Notes on the proof of the van der Waerden permanent conjecture

Vicente Valle Martinez
vvalle@iastate.edu

Follow this and additional works at: https://lib.dr.iastate.edu/creativecomponents
Part of the Discrete Mathematics and Combinatorics Commons

Recommended Citation
Valle Martinez, Vicente, "Notes on the proof of the van der Waerden permanent conjecture" (2018). Creative Components. 12.
https://lib.dr.iastate.edu/creativecomponents/12

This Creative Component is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Creative Components by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Notes on the proof of the van der Waerden permanent conjecture

by

Vicente Valle Martinez

A Creative Component submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Mathematics

Program of Study Committee:
Sung Yell Song, Major Professor
    Steve Butler
    Jonas Hartwig
    Leslie Hogben

Iowa State University
Ames, Iowa
2018

Copyright © Vicente Valle Martinez, 2018. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>iii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>v</td>
</tr>
<tr>
<td>CHAPTER 1. Introduction and Preliminaries</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Combinatorial interpretations</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Computational properties of permanents</td>
<td>4</td>
</tr>
<tr>
<td>1.3 Computational complexity in computing permanents</td>
<td>8</td>
</tr>
<tr>
<td>1.4 The organization of the rest of this component</td>
<td>9</td>
</tr>
<tr>
<td>CHAPTER 2. Applications: Permanents of Special Types of Matrices</td>
<td>10</td>
</tr>
<tr>
<td>2.1 Zeros-and-ones matrices</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Latin squares</td>
<td>14</td>
</tr>
<tr>
<td>2.3 Doubly stochastic matrices</td>
<td>15</td>
</tr>
<tr>
<td>CHAPTER 3. Proof of van der Waerden’s Permanent Conjecture</td>
<td>18</td>
</tr>
<tr>
<td>3.1 Permanents and quadratic forms</td>
<td>18</td>
</tr>
<tr>
<td>3.2 Key inequalities in permanents</td>
<td>21</td>
</tr>
<tr>
<td>3.3 Egorychev’s proof of van der Waerden’s conjecture</td>
<td>23</td>
</tr>
<tr>
<td>3.4 Applications of van der Waerden’s result</td>
<td>25</td>
</tr>
<tr>
<td>CHAPTER 4. Interesting Open Problems in Related Subjects</td>
<td>27</td>
</tr>
<tr>
<td>APPENDIX A. Additional Material</td>
<td>31</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1.1  Combinatorial interpretations of the permanent of a $3 \times 3$ matrix . 4
Figure 1.2  Permanent sum formula ($n = 3$) . . . . . . . . . . . . . . . . . . . 6
Figure 1.3  Permanent of matrix with row or column zero ($n = 3$) . . . . . . 7
Figure 2.1  Solutions to the derangement and ménage problems using permanents 12
Figure 2.2  Latin square examples . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 15
ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Sung-Yell Song for his guidance and support in the completion of this work. Dr. Song was always available for assistance and right from the outset helped me determine the scope of this project. Further, his dedication and key insights on the mathematical content of the proofs made it possible to steer this project to success.

I would also like to take an opportunity to thank my committee members Dr. Steve Butler, Dr. Jonas Hartwig, and Dr. Leslie Hogben for their valuable feedback and guidance in the review of this work as well as for their example, support and guidance throughout my graduate studies.
The permanent of an $n \times n$ matrix $A = (a_{ij})$ with real entries is defined by the sum

$$\sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}$$

where $S_n$ denotes the symmetric group on the $n$-element set $\{1, 2, \ldots, n\}$. In this creative component we survey some known properties of permanents, calculation of permanents for particular types of matrices and their applications in combinatorics and linear algebra. Then we follow the lines of van Lint’s exposition of Egorychev’s proof for the van der Waerden’s conjecture on the permanents of doubly stochastic matrices. The purpose of this component is to provide elementary proofs of several interesting known facts related to permanents of some special matrices. It is an expository survey paper in nature and reports no new findings.
CHAPTER 1. Introduction and Preliminaries

In order to quantify properties of matrices, it can be useful to study functions taking values from the matrix entries. This paper focuses on one such function, the permanent, as well as its properties and applications in linear algebra and combinatorial matrix theory. The permanent was introduced independently by Binet and Cauchy in 1812 and received renewed interest after new developments by Egorychev and Falikman after the resolution of van der Waerden’s conjecture. The permanent is defined as a polynomial in the entries of the matrix as follows.

**Definition 1** The permanent of an $n \times n$ matrix $A = (a_{ij})$ is defined by the sum

$$\text{per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}$$

where $S_n$ denotes the symmetric group on the set $[n] := \{1, 2, \ldots, n\}$ of $n$ elements.

Mathematicians have raised the question of what is counted by the sum and how the output of the function varies as we perturb the values of entries, rows, or columns of the original matrix. We aim to discuss the main properties and interpretations of the permanent of a given matrix, and some of the questions that may arise from their study. In the remainder of this chapter, we illustrate some of the common applications of permanents that arise in the study of combinatorial enumeration problems.

1.1 Combinatorial interpretations

We discuss two combinatorial interpretations of the permanent of a matrix $A$. We begin this section by recalling some basic definitions and standard graph terminology.
A graph $G$ is a pair consisting of a finite vertex set $V(G)$ and a finite edge set $E(G)$ (possibly empty) together with a relation that associates with each edge two vertices (not necessarily distinct) called its endvertices. An edge whose endvertices are equal is called a loop. If edges have the same endvertices they are called multiple edges. A graph is called simple if it has no loops or multiple edges. When two vertices are the endvertices of an edge we say that they are adjacent. A neighborhood of a vertex is the set of vertices adjacent to the vertex. A vertex is incident an edge if it is an endvertex of the edge. The degree of a vertex is the number of edges incident to the vertex. A path is a simple graph all of whose vertices are incident to one or two edges, and its vertices can be ordered to make a sequence so that two vertices are adjacent if and only if they are consecutive in the sequence of vertices. A path with an equal number of vertices and edges, (such a path is said to be closed with each vertex of degree two) is called a cycle. A set of vertices is said to be independent if no two distinct vertices in the set are adjacent. A graph is bipartite if its vertex set is the union of two disjoint independent sets called partite sets. A complete bipartite graph, denote by $K_{s,t}$ is a bipartite graph such that partite sets have size $s$ and $t$, and two vertices are adjacent if and only if they are in different partite sets. Given a graph, we can obtain a directed (or oriented) graph by replacing all edges of a graph $G$ by ordered pairs of vertices, (i.e., $\{x,y\} \in E(G)$ by $(x,y)$ or $(y,x)$). By definition, a directed graph (digraph, in short) is a pair $G = (V,E)$ where $V$ is a vertex set and $E \subseteq V \times V$ is a set of ordered pairs $(x,y)$ called (directed) edges.

The permanent as a sum of weights of cycle covers. A weighted graph is a (di)graph with (nonnegative) numerical labels on the edges. A cycle cover of a digraph is a collection of vertex-disjoint cycles that covers (i.e. their edges are incident to) all of the vertices in the graph. When the graph is weighted, define the weight of a cycle cover to be the product of the weights of the edges in each cycle.

