Permanents and determinants with generic noncommuting entries

Joy P. Fillmore & S. Gill Williamson

To cite this article: Joy P. Fillmore & S. Gill Williamson (1986) Permanents and determinants with generic noncommuting entries, Linear and Multilinear Algebra, 19:4, 321-334, DOI: 10.1080/03081088608817727

To link to this article: http://dx.doi.org/10.1080/03081088608817727

Published online: 30 May 2007.

Submit your article to this journal

Article views: 22

View related articles

Citing articles: 2 View citing articles
Permanents and Determinants with Generic Noncommuting Entries

JOY P. FILLMORE and S. GILL WILLIAMSON
University of California, San Diego, La Jolla, California 92093

(Received August 12, 1985)

This paper studies some basic combinatorial properties of matrix functions of generic matrices. A generic matrix is one with entries from a free associative algebra, over a field, and on a finite set of non-commuting variables (i.e., a tensor algebra). The principal tools are shuffle products. Generic column and row permanents are defined and analogs of the Laplace and Cauchy-Binet theorems are derived in terms of shuffles. In this setting, the generic permanents include as special cases all of the classical matrix functions: Schur matrix functions, determinants, and permanents. 1980 Mathematics Classification 05.15.

Keywords: Shuffle product, generic matrix functions, minor expansions, Laplace Expansion Theorem, Cauchy-Binet Theorem, permanents, determinants, tensor algebra, matrices with non-commuting entries.

1. INTRODUCTION

We record here some results in combinatorial matrix theory. These results are motivated by studies [1, 8], where a need for computational techniques in certain areas of linear and multilinear algebra arose that were not a consequence of the classical framework. In particular, we deal here with determinants of matrices with noncommuting entries such as arise classically from linear algebra over general rings. The usual determinants of such matrices necessarily take on values which commute and are multilinear functions of both rows and columns. In many applications, however, these determinants do more than just provide a number. They serve to organize and provide essentially combinatorial insights into calculations. It is this latter aspect that admits a generalization to the noncommuting