A by deleting the i-th row and the j-th column of the matrix A; $A = (a_1, \dots, a_n)$, where a_j is the j-th column of the matrix A; $A\begin{pmatrix} i, & j \\ x, & y\end{pmatrix}$ is the n × n matrix obtained from the n × n matrix A by replacing the i-th column by the n-vector x and the j-th column by the n-vector y.

3°. The following conjecture has been formulated by van der Waerden ([1]; [2], p. 155, Conjecture 1): if $A \in \Omega_n$, then

$$\operatorname{per} A \geqslant \frac{n!}{n^n},$$

and the equality is attained if and only if $A = J_n$. The van der Waerden conjecture and the various aspects of its solution have been devoted a large number of investigations (see the surveys in [2, 3]); there one gives a list of solved and unsolved conjectures and problems for permanents (at the moment of the compilation of the lists). In [5] it is proved that: a) if $A = (a_{ij})$ is a minimizing matrix, then per A(i/j) = per A for all $a_{ij} > 0$; b) if A is a positive minimizing matrix, then $A = J_n$. Marcus and Newman have formulated the following conjecture ([3], Conjecture 11; [2], pp. 156-157, Conjecture 11): there exists no matrix $A = (a_{ij}) \in \Omega_n$ such that

$$\begin{aligned} &\operatorname{per} A(i/j) = \operatorname{per} A \quad \text{for all} \quad i, j \in \overline{1, n}, \text{ for some} \quad a_{ij} \neq 0, \\ &\operatorname{per} A(i/j) \geqslant \operatorname{per} A \quad \text{for all} \quad i, j \in \overline{1, n}, \text{ for some} \quad a_{ij} = 0, \end{aligned}$$

and per A(i/j) > per A for some pair $s, t \in \overline{1, n}$, for which $a_{st} = 0$.

Making use of the results of [5], London has proved [6] that if A is a minimizing matrix, then

$$\operatorname{per} A(i/j) \geqslant \operatorname{per} A \ \text{for all} \quad i, j \in \overline{1, n}. \tag{1}$$

We also know the following

Proposition 1. (See, for example, [2], p. 101, Problem 18). Let A be an arbitrary minimizing matrix for which we have the equality

$$\operatorname{per} A(i/j) = \operatorname{per} A \ \text{for all} \quad i, j \in \overline{1, n}. \tag{2}$$

Then, $A = J_n$.

4°. The purpose of the present paper is the proof of the following statements.

THEOREM 1.

a) $\min_{X\in\Omega_n} \operatorname{per} X = \frac{n!}{n^n}$.

b) $X \in \Omega_n$, $\operatorname{per} X = \frac{n!}{n^n}$ if and only if $X = J_n$.

THEOREM 2. If $A \in \Omega_n$ and the inequalities (1) hold, then the equalities (2) hold.

LEMMA (a geometric inequality for the permanent).

a) Let $f_1, \dots, f_{n-1}$ be n-vectors with nonnegative elements and let g be an n-vector with real elements. Then

$$\operatorname{per}^2 (f_1, \dots, f_{n-2}, f_{n-1}, g) \geqslant \operatorname{per} (f_1, \dots, f_{n-2}, f_{n-1}, f_{n-1}) \operatorname{per} (f_1, \dots, f_{n-2}, g, g). \tag{3}$$

b) Let $f_1, \dots, f_{n-1}$ be n-vectors with positive elements. Equality in (3) is attained if and only if $g = \lambda f_{n-1}$, λ being a constant.