Remark 1 Let $G$ be the digraph with ‘weighted’ adjacency matrix $A$; that is, $A$ is the matrix
whose rows and columns are indexed by the vertex set \( V(G) = [n] \) and its \((i,j)\)-entry, \( a_{ij} \), is the weight of the edge from vertex \( i \) to vertex \( j \). There is a natural bijection between a permutation \( \sigma \in S_n \) and a cycle cover of \( G \). In a cycle cover \( G \) every vertex \( v_i \in V \) has a successor \( v_{\sigma(i)} \) and \( \sigma \) is a permutation of \([n]\). Conversely, the matrix corresponding permutation \( \sigma \in S_n \) of the set \( V = [n] \) can be made to form a digraph \( G = (V, E) \) with directed edges \((v_i, v_{\sigma(i)})\) which is a cycle cover. Then \( \text{per}(A) \) is the sum of the weights of all cycle covers of \( G \).

The permanent as the sum of weights of perfect matchings. A matching of a graph \( G = (V, E) \) is a set of pairwise non-incident edges. A perfect matching is a matching incident to all vertices of the graph. Define the weight of a perfect matching to be the product of the weights of the edges in the perfect matching.

**Remark 2** Let \( G \) be the complete bipartite graph \( K_{n,n} \) with vertex set \( \{x_1, \ldots, x_n\} \cup \{y_1, \ldots, y_n\} \) and undirected edges with weight \( a_{ij} \) between \( x_i \) and \( y_j \). There is a natural bijection between permutations \( \sigma \in S_n \) and a perfect matching of \( G \) given by the edges \( \{x_i, y_{\sigma(i)}\} \), \( i \in [n] \).

Indeed, a permutation matrix corresponding to \( \sigma \in S_n \) is a perfect matching in \( G \) on \( 2n \) vertices where the partite sets \( X \) and \( Y \) are the row indices and column indices, respectively, and with undirected edges of the form \((x_i, y_{\sigma(i)}) = (x_{\sigma(i)}, y_i)\) where \( x_i \in X \) and \( y_i \in Y \). Conversely, a perfect matching of a bipartite graph on \( 2n \) vertices corresponds to a permutation matrix by the previous construction since each row index of the adjacency matrix is in a one-to-one correspondence with the column indices. Then \( \text{per}(A) \) is the sum of the weights of all perfect matchings of a graph.

Figure 1.1 shows the terms in the permanent of a \( 3 \times 3 \) matrix, along with the matrix representation of the permutations corresponding to each summand and their interpretations in terms of weight of cycle cover and weight of perfect matching listed above. Note the difference in labels for each case.
4

\[
\text{per} \left( \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \right) = ae_i + bfg + cdh + ceg + bdi + afh
\]

Figure 1.1 Combinatorial interpretations of the permanent of a 3 \times 3 matrix

1.2 Computational properties of permanents

The formula for the permanent bears much resemblance to that of the determinant:

\[
\text{det}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} \text{sgn}(\sigma) a_{i\sigma(i)}
\]

Despite their apparent similarity, as it is expanded in Section 1.3, these mathematical objects have very different computational complexity.

**No sign cancellation.** Due to the absence of the \(\text{sgn}(\sigma)\) term, the permanent of a matrix with repeated (or linearly dependent) columns or rows need not be zero as it is in the case for the determinant.

**Permutation invariance.** Since the permanent sums over all elements of the symmetric group \(S_n\), the permanent of a matrix is invariant under permutation of rows or columns. That is, if \(P_1\) and \(P_2\) are permutation matrices, \(\text{per}(A) = \text{per}(P_1AP_2)\) and \(\text{per}(A) = \text{per}(A^T)\).
**Permanent of a product.** A major distinction from the determinant is that the permanent does not have the ‘product property’.

\[
\text{per}(AB) \neq \text{per}(A) \cdot \text{per}(B).
\]  

Equality fails to hold as early as \( n = 2 \) as it can be verified by the substitution \( A = B = J \) where \( J \) is the all-ones matrix: \( \text{per}(J^2) = 8 \) whereas \( \text{per}(J)^2 = 4 \).

However, if a matrix \( M \) is an upper (equivalently lower) triangular block matrix

\[
M = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}
\]

in which both \( A \) and \( D \) are square,

\[
\text{per}(M) = \text{per}(A) \cdot \text{per}(D).
\]

In what follows, an \( n \times n \) matrix \( M \) containing an \( s \times (n - s) \) submatrix of all zero entries for any \( s = 1, 2, \ldots, n - 1 \), will be called (partial) decomposable. If there is no such \( s \), \( A \) will be called fully indecomposable.

**Permanent of a sum.** The permanent of a sum of matrices can be calculated as the sum over complementary matrix entries as follows (See Theorem 1.4 in (19)).

\[
\text{per}(A + B) = \sum_{r,c \subseteq [n], |r|=|c|} \text{per}(A[r,c]) \cdot \text{per}(B[\bar{r}, \bar{c}])
\]  

Where the permanent of a \( 0 \times 0 \) matrix is assumed to be 1. The notation \( A[r,c] \) indicates the submatrix of \( A \) removed by the selection of rows and columns indexed by \( r \) and \( c \), respectively. Also we denote by \( \bar{r} \) the complement of \( r \) in \( [n] \).

We note that \( \text{per}(A) \) and \( \text{per}(B) \) appear on the right side of (1.2) when \( |r| = |c| = n \) and \( |r| = |c| = 0 \), respectively. In particular, if all permanents involved were positive (not generally the case) then we would have subadditivity, i.e. \( \text{per}(A + B) \leq \text{per}(A) + \text{per}(B) \).
There are nevertheless some properties of determinant that have their equivalent formulation in terms of permanents. For instance, the permanent can be calculated by expanding along a row or column:

### Laplace expansion by (permanental) minors along a row or column.

Suppose $A$ is an $n \times n$ matrix and $A_{ij}$ denotes $A[\{i\}, \{j\}]$, the submatrix obtained by deleting the $i$th row and $j$th column. Then, for $i = 1, 2, \ldots, n$

$$\text{per}(A) = \sum_{j=1}^{n} a_{ij} \text{per}(A_{ij}) = \sum_{i=1}^{n} a_{ij} \text{per}(A_{ij}).$$

### Scalar multiple of row or column.

Let $A'$ be obtained by multiplying a row or column of $A$ by some scalar $\alpha$, then $\text{per}(A') = \alpha \cdot \text{per}(A)$. It then follows that $\text{per}(\alpha A) = \alpha^n \text{per}(A)$ for an $n \times n$ matrix.

This property can also be seen from the interpretation in terms of weight of cycle covers and weight of perfect matchings shown previously: Each element in a row or column of $A$ appears once in each of the terms summed in the definition. In particular, if a row or column is zero, as highlighted in Figure 1.3, then the permanent is zero: Any matching
must use \( x_1 \), but then the weight includes a multiplicative factor of zero. Similarly, any cycle cover must visit \( v_1 \), making the weight zero.

![Figure 1.3 Permanent of matrix with row or column zero (\( n = 3 \))](image)

This observation about zero entries generalizes to the following observation. Let \( A \) be an \( n \times n \) nonnegative matrix. Then \( \text{per}(A) = 0 \) if and only if \( A \) contains an \( s \times (n - s + 1) \) submatrix of zeros, for some \( s \in \{1, 2, \ldots, n\} \).

**Multilinearity in rows and columns.** Given a matrix \( A \), let \( A' \) be the matrix obtained by replacing the elements in the \( i \)th row of \( A \) by \( b_{ij} \) \( (j = 1, 2, \ldots, n) \), and let \( A'' \) denote the matrix obtained by replacing the \( i \)th row entries \( a_{ij} \) by \( a_{ij} + b_{ij} \). Then,

\[
\text{per}(A'') = \text{per}(A) + \text{per}(A').
\]

The following is a result of Binet and Cauchy for the product of two matrices as stated in the linear algebra handbook (cf. (8, Sec. 31-2)).

**Theorem 1** (Binet-Cauchy formula for permanents) If \( A \) and \( B \) are \( m \times n \) and \( n \times m \) matrices, respectively, with \( m \leq n \), then

\[
\text{per}(AB) = \sum_{\alpha} \frac{1}{\mu(\alpha)} \text{per}(A[\{1, 2, \ldots, m\}, \alpha]) \cdot \text{per}(B[\alpha, \{1, 2, \ldots, m\}]).
\]

The sum ranges over all nondecreasing sequences \( \alpha \) of \( m \) integers chosen from \( \{1, 2, \ldots, n\} \) and \( \mu(\alpha) = \alpha_1!\alpha_2!\ldots\alpha_n! \) where \( \alpha_i \) denotes the number of occurrences of integer \( i \) in \( \alpha \).

Finally, we present the following result due to Ryser which is an application of the generalized inclusion-exclusion principle.
Theorem 2 (Ryser’s formula) Given an $n \times n$ matrix $A$,

$$\text{per}(A) = (-1)^n \sum_{S \subseteq \{1, \ldots, n\}} (-1)^{|S|} \prod_{i=1}^{n} \sum_{j \in S} a_{ij}. \quad (1.3)$$

It can be evaluated using $O(2^{n-1}n^2)$ operations. This can be improved on by a factor of two by enumerating the subsets $S$ on Ryser’s formula in gray code order. We refer the interested reader to (22).

1.3 Computational complexity in computing permanents

In the Gaussian elimination algorithm, an $n \times n$ matrix $A$ can be written in the form $A = E_1 \cdots E_k T$ where $T$ is upper-triangular and $E_i$ are elementary matrices corresponding to row operations, and $k$ is a polynomial in $n$. Then, since the determinant of elementary matrices and triangular matrices can also be calculated in polynomial-time (as the product of the diagonal entries), $\det(A)$ can be calculated in polynomial-time. On the other hand, the permanent does not satisfy the product property given in (1.1) and therefore this computation technique does not apply. In fact, a result of Valiant (25) proves it belongs to the #P-complexity class, associated with enumerating decision problems in NP.

Calculating every permutation in $S_n$ from the definition becomes computationally infeasible as $n$ grows large. Computing the permanent by using the Laplace expansion formula across a row or column does not give much advantage, and often formula 1.3 is preferred.

Because of the computational complexity, approximation algorithms have been recently studied for matrices with special properties, and by counting random perfect matchings sampling from probability distributions. Karmarkar et al (11) give the best approximation algorithm known to date with a runtime that grows exponentially with the size of the matrix. A survey of results in this area of research is given by Luby in (14).
1.4 The organization of the rest of this component

Chapter 2 of this component contains permanent values for several special classes of matrices including zeros-and-ones matrices, nonnegative matrices and stochastic matrices. The König’s theorem on maximum matching number and minimum covering number of a finite bipartite graph and von Neumann and Birkhoff theorem on matrices having the same row sums and column sums are discussed in this chapter. In Chapter 3, we turn our attention specifically to the van der Waerden’s Permanent Conjecture, concerning a bound of the permanent values of doubly stochastic matrices. We conclude this component with a discussion with a possible future research problems in the applications of permanents of matrices in Chapter 4.
CHAPTER 2. Applications: Permanents of Special Types of Matrices

2.1 Zeros-and-ones matrices

For matrices $A = (a_{ij})$ with $a_{ij} \in \{0, 1\}$, the combinatorial interpretations given before have applications in counting problems in graphs. Indeed, if $A = (a_{ij})$ is the adjacency matrix of a (di)graph $G$, then each nonzero cycle cover has weight 1 and thus the permanent counts the number of vertex-disjoint cycle covers of $G$.

Let $G = (X \cup Y, E)$ be a bipartite graph with $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$. Then its adjacency matrix $A$ can be expressed as $A = \begin{pmatrix} 0 & B \\ B^\top & 0 \end{pmatrix}$ where $B = (b_{ij})$ is the $n \times n$ matrix with $b_{ij} = 1$ if and only if $\{x_i, y_j\} \in E$. It is then clear that $\text{per}(A) = (\text{per}(B))^2$ thus $\text{per}(A)$ is the square of the number of perfect matchings.

In regard to the permanent, for a nonnegative $n \times n$ matrix $A$ with row sums $r_1, r_2, \ldots, r_n$,

$$\text{per}(A) \leq r_1 r_2 \cdots r_n.$$ 

The following bound was proved in (3) and (17).

**Theorem 3** (Minc-Brégman inequality) Let $A$ be an $n \times n$ matrix with $a_{ij} \in \{0, 1\}$ and with row sums $r_1, r_2, \ldots, r_n$, then

$$\text{per}(A) \leq \prod_{i=1}^{n} (r_i!)^{1/r_i}.$$ 

By an application of the arithmetic-geometric mean inequality, this implies the already known $\text{per}(A) \leq \prod_{i=1}^{n} \frac{r_i+1}{2}$ also proved by (17).
The interpretation of permanents counting the weight of perfect matchings in a bipartite graph given in Section 1.1 has formulations useful in counting problems. Below are some examples:

**Example 1** (Worker assignments) The permanent of a $n \times n$ zeros-and-ones matrix $A$ counts the number of possible unique (no two workers assigned to the same job, and no two jobs assigned to the same worker) assignments of $n$ workers to $n$ tasks where $A$ is the matrix with $a_{ij} = 1$ if worker $j$ is qualified for task $i$ and $a_{ij} = 0$ otherwise.

**Example 2** (Derangement problem and ménage problem) The interpretation of the permanent in terms of counting perfect matchings in a bipartite graph for zeros-and-ones matrices can be used in counting the number of permutations with restricted positions. Let $A = (a_{ij})$ be the matrix with $a_{ij} = 1$ if $i \to j$ is allowed in a permutation and $a_{ij} = 0$ otherwise. Then \( \text{per}(A) \) counts the permutations of an $n$-set that satisfy all the restrictions.

The derangement problem asks to count the number of permutations of an $n$-set that leave no fixed element. The solution is then given by \( \text{per}(J - I) \) which allows for the following approximation:

\[
\text{per}(J - I) = n! \left( 1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + (-1)^n \frac{1}{n!} \right) \approx \frac{n!}{e}.
\]

The ménage problem asks for the number of ways that $n$ couples can sit around a circular table, alternating between men and women so that nobody sits next to their spouse. The number of ways to do this is $2n! M_n$ where $M_n$ is called the ménage number.

Similarly as in with the derangement problem, the equivalent solution to the ménage problem is given by $M_n = \text{per}(J - I - E)$ where $J$ is the all-ones matrix, $I$ is the identity matrix, and $E$ is the cyclic permutation matrix with entries $(e_{ij}) = 1$ if $j = i + 1$ mod $n$ and 0 otherwise.
Derangements \((n = 3)\)
\[
\text{per}\left( \begin{array}{ccc} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{array} \right) = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!} = 3!(1 - 1 + \frac{1}{2} - \frac{1}{6}) = 2
\]

Ménage \((n = 3)\)
\[
\text{per}\left( \begin{array}{ccc} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right) = 2(n!) \sum_{k=0}^{n} (-1)^k \frac{2n}{2n-k} \binom{2n-k}{k} (n-k)! = 1
\]

Figure 2.1 Solutions to the derangement and ménage problems using permanents

Permanents and system of distinct representatives. Let \(\mathcal{A} = \{A_1, A_2, \ldots, A_n\}\) be a collection of \(n\) subsets of \([n]\). The incidence matrix \(Q = (q_{ij})\) of \(\mathcal{A}\) is defined by \(q_{ij} = 1\) if \(i \in A_j\) and \(q_{ij} = 0\) otherwise. A system of distinct representatives (SDRs) of \(\mathcal{A}\) is a set of elements \(x_i \in A_i\) such that \(x_i \neq x_j\) for all distinct \(i, j \in [n]\).