5°. The result of the lemma presents an independent interest for the investigation of the permanents, nonnegative matrices and also in information theory. This result is based on the representation of the permanent in terms of mixed discriminants and follows directly from Aleksandrov's known results on mixed discriminants, for whose formulation it is necessary to introduce the following definitions (we adhere to the notations of Sec. 1 of [4]). Let $f_1, \ldots, f_m$ be quadratic forms of n variables:

$$f_k = \sum_{i,j=1}^{n} a_{ij}^{(k)} x_i x_j.$$

Their linear combination $f = \lambda_1 f_1 + \ldots + \lambda_m f_m$ is also a quadratic form of the same n variables with coefficients $a_{ij} = \lambda_1 a_{ij}^{(1)} + \ldots + \lambda_m a_{ij}^{(m)}$. The discriminant $D(f, \ldots, f)$ of the form f is the homogeneous polynomial of degree n with respect to $\lambda_1, \ldots, \lambda_m$:

$$D(f, \ldots, f) = \sum_{k_1, \ldots, k_n} \lambda_{k_1} \ldots \lambda_{k_n} D(f_{k_1}, \ldots, f_{k_n}).$$

Here the coefficient $D(f_{k_1}, \ldots, f_{k_n})$ of the product $\lambda_{k_1}, \ldots, \lambda_{k_n}$, called the mixed discriminant of the quadratic forms $f_{k_1}, \ldots, f_{k_n}$, depends only on the forms $f_{k_1}, \ldots, f_{k_n}$ and it is chosen in such a way that it does not depend on the orders of the forms $f_{k_1}, \ldots, f_{k_n}$, while the summation is taken over all indices $k_1, \ldots, k_n$ which run independently through all the values from 1 to m.

THEOREM. (Aleksandrov's inequality for mixed discriminants [4].) a) Let $f_1, \ldots, f_{n-1}$ be positive definite quadratic forms and let g be a quadratic form. Then

$$D^2(f_1, \ldots, f_{n-2}, f_{n-1}, g) \geqslant D(f_1, \ldots, f_{n-2}, f_{n-1}, f_{n-1}) D(f_1, \ldots, f_{n-2}, g, g). \tag{4}$$

b) Equality in (4) is attained if and only if $g = \lambda f_{n-1}$, λ being a constant.

Since a nonnegative definite form is the limit of a sequence of positive definite forms, it follows that statement a) of the theorem is valid also for the case when $f_1, \ldots, f_{n-1}$ are nonnegative definite forms.

We prove that

$$\operatorname{per}\begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} = D(f_1, \ldots, f_n) n!,$$

where f_i is the quadratic form with the matrix

$$\begin{pmatrix} a_{1i} & & 0 \\ & \ddots & \\ 0 & & a_{ni} \end{pmatrix}.$$

Indeed, in this case we have

$$f = \lambda_1 f_1 + \ldots + \lambda_n f_n = \sum_{i=1}^{n} (\lambda_1 a_{i1} + \ldots + \lambda_n a_{in}) x_i^2;$$

$$D(f, \ldots, f) = \det\begin{pmatrix} \lambda_1 a_{11} + \ldots + \lambda_n a_{1n} & & 0 \\ & \ddots & \\ 0 & & \lambda_1 a_{n1} + \ldots + \lambda_n a_{nn} \end{pmatrix} = \prod_{i=1}^{n} (\lambda_1 a_{i1} + \ldots + \lambda_n a_{in}),$$

and, as it can be easily seen, the coefficient of $\lambda_1, \ldots, \lambda_n$ in the last expression is equal to per A. Now, inequalities (3) follows directly from the inequalities (4).

Remark 1. Inequalities (3) can be carried over immediately to the permanent of a rectangular matrix (for the definition and the properties of the permanent of a rectangular matrix, see, for example, [9]). In [4] (Sec. 3), one has obtained the following statement: let $f_1, \ldots, f_n$ be given positive forms of n variables. We have the inequality

$$\{D(f_1, \ldots, f_n)\}^m \geqslant \prod_{k=1}^{m} D(f_k, \ldots, f_k, f_{m+1}, \ldots, f_n), \tag{5}$$

where equality prevails if and only if all the forms $f_1, \ldots, f_k$ are proportional to each other. From the representation of the permanent in terms of mixed discriminants and from the inequality (5) we obtain by continuity the following proposition.