**Proposition 1** Let \(A\) be the \(n \times n\) zeros-and-ones matrix whose rows and columns are indexed by the elements of \([n]\) and its \((i, j)\)-entry is one if \(i \in A_j\), otherwise zero. Then the \(\text{per}(A)\) equals to the number of SDRs of \(\mathcal{A}\).

**Proof**: In the evaluation of the permanent of the incidence matrix \(A\), the product corresponding to a permutation \(\pi\) is zero unless \(\pi(i) \in A_i\) for all \(i\), when it is 1. In this case \((\pi(1), \pi(2), \ldots, \pi(n))\) is an SDR for \((A_1, A_2, \ldots, A_n)\). Conversely, any SDR arises from such a permutation.

\(\square\)

**König’s theorem on bipartite graphs.** We first recall that given a graph \(G\), a set of vertices is called a vertex cover if every edge of \(G\) is incident with a vertex in the set.

**Theorem 4** (König (1931), Egerváry (1931)) If \(G\) is a finite bipartite graph, then the maximum size of a matching in \(G\) equals the minimum size of a vertex cover of \(G\).
Proof: Let us use the following parameters.

$$\alpha(G) := \max\{|M| : M \text{ is a matching in } G\};$$

$$\beta(G) := \min\{|W| : W \text{ is a vertex cover of } G\}.$$ 

Then it is clear that $\alpha(G) \leq \beta(G)$; and so, it suffices to show that $\beta(G) \leq \alpha(G)$. Let $X$ and $Y$ be a bipartition of the vertex set $V$ of $G$. We use the induction on $|V|$. If $|V| = 2$, then either $\beta(G) = \alpha(G) = 0$ or $\beta(G) = \alpha(G) = 1$. Without loss of generality, we now assume that $G$ is a graph with $|V(G)| \geq 3$, and consider the following two cases.

Case 1. Suppose there is a minimum cover $Q = R \cup T$ such that $R = Q \cap X \neq \emptyset$ and $T = Q \cap Y \neq \emptyset$. Then there is no edge linking any vertex from $X - R$ to any of $Y - T$ since $R \cup T$ is a cover. Let $H := (R \cup (Y - T), E(H))$ where $E(H)$ is the set of edges of $G$ with both endvertices in $R \cup (Y - T)$. Similarly, we define $H' := (T \cup (X - R), E(H'))$ where $e \in E(H')$ if and only if $e \cap T \neq \emptyset$ and $e \cap (X - R) \neq \emptyset$. Then clearly, $|V(H)| < |V(G)|$ and $|V(H')| < |V(G)|$, where $V(H), V(H')$ and $V(G)$ are the vertex sets of corresponding graphs $H, H'$ and $G$, respectively. Moreover, if $Q_1$ is a minimum cover of $H$, then $Q_1 \cup T$ is a cover of $G$. In this case, we have $|Q_1| + |T| \geq \beta(G) = |R| + |T|$ which implies that $|Q_1| \geq |R|$. Since $R$ is a cover of $H$, this means $|R| = |Q_1|$. By induction hypothesis, $\alpha(H) = \beta(H) = |R| = |Q_1|$. So, there exists a maximum matching $M_H$ of size $|R|$ in $H$. Similarly, we see that there exists a maximum matching $M_{H'}$ of size $|T|$ in $H'$. Now, we have a matching $M_H \cup M_{H'}$ for $G$ with

$$\alpha(G) \geq |M_H \cup M_{H'}| = |M_H| + |M_{H'}| = |R| + |T| = \beta(G).$$

Case 2. Suppose that every minimum cover is either contained in $X$ or contained in $Y$. Choose an edge $e_0$ with $e_0 = \{x, y\}$ for some $x \in X, y \in Y$, and define $H = (V(H), E(H))$ as follows:

$$V(H) := (X - \{x\}) \cup (Y - \{y\}),$$

$$E(H) := E(G) - \{e_0\} - \{e \in E(G) : e \cap \{x, y\} \neq \emptyset\},$$
where every edge of $H$ has the same endpoints as in $G$. Then $H$ is a bipartite graph with $|V(H)| = |V(G)| - 2$.

Since $x, y$ can be added to any minimum cover for $H$ to give a cover for $G$ (this cover is not necessarily a minimum cover of $G$), we have $\beta(G) < \beta(H) + 2$ or $\beta(G) \leq \beta(H) + 1$. Therefore, if $M_H$ is a maximum matching in $H$, then $M = M_H \cup \{e_0\}$ is a matching in $G$. As $|M_H| = \alpha(H) = \beta(H)$ by induction hypothesis, we obtain the inequality $\alpha(G) \geq |M| = |M_H| + 1 = \beta(H) + 1 \geq \beta(G)$ as desired. This completes the proof.

\[\square\]

In a bipartite graph $G = (X \cup Y, E)$, if a matching $M$ saturates $X$; that is, $|M| = |X|$, then for every $S \subseteq X$, there must be at least $|S|$ vertices in $Y$ that have neighbors in $S$, since the vertices matched to $S$ must be chosen from $N(S)$. Thus $|N(S)| \geq |S|$ is a necessary condition. The condition “For every $S \subseteq X$, $|N(S)| \geq |S|$” is called Hall’s condition. Hall proved (in what is known as Hall’s Marriage Theorem) that this obvious necessary condition is also sufficient in general.

In a bipartite graph $G = (X \cup Y, E)$ with $|X| = |Y| = n$, if a matching $M$ saturates $X$, this matching is a perfect matching. In this case, the set of endvertices of the edges in the matching in $Y$ gives an SDR for the family $(N(x_1), N(x_2), \ldots, N(x_n))$ of the neighborhoods of $x_1, x_2, \ldots, x_n \in X$.

### 2.2 Latin squares

We define a $k \times n$ Latin rectangle for $k \leq n$, to be a $k \times n$ array with entries from a set $[n] = \{1, 2, \ldots, n\}$ with the property that each entry occurs exactly once in each row and at most once in each column.

Equivalently, a map $L : [k] \times [n] \to [n]$ is called a $k \times n$ Latin rectangle on $[n]$ if (i) for each $i, 1 \leq i \leq k$, $L(i, 1), L(i, 2), \ldots, L(i, n)$ are all distinct, and (ii) for each $j, 1 \leq j \leq n$, $L(1, j), L(2, j), \ldots, L(k, j)$ are all distinct. If $k = n$, we call it a Latin square.

We note that
(a) Given a $k \times n$ Latin rectangle with $k < n$, there are at least $(n - k)!$ ways to add a row to form a $(k + 1) \times n$ Latin rectangle.

(b) The number of Latin squares of order $n$, denoted $L(n)$, is at least $\prod_{k=1}^{n} k!$.

Figure 2.2 Latin square examples

Figure 2.2 shows two Latin squares. The second can be interpreted as a Latin square in two ways, choosing colors or chess pieces for symbols. As it will be shown in the following section, the bound given above for extending a Latin rectangle can be improved on with the help of permanents and the result of van der Waerden.

2.3 Doubly stochastic matrices

A doubly stochastic matrix $A = (a_{ij})$ is an $n \times n$ matrix with nonnegative entries for which $\sum_{j=1}^{n} a_{ij} = 1$ and $\sum_{i=1}^{n} a_{ij} = 1$. By the following theorem of von Neumann and Birkhoff, it is well known that every doubly stochastic matrix can be expressed as a convex combination of permutation matrices.

Theorem 5 (von Neumann, Birkhoff (1946)). Let $A$ be an $n \times n$ nonnegative matrix such that all row sums and column sums are equal. Then there exist positive real numbers $\alpha_1, \alpha_2, \ldots, \alpha_m$, and permutation matrices $P_1, P_2, \ldots, P_m$ such that $A = \sum_{i=1}^{m} \alpha_i P_i$.

One proof of this theorem relies on an important result, the ‘Marriage Theorem’ by P. Hall in 1935 and a theorem on matchings and coverings by König and Egerváry. We give a sketch of the proof to demonstrate the connections to the theory of matchings and SDRs.
**Proof:** We use strong induction on the number of positive entries of \( A \). Let each row sum be \( s > 0 \), and let

\[
p(A) := |\{(i, j) : a_{ij} > 0\}|.
\]

Then a nontrivial \( n \times n \) matrix \( A \) with equal row sums and column sums has at least \( n \) and at most \( n^2 \) positive entries. When \( p(A) = n \), it is clear that each row and each column has single positive entry, in which case the entry must be \( s \); and so, \( A = sP \) for a permutation matrix \( P = (p_{ij}) \) defined by \( p_{ij} = 1 \) if and only if \( a_{ij} = s \) and \( p_{ij} = 0 \) otherwise.

Now let \( A \) be a matrix with \( p(A) = k \) for any \( k \) (\( n < k \leq n^2 \)). We assume that the theorem is valid for all matrices having less than \( k \) positive entries. Let \( \mathcal{G} = (\mathcal{R} \cup \mathcal{C}, E) \) be the bipartite graph defined by: \( \mathcal{R} = \{r_1, r_2, \ldots, r_n\} \), the set of row indices; \( \mathcal{C} = \{c_1, c_2, \ldots, c_n\} \), the set of column indices for \( A \); and \( \{r_i, c_j\} \in E \) if and only if \( a_{ij} > 0 \). By König’s Theorem, we know that the maximum matching number, denoted by \( \alpha(G) \) equals the minimum covering number, denoted by \( \beta(G) \). Let \( \mathcal{R}' \cup \mathcal{C}' \) be a minimum cover of \( \mathcal{G} \) where \( \mathcal{R}' \subseteq \mathcal{R} \) and \( \mathcal{C}' \subseteq \mathcal{C} \). Let \( s \) be the common column sum. Then \( |\mathcal{R}'|s + |\mathcal{C}'|s \geq \sum_{i,j} a_{ij} = ns \), or equivalently \( |\mathcal{R}' \cup \mathcal{C}'| \geq n \). Hence, \( \alpha(G) = n \), and \( \mathcal{G} \) has a perfect matching \( M = \{(r_i, c_{\sigma(i)}) : 1 \leq i \leq n\} \) for some \( \sigma \in S_n \). That is, there exists a permutation \( \sigma : [n] \to [n] \) such that \( a_{i\sigma(i)} > 0 \) for all \( i, 1 \leq i \leq n \). Let \( \alpha_1 = \min\{a_{i\sigma(i)} : 1 \leq i \leq n\} \), and \( P_1 = (p_{ij}) \) where \( p_{ij} = 1 \) if \( j = \sigma(i) \) and \( p_{ij} = 0 \) otherwise. Then \( A' := A - \alpha_1 P_1 \) has equal row and column sums, \( s - \alpha_1 \), and entries are nonnegative with \( p(A') < p(A) \). So by induction, \( A' \) can be written as a linear combination of permutation matrices and so does \( A \).

\[ \square \]

Therefore, a doubly stochastic matrix is a convex combination of permutation matrices.

**Corollary 1** The matrix \( A \) is an \( n \times n \) doubly stochastic matrix if and only if there exist nonnegative numbers \( t_\pi \) such that

\[
A = \sum_{\pi \in S_n} t_\pi P_\pi \quad \text{and} \quad \sum_{\pi \in S_n} t_\pi = 1.
\]
We refer to the nonnegative numbers $t_\pi$ as the coefficients of the decomposition. Such a decomposition may not be unique and the number of positive coefficients is less than $n!$ in general. The fact that there is at least one constructible decomposition with no more than $(n - 1)^2 + 1$ permutation matrices was known by Marcus and Ree (16). \(^1\)

Let $\Omega_n$ denote the set of doubly stochastic matrices of size $n$.

**Proposition 2** If $A \in \Omega_n$ then $\text{per}(A) \leq 1$, with equality if and only if $A$ is a permutation matrix.

Permanents of doubly stochastic matrices have applications in problems arising in statistical physics such as the following: Suppose $n$ unlabeled identical objects or atoms are placed (bijectively) into $n$ labeled containers or positions in a crystalline lattice. If after some process the probability that the item in the $i$th container has moved to the $j$th container with probability $P = (p_{ij})$. Then $P \in \Omega_n$ and $\text{per}(P)$ is the probability that we end up with one ball in each bucket (referred to as the permanence of the initial state in (8)).

In 1926, B. L. van der Waerden asked whether the minimum permanent among all square doubly stochastic matrices was $\frac{n!}{n^n}$. This conjecture remained open until proved in 1980 by G. P. Egorychev and was also proved independently by D. I. Falikman in 1981, and is the focus of the next chapter.

**Conjecture 6** (van der Waerden, 1926) The permanent of a doubly stochastic matrix $A$ is greater than equal to $\frac{n!}{n^n}$, and equality holds if and only if $A = \frac{1}{n}J$ where $J$ is the $n \times n$ all-ones matrix.

\(^1\)The problem of computing the representation with minimum number of summands was shown to be NP-hard problem, but some heuristics for computing it are known. See (7).
CHAPTER 3. Proof of van der Waerden’s Permanent Conjecture

We now specifically turn our attention to the well-known proof of van der Waerden’s permanent conjecture by Egorychev (9). We will discuss some key ingredients of Egorychev’s proof by following the lines of van Lint and Knuth’s expositions of the proof (cf. (12; 23; 24)).

Egorychev’s proof is based on the observation that the permanent is a continuous function defined on the set \( \Omega_n \) of doubly stochastic matrices and \( \Omega_n \) is a closed and bounded subset of the set \( M_n(\mathbb{R}) \) of \( n \times n \) matrices with real entries, as an \( n^2 \)-dimensional space over the reals. By von Neumann and Birkhoff’s Theorem 5, for \( A \in \Omega_n \),

\[
\per(A) = \sum_{\pi \in S_n} \prod_{i=1}^{n} a_{i\pi(i)} \geq \sum_{\pi \in S_n} \prod_{i=1}^{n} t_{\pi(i)} = \sum_{\pi \in S_n} t_{\pi}^n
\]

for some nonnegative numbers \( t_{\pi} \) with \( \sum_{\pi \in S_n} t_{\pi} = 1 \). So the minimal permanent cannot be zero. Then, results associated with permanents of submatrices of a doubly stochastic matrix with minimal permanent force all entries of the minimal matrix to be \( 1/n \). According to the van der Waerden’s conjecture, there is only one minimal element. We will discuss the key ingredients for this line of the proof.