Proposition. a) Let $A = (a_1, \ldots, a_n)$ be a nonnegative $n \times n$ matrix. We have the inequality

$$\{\operatorname{per}(a_1, \ldots, a_n)\}^m \geqslant \prod_{k=1}^{m} \operatorname{per}(a_k, \ldots, a_k, a_{m+1}, \ldots, a_n). \tag{6}$$

b) Let $A = (a_1, \ldots, a_n)$ be a nonnegative $n \times n$ matrix. Equality in (6) is attained if and only if the n-vectors $a_1, \ldots, a_n$ are proportional to each other.

Remark 2. The possibility of the proof of the result of the lemma, using the relation between the permanent and Minkowski's mixed volumes, has been mentioned by the author in [10]. The role of the geometric inequalities (3) in the proof of the Marcus–Newman and van der Waerden conjectures for the permanent is the same as the role of their corresponding Brunn–Minkowski and Aleksandrov–Fenchel inequalities for mixed volumes, yielding the solution of many important extremal and uniqueness problems (see the survey in [11]).

6°. Proof of Theorem 2. By virtue of Birkhoff's known theorem, a matrix $A \in \Omega_n$ can be represented in the form of a convex combination of permutation matrices. From here it follows that if $A \in \Omega_n$, then

$$\operatorname{per} A \geqslant 0.$$

By the Laplace formula for permanents, applied to the elements of the j-th column, we have

$$\operatorname{per} A \begin{pmatrix} i, & j \\ a_i, & a_i \end{pmatrix} = \sum_k a_{ki} \operatorname{per} A(k/j), \tag{7}$$

and, similarly,

$$\operatorname{per} A \begin{pmatrix} i, & j \\ a_j, & a_j \end{pmatrix} = \sum_k a_{kj} \operatorname{per} A(k/i). \tag{8}$$

Making use of the conditions of the theorem, of the geometric inequality (3) for the permanents

$$\operatorname{per}^2 A \begin{pmatrix} i, & j \\ a_i, & a_j \end{pmatrix} \geqslant \operatorname{per} A \begin{pmatrix} i, & j \\ a_i, & a_i \end{pmatrix} \operatorname{per} A \begin{pmatrix} i, & j \\ a_j, & a_j \end{pmatrix}, \quad i, j = \overline{1, n},$$

and of the equalities (7), (8), we obtain the following system of inequalities relative to the numbers $\operatorname{per} A(i/j)$, $i, j = \overline{1, n}$:

$$\begin{gathered} \operatorname{per}^2 A \geqslant \Big(\sum_k a_{ki} \operatorname{per} A(k/j)\Big)\Big(\sum_k a_{kj} \operatorname{per} A(k/i)\Big), \\ \operatorname{per} A(i/j) \geqslant \operatorname{per} A \geqslant 0, \end{gathered} \tag{9}$$

for all $i, j \in \overline{1, n}$.

We show that the inequalities (9) have the solution

$$\operatorname{per} A(i/j) = \operatorname{per} A \quad \text{for all} \quad i, j \in \overline{1, n}.$$

We assume the opposite, i.e., there exists at least one pair $r, s \in \overline{1, n}$ such that

$$\operatorname{per} A(r/s) > \operatorname{per} A.$$

Since $A \in \Omega_n$, there exists $t \in \overline{1, n}$ such that $a_{rt} > 0$. Then, by virtue of (9) we have

$$\operatorname{per}^2 A = \operatorname{per}^2 A \begin{pmatrix} s, & t \\ a_s, & a_t \end{pmatrix} \geqslant \Big(\sum_k a_{ks} \operatorname{per} A(k/t)\Big)\Big(\sum_k a_{kt} \operatorname{per} A(k/s)\Big)$$

$$= \Big(\sum_k a_{ks} \operatorname{per} A(k/t)\Big)\Big(\sum_{\substack{k=1 \\ k \neq r}}^n a_{kt} \operatorname{per} A(k/s) + a_{rt} \operatorname{per} A(r/s)\Big) > \Big(\sum_{k=1}^n a_{ks} \operatorname{per} A\Big)\Big(\sum_{k=1}^n a_{kt} \operatorname{per} A\Big) = \operatorname{per}^2 A,$$

where the strict inequality follows from the inequalities $\operatorname{per} A(i/j) \geq \operatorname{per} A$, $a_{rt} > 0$, $\operatorname{per} A(r/s) > \operatorname{per} A$, $\operatorname{per} A > 0$. The obtained contradiction proves Theorem 2.