3.1 Permanents and quadratic forms

To build up to the proof of the main theorem, we introduce some preliminary results by diverting our attention to properties of quadratic forms.

A quadratic form \( f(x_1, \ldots, x_n) \) of \( n \) variables is an expression of the form

\[
f(x_1, \ldots, x_n) = \sum_{1 \leq i, j \leq n} f_{ij} x_i x_j
\]
defined by some \( n \times n \) matrix \( F = (f_{ij}) \). The matrix used is not necessarily unique. Take for instance the matrices below

\[
F_1 = \begin{pmatrix} 1 & 0 & -2 & -2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ -1 & -1 & 0 & 1 \end{pmatrix}
\]

both of which correspond to the quadratic form

\[
f(x) = f(x_1, x_2, x_3, x_4) = x_1^2 + x_2^2 - 2x_1x_3 + x_3^2 - 2x_1x_4 - 2x_2x_4 + x_4^2.
\]

In general, the coefficient of the term \( x_i x_j \) in the quadratic form \( f(x_1, \ldots, x_n) \) is \( f_{ij} + f_{ji} \) for \( i \neq j \) so that we may assume without loss of generality that the matrix is symmetric about the diagonal with \( f_{ij} = f_{ji} = \frac{f_{ij} + f_{ji}}{2} \).

Let \( x = Ty \) for a linear transformation \( T \) and a (column) vector \( y \) of variables. Then note that with \( G = T^\top F T \)

\[
f(x) = x^\top F x = y^\top T^\top F T y = y^\top G y = g(y).
\]

This means we may turn a quadratic form into another one by means of a linear transformation of the variables. The following lemma illustrates a way to do this.

**Lemma 1** Let \( F \in M_n(\mathbb{R}) \) and let \( f \) be the quadratic form defined by \( f(x) = x^\top F x = \sum_{i,j} f_{ij} x_i x_j \) for \( x \in \mathbb{R}^n \). Let the vector \( a = (a_1, a_2, \ldots, a_n)^T \in \mathbb{R}^n \) be such that \( a_1 \neq 0 \) and \( f(a) = c \neq 0 \). Let \( T \in M_n(\mathbb{R}) \) be the nonsingular matrix

\[
T = \begin{bmatrix}
  a_1 & -a_1 \sum_{k=1}^n f_{2k} a_k / c & \cdots & -a_1 \sum_{k=1}^n f_{jk} a_k / c & \cdots & -a_1 \sum_{k=1}^n f_{nk} a_k / c \\
  a_2 & 1 - a_2 \sum_{k=1}^n f_{2k} a_k / c & \cdots & -a_2 \sum_{k=1}^n f_{jk} a_k / c & \cdots & -a_2 \sum_{k=1}^n f_{nk} a_k / c \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  a_j & -a_j \sum_{k=1}^n f_{2k} a_k / c & \cdots & 1 - a_j \sum_{k=1}^n f_{jk} a_k / c & \cdots & -a_j \sum_{k=1}^n f_{nk} a_k / c \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  a_n & -a_n \sum_{k=1}^n f_{2k} a_k / c & \cdots & -a_n \sum_{k=1}^n f_{jk} a_k / c & \cdots & 1 - a_n \sum_{k=1}^n f_{nk} a_k / c
\end{bmatrix}.
\]
Then the transformation defined by \( x = Ty \) makes \( f(x) = cy_1^2 + g(y_2, y_3, \ldots, y_n) \), where \( g \) is a quadratic form in \( \mathbb{R}^{n-1} \) in \( n - 1 \) variables.

In this case, \( T^{-1} \) is given by

\[
T^{-1} = \begin{bmatrix}
\sum_{k=1}^{n} f_{1k}a_k/c & \sum_{k=1}^{n} f_{2k}a_k/c & \cdots & \sum_{k=1}^{n} f_{jk}a_k/c & \cdots & \sum_{k=1}^{n} f_{nk}a_k/c \\
-a_2/a_1 & 0 & \cdots & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & \cdots & 0 \\
0 & \cdots & 0 & -a_j/a_1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & 0 & \cdots & -a_n/a_1
\end{bmatrix}.
\]

If the matrix corresponding to the linear transformation is invertible, we call two such quadratic forms \( f \) and \( g \) as equivalent. Since we may write \( x^\top Fx \) or \( z^\top Gz \) interchangeably by means of a suitable transformation so that the quadratic forms are equivalent, one may ask if there is a convenient choice of \( z \) and \( G \) (change of variables). To answer this, recall that a real symmetric matrix can be diagonalized by a real orthogonal matrix, therefore the quadratic form may be written as \( y^\top \Lambda y = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \) where \( \Lambda \) is a diagonal matrix after applying some real orthogonal transformation \( y = Ux \) and \( \lambda_i \) are the \( \pm \) of eigenvalues of \( F \) for \( i = 1, 2, \ldots, n \). As a consequence, every quadratic form in \( n \) variables can be reduced to the form

\[ c_1 x_1^2 + \cdots + c_n x_n^2 \]

by a nonsingular linear transformation. Furthermore, we may rescale by yet another linear transformation \( z = Ry \) with matrix \( R = diag \left( |\lambda_1|^{-\frac{1}{2}}, \ldots, |\lambda_n|^{-\frac{1}{2}} \right) \), effectively normalizing the coefficients \( \lambda_i \).

The above guarantees existence of a convenient form to write any quadratic form after some linear transformations. The Sylvester’s Law of Inertia theorem assures us that this representation is unique for real symmetric matrices. That is, two real symmetric matrices are congruent if and only if they have the same number of positive eigenvalues and the same number of negative eigenvalues. A longer discussion of this can be found in (10).
3.2 Key inequalities in permanents

In this section we will follow the lines of proof by van Lint (23) for a particular inequality for permanents originally due to Aleksandrov (1).

**Theorem 7** Let $a_1, a_2, \ldots, a_{n-1}$ be row vectors in $\mathbb{R}^n$ with positive coordinates and $b \in \mathbb{R}^n$ be an arbitrary vector. Then

$$(\text{per}(a_1, \ldots, a_{n-1}, b))^2 \geq \text{per}(a_1, \ldots, a_{n-1}, a_{n-1}) \cdot \text{per}(a_1, \ldots, a_{n-2}, b, b)$$

and equality holds if and only if $b = \lambda a_{n-1}$ for some constant $\lambda$.

J. H. van Lint’s proof uses the following notion of a Lorentz space. The space $\mathbb{R}^n$ is called a Lorentz space if a symmetric inner product $\langle x, y \rangle = x^t Q y$ has been defined such that $Q$ has one positive eigenvalue and $n-1$ negative eigenvalues. In a Lorentz space a vector $x$ is called positive (negative) if $\langle x, x \rangle$ is positive (negative); and isotropic if $\langle x, x \rangle = 0$.

**Lemma 2** If $a$ is a positive vector in a Lorentz space and $b$ is arbitrary then

$$\langle a, b \rangle^2 \geq \langle a, a \rangle \cdot \langle b, b \rangle$$

and equality holds if and only if $b = \lambda a$ for some constant $\lambda$.

**Proof**: We observe that by Sylvester’s law of inertia there is no 2-dimensional subspace of a Lorentz space containing a positive vector. As a consequence, if $b$ is not a multiple of $a$ then the plane spanned by $a$ and $b$ contains an isotropic vector and a negative vector. If we consider the inner product

$$\langle a + \lambda b, a + \lambda b \rangle = \langle b, b \rangle \lambda^2 + 2\langle a, b \rangle \lambda + \langle a, a \rangle$$

as a quadratic form in variable $\lambda$, since this form is zero respectively negative for suitable values of $\lambda$, it must have a positive discriminant.

$\square$
We now define an inner product on \( \mathbb{R}^n \) for given vectors \( a_1, a_2, \ldots, a_{n-2} \in \mathbb{R}^n \) with positive coordinates, by

\[
\langle x, y \rangle := \text{per}(a_1, \ldots, a_{n-2}, x, y),
\]

or equivalently,

\[
\langle x, y \rangle := x^tQy,
\]

where \( Q = (q_{ij}) \) is given by

\[
q_{ij} := \text{per}(a_1, \ldots, a_{n-2}, e_i, e_j) = \text{per}(A[\{n-1, n\}, \{i, j\}]),
\]

where \( A \) is a matrix with rows \( a_1, \ldots, a_n \). With this inner product, \( \mathbb{R}^n \) is a Lorentz space that suffices to assert Theorem 7 because of Lemma 2.

**Lemma 3** \( \mathbb{R}^n \) with the inner product given by Equation (3.1) is a Lorentz space.

**Proof**: We use induction on \( n \). For \( n = 2 \), we have \( Q = J - I \) and the statement follows. Now assume that the statement is true for \( \mathbb{R}^{n-1} \). We shall show that \( Q \) does not have the eigenvalue zero. Suppose \( Qc = 0 \) for some vector \( c \in \mathbb{R}^n \). Then

\[
\text{per}(a_1, \ldots, a_{n-2}, c, e_i) = 0 \quad \text{for} \quad 1 \leq i \leq n.
\]

Consider the \( (n - 1) \times (n - 1) \) matrices \( A_{ni} \) where

\[
A = (a_1, \ldots, a_{n-3}, x, y, e_i)
\]

and apply the induction hypothesis and Lemma 2. Since \( a_{n-2} \) has positive coordinates, it follows from (2) that for \( 1 \leq i \leq n \),

\[
\text{per}(a_1, \ldots, a_{n-3}, c, c, e_i) \leq 0
\]

and for each \( i \) equality holds if and only if all coordinates of \( c \) except \( c_i \) are zero. The assumption \( Qc = 0 \) therefore implies that \( c = 0 \).

Let \( j = (1, 1, \ldots, 1) \). We consider the inner product \( x^tQ_\theta y \) defined by

\[
x^tQ_\theta y := \text{per}((1 - \theta)j + \theta a_1, \ldots, (1 - \theta)j + \theta a_{n-2}, x, y).
\]
For every $\theta \in [0,1]$, this satisfies the condition of the theorem. Hence $Q_\theta$ does not have the eigenvalue 0. Hence $Q = Q_1$ has the same number of positive eigenvalues as $Q_0$ and since $Q_0$ is a multiple of $J - I$ this number is 1.

$\Box$

The following result by London on the minimal matrix in $\Omega_n$ is of relevance in the rest.

**Theorem 8** If $A \in \Omega_n$ is a minimal matrix, then

$$\text{per}(A_{ij}) \geq \text{per}(A) \quad \text{for any } i, j \in [n].$$

**Proof:** For the proof, we will refer the reader to London (13) and Minc (20). We just remark that a proof can be deduced from a number of facts on minimal matrices, most of which are easily accessible from the mentioned references. We will state them without proof.

(a) If $A \in \Omega_n$, then $0 < \text{per}(A) \leq 1$. The equality holds if and only if $A$ is a permutation matrix.

(b) If $A \in \Omega$ is a minimal matrix, then $A$ is fully indecomposable; that is, no $s \times (n - s)$ submatrix of $A$ for any $s = 1, 2, \ldots, n - 1$ is a zero matrix.

(c) If $A = (a_{ij}) \in \Omega$ is a minimal matrix and $a_{kl} > 0$, then $\text{per}(A_{kl}) = \text{per}(A)$.

$\Box$

### 3.3 Egorychev’s proof of van der Waerden’s conjecture

Now we can put all the ingredients for the proof of van der Waerden’s conjecture to make it a theorem by following Egorychev’s proof. The proof is based on two results that we discussed in the previous section. We recall them here:

(a) (Theorem 8) The permanent of a minimal matrix in $\Omega_n$ cannot exceed the permanent of any of its $(n - 1) \times (n - 1)$ submatrix.

Many of the ideas of the component including the details of this result are due to Marcus and Newman (15) as well as Minc (17; 18; 19; 20).
(b) (Theorem 7) If $A \in \Omega_n$ is with rows $a_1, a_2, \ldots, a_{n-1}, a_n$, all of which except possibly one ($a_n$, say) are nonnegative, then

$$(\text{per}(A))^2 \geq \text{per}(a_1, a_2, \ldots, a_{n-2}, a_{n-1}) \cdot \text{per}(a_1, a_2, \ldots, a_{n-2}, a_n, a_n).$$

If $a_1, a_2, \ldots, a_{n-1}$ are positive, then the equality holds if and only if $a_{n-1} = a_n$. An inequality analogous to this (which uses $a_{n-1}$ and $a_n$), together with the condition for equality, holds for any pair of rows of $A$.

Egorychev first proved the following key result: All the permanents of $(n-1) \times (n-1)$ submatrix of an $n \times n$ minimal matrix are equal to the permanent of the matrix.

**Theorem 9** If $A \in \Omega_n$ is minimal, then $\text{per}(A_{ij}) = \text{per}(A)$ for any $i, j \in [n]$.

**Proof:** Let $A = (a_{ij})$ be a minimal matrix in $\Omega_n$, and let $A_{ij}$ denote the submatrix of $A$ obtained by deleting its $i$th row and $j$th column. By Theorem 8, $\text{per}(A_{ij}) \geq \text{per}(A)$ for all $i$ and $j$. Suppose that for some $k$ and $l$ the inequality is strict, $\text{per}(A_{kl}) > \text{per}(A)$. Let $a_{hl}$ be a positive entry in the $l$th column of $A$, $h \neq k$. Then, by the inequality we would have

$$(\text{per}(A))^2 \geq \text{per}(a_1, \ldots, a_h, \ldots, a_{k-1}, a_h, a_{k+1}, \ldots, a_n) \cdot \text{per}(a_1, \ldots, a_{h-1}, a_k, a_{h+1}, \ldots, a_k, \ldots, a_n)$$

$$= \left( \sum_{j=1}^{n} a_{hj} \text{per}(A_{kj}) \right) \left( \sum_{j=1}^{n} a_{kj} \text{per}(A_{hj}) \right)$$

$$> (\text{per}(A))^2,$$

since

$$a_{hj} \text{per}(A_{kj}) \geq a_{hj} \text{per}(A), \quad a_{kj} \text{per}(A_{hj}) \geq a_{kj} \text{per}(A),$$

for $j = 1, \ldots, n$, and $a_{hl} \text{per}(A_{kl}) > a_{kl} \text{per}(A)$.

$\Box$

Theorem 9 implies the following.