7°. From London's results [6] [see inequalities (1)] and Theorem 2 we obtain the following proposition.

Proposition 2. If A is a minimizing matrix, then equalities (2) hold for A.

From Propositions 1 and 2 there follows the validity of Theorem 1.

The proof of statement b) of [5], given in 3°, is, in our opinion, somewhat tedious and complicated (the complexity of the construction, the limiting process, etc.). In the following section we give another short geometric proof of this statement. This proof, with the use of Proposition 2 and of the geometric inequality (3) for the permanent, allows us to obtain a more transparent, geometric proof of Theorem 1.

8°. First we prove

Proposition 3. (See [6], Lemma 1). Let $0 \leq \lambda \leq 1$, let A be a minimizing matrix and for each $i, j \in \overline{1, n}$ the matrix $A' = (a_1', \ldots, a_n')$, where $a_i' = \lambda a_i + (1-\lambda)a_j$, $a_j' = \lambda a_j + (1-\lambda)a_i$ and $a_k' = a_k$ ($k \neq i, j$). Then: a) $A' \in \Omega_n$; b) $\operatorname{per} A' = \operatorname{per} A$; c) A' satisfies the equalities (2).

The condition a) is obvious; condition b) follows directly from the multilinearity and the symmetry of the function per A relative to the column vectors of the matrix A, from Proposition 2 and from Eqs. (7), (8); condition c) follows at once from a), b) and Proposition 2.

Now we show that if A is a minimizing matrix, then $A = J_n$. Clearly, the matrix A cannot be (to within a permutation of rows and columns) a matrix of the form

$$A = \begin{pmatrix} 1 & 0 \\ 0 & A_{n-1} \end{pmatrix},$$

where $A_{n-1} \in \Omega_{n-1}$; in this case, according to Proposition 2, $\operatorname{per} A = \operatorname{per} A(1/i) = 0$ ($i = 2, \ldots, n$), which is a contradiction. We fix some column of the matrix A, for example the n-th one, and we show that

$$a_n = \begin{pmatrix} 1/n \\ 1/n \\ \vdots \\ 1/n \end{pmatrix}.$$

Making use of Proposition 3, it is easy to see that the matrix A can be transformed in $n-2$ steps into the matrix $B = (b_1, \ldots, b_{n-1}, a_n)$, satisfying the properties a)-c) of Proposition 3 and such that the elements of its first row, excluding perhaps a_{1n}, are equal among themselves. We set $d = (a_{11} + \ldots + a_{1n-1})/(n-1)$ and $a_{1i_1} = \min\limits_{1 \leqslant i \leqslant n-1} a_{1i}$, $a_{1i_2} = \max\limits_{1 \leqslant i \leqslant n-1} a_{1i}$, $\lambda = (d - a_{1i_2})/(a_{1i_1} - a_{1i_2}) \leqslant 1$. With the aid of the transformation from Proposition 3 for $i = i_1$, $j = i_2$ we obtain $a_{1i_1}' = d$, etc. We note that the mentioned transformation, applied to the first $n-1$ columns of the matrix B, does not change the elements of the first row. Applying the same process to the elements of the second row, we achieve that its elements, excluding perhaps a_{2n}, will be equal among themselves, etc. As a result one obtains a matrix $C = (c_1, \ldots, c_{n-1}, a_n)$ in each row of which the first $n-1$ elements are equal among themselves (and positive) and the matrix C satisfies the conditions a)-c) of Proposition 3. From the condition of the strict equality in (3) $C \in \Omega_n$ it follows immediately that $c_1 = \ldots = c_{n-1} = a_n$, i.e., $A = J_n$. The proof is concluded.