**Lemma 4** Let $A = (a_1, a_2, \ldots, a_n) \in \Omega_n$ be a minimal matrix. Then the matrix $B$ obtained by replacing both $a_i$ and $a_j$ by $\frac{1}{2}(a_i + a_j)$, is also minimal in $\Omega_n$; and thus, Theorem 9 applies to $B$ as well.
Proof: It is obvious that $B \in \Omega_n$. Since $\text{per}(B) = \sum_{l=1}^{n} b_{kl} B_{kl}$ for any $k \in [n]$, and Theorem 9 we have

$$
\text{per}(B) = \text{per}(a_1, \ldots, a_{i-1}, \frac{1}{2}(a_i + a_j), a_{i+1}, \ldots, a_{j-1}, \frac{1}{2}(a_i + a_j), a_{j+1}, \ldots, a_n) \\
= \frac{1}{4} \{ \text{per}(a_1, \ldots, a_i, \ldots, a_i, \ldots, a_n) + \text{per}(a_1, \ldots, a_i, \ldots, a_{j-1}, a_j, \ldots, a_n) + \text{per}(a_1, \ldots, a_j, \ldots, a_j, \ldots, a_n) \} \\
= \frac{1}{4} \left\{ \sum_{l=1}^{n} a_{il} \text{per}(A_{jl}) + \text{per}(A) + \text{per}(A) + \sum_{l=1}^{n} a_{jl} \text{per}(A) \right\} \\
= \frac{1}{4} \left\{ \sum_{l=1}^{n} a_{il} \text{per}(A) + \text{per}(A) + \text{per}(A) + \sum_{l=1}^{n} a_{jl} \text{per}(A) \right\} \\
= \text{per}(A).
$$

□

Application of Lemma 4 together with the condition for equality in Aleksandrov’s inequality, Egorychev proved the following.

**Theorem 10** If $S \in \Omega_n$, then $\text{per}(S) \geq \frac{n!}{n^n}$, where equality holds if and only if $S = \frac{1}{n} J_n$.

Proof: Let $a_i$ be an arbitrary row of an $n \times n$ minimal matrix $A$. By suitably averaging the other rows a finite number of times we obtain a minimal matrix $B$ whose $i$th row is $a_i$ and whose other rows $b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n$, are positive.

$$
\text{per}(b_1, \ldots, b_j, \ldots, b_{i-1}, b_j, b_{i+1}, \ldots, b_n) \cdot \text{per}(b_1, \ldots, b_{i-1}, a_i, b_{i+1}, \ldots, a_i, \ldots, b_n) = (\text{per}(B))^2
$$

for $j = 1, \ldots, i - 1, i + 1, \ldots, n$, and therefore $b_j = a_i$ for $j = 1, \ldots, i - 1, i + 1, \ldots, n$. Thus all the entries in $a_i$ are $1/n$ and, since $a_i$ was an arbitrary row of $A$, it follows that $A = \frac{1}{n} J_n$. This proves the van der Waerden’s conjecture, becoming Egorycev’s theorem.

□

### 3.4 Applications of van der Waerden’s result

We return to the remark made in Section 2.2 on how this result can be used in bounding the number of ways to extend a Latin rectangle.
Let $k < n$ and let $L$ be a $k \times n$ Latin rectangle on symbols $\{1, \ldots, n\}$. To count the number of ways to extend $L$ to a $(k + 1) \times n$ Latin rectangle, let $B = (b_{ij})$ be the matrix with $b_{ij} = 1$ if $i$ does not appear in column $j$ of the rectangle and $b_{ij} = 0$ otherwise. Then $\text{per}(B)$ counts the number of ways to extend $L$. The row and column sums of matrix $B$ are $n - k$, so $\frac{1}{n-k}B$ is doubly stochastic. By van der Waerden’s result, this permanent is at least $(n - k)^n \frac{n!}{n^n}$ so we have

$$L(n) \geq n! \prod_{k=1}^{n-1} \left\{ (n-k)^n \frac{n!}{n^n} \right\} = \frac{(n!)^{2n}}{n^{n^2}}.$$
CHAPTER 4. Interesting Open Problems in Related Subjects

We find the following open problems interesting.

1. Bapat and Brualdi proposed the following problem on positive semidefinite matrices (PSD) that may be viewed as an analog of van der Waerden’s conjecture. Their problem posted in (2) introduced the question of the minimum for the mixed discriminant over the larger set

\[ \mathcal{D} = \{(A_1, \ldots, A_n) : A_i \text{ is PSD } \forall i, \sum_{i=1}^{n} A_i = I\}. \]

2. A recent result of (6) found a negative answer to a conjecture made by Drury, but the following question originally raised by (5) remains open.

**Conjecture 11** Let \(A\) and \(B\) be \(n \times n\) positive semi-definite matrices. Then

\[ \text{per}(A \circ B) \leq (\text{per}A)(\text{per}B) \]

where \(A \circ B = (a_{ij}b_{ij})\).

3. The Dittert-Hajec conjecture asserts that \(\frac{1}{n}J\) is the maximum of the function

\[ \prod_{i=1}^{n} \left( \sum_{j=1}^{n} a_{ij} \right) + \prod_{j=1}^{n} \left( \sum_{i=1}^{n} a_{ij} \right) - \text{per}(A) \]

over the set of \(n \times n\) nonnegative matrices.

4. A number of conjectures originally listed in the monographs by Minc remain open.

We refer the interested reader to the bibliography.
Bibliography


APPENDIX A. Additional Material

SAGE Code

# testing some properties of permanent
def repc(A, i, a): # replaces i-th COLUMN of A with vector a
    New=A
    entries=A.dimensions()[0];
    if entries != len(a):
        return 0
    else:
        for k in range(entries):
            New[k, i]=a[k]
    return New
def repr(A, i, a): # replaces i-th ROW of A with vector a
    New=A
    entries=A.dimensions()[1];
    if entries != len(a):
        return 0
    else:
        for k in range(entries):
            New[i, k]=a[k]
    return New
# \( \text{per}(A)^2 < \text{per}(A') \cdot \text{per}(A'') \)

\[
\text{GP} = \text{graphs.PetersenGraph()}
\]

\[
\text{AP} = \text{GP}.\text{adjacency_matrix}()
\]

\[
\text{c9} = [0, 0, 0, 1, 0, 1, 1, 0, 0] \quad \# \text{[9]th row/column of this matrix}
\]

\[
\text{c8} = [0, 0, 0, 1, 0, 1, 1, 0, 0, 0] \quad \# \text{[8]th row/column of this matrix}
\]

\[
\text{AP.\text{permanent}()} 
\]

\[
\text{repc}(\text{AP}, 8, \text{c9}).\text{permanent}()
\]

\[
\text{repc}(\text{AP}, 9, \text{c9}).\text{permanent}()
\]

# symbolic matrix for computations

\[
x1, x2, x3, x4, x5, x6, x7, x8=\text{var('x1 x2 x3 x4 x5 x6 x7 x8')}
\]

\[
x9, x10, x11, x12, x13, x14, x15, x16=\text{var('x9 x10 x11 x12 x13 x14 x15 x16')}
\]

\[
\text{symbolic=matrix([[x1, x2, x3, x4], [x5, x6, x7, x8], [x9, x10, x11, x12], [x13, x14, x15, x16]])}
\]

\[
\text{show(symbolic)}
\]

# two algorithms for computing the permanent

\[
\text{symbolic.\text{permanent}(algorithm='Ryser')} \quad \# \text{default}
\]

\[
\text{symbolic.\text{permanent}(algorithm='ButeraPernici')} \quad \# \text{for sparse matrices}
\]