9°. The permanent arises in the solutions of a series of important mathematical and physical problems [2], making use of its fundamental characteristic property, namely it gives the number of systems of distinct representatives of sets. In spite of the wide applicability of the permanent, there exist difficulties related to its computation and in connection with this there follows the importance of the estimation of the permanent for classes of matrices. In this circle of questions, van der Waerden's problem occupies a key position. From the statement of the lemma, from our proofs of Theorems 1 and 2 and from the validity of the van der Waerden and the Marcus–Newman conjecture, there follows the validity of a series of statements which have arisen in connection with the attempts of settling these conjectures and their generalizations (see [2], Sec. 8.4). These results, together with the description of the geometric characteristic of the permanent as Minkowski's mixed volume, the natural generalization of the van der Waerden conjecture to the multidimensional case and to mixed volumes, lower bounds for some combinatorial functions, other than the permanent, will be considered in the future. Currently, we can formulate a corollary to Theorem 1 in which we obtain lower bounds for certain combinatorial quantities which admit a representation in terms of the permanents of block doubly stochastic matrices of 0 and 1. These estimates, stated earlier by other authors (see [2], Sec. 8.2; [7, 8]), improve in an essential manner the known estimates.

COROLLARY. Let $L(r, n)$ be the number of Latin rectangles, $r \leq n$, and let $L(n, n)$ be the number of Latin squares of order n; let $N(v)$ be the number of non-isomorphic Steiner triple systems of order v; let λ_d be the key constant in the d-dimensional dimer problem. Then

$$L(r, n) \geqslant (n!)^r n^{n(1-r)} \prod_{t=1}^{r-1} (n-t)^n,$$

$$L(n, n) \geqslant (n!)^{2n} n^{-n^2};$$

$$N(v) \geqslant (e^{-5} v)^{v^2/6}$$

and for each $\varepsilon > 0$ and for sufficiently large v we have

$$N(v) \geqslant ((3^{-3/2}e^{-2} - \varepsilon)\, v)^{v^2/6};$$

$$\lambda_d \geqslant \frac{1}{2} \log d - 0{,}153.$$

Since [12]

$$N(v) \leqslant (e^{-1/2}v)^{v^2/6}$$

and [7]

$$\lambda_d \leqslant \frac{1}{2} \ln d,$$

it follows that for $v \to \infty$ we have

$$\ln N(v) \infty (v^2/6) \ln v$$

and for $d \to \infty$ we have

$$\lambda_d \infty 1/2 \ln d.$$

LITERATURE CITED

1. B. L. van der Waerden, "Aufgabe 45," Jber. Deutsch. Math. Verein., 35, 117 (1926).
2. H. Minc, Permanents, Encyclopedia of Mathematics and Its Applications, Vol. 6, Addison-Wesley, Reading, Mass. (1978).
3. M. Marcus and H. Minc, "Permanents," Am. Math. Mon., 72, 577-591 (1965).
4. A. D. Aleksandrov, "On the theory of mixed volumes of convex bodies, Part IV," Mat. Sb., 3, 227-251 (1938).
5. M. Marcus and M. Newman, "On the minimum of the permanent of a doubly stochastic matrix," Duke Math. J., 26, 61-72 (1959).
6. D. London, "Some notes on the van der Waerden conjecture," Linear Algebra Appl., 4, 155-160 (1971).
7. J. M. Hammersley, "An improved lower bound for the multidimensional dimer problem," Proc. Cambridge Phil. Soc., 64, 455-463 (1968).
8. R. M. Wilson, "Nonisomorphic Steiner triple systems," Math. Z., 135, No. 4, 303-313 (1974).
9. H. J. Ryser, Combinatorial Mathematics, Wiley, New York (1963).
10. G. P. Egorychev, "New formulas for the permanent," Dokl. Akad. Nauk SSSR, 254, No. 4, 784-787 (1980).
11. Yu. D. Burago and V. A. Zalgaller, Geometric Inequalities [in Russian], Nauka, Leningrad (1980).
12. J. Doyen and G. Valette, "On the number of non-isomorphic Steiner triple systems," Math. Z., 120, No. 2, 178-192 (1971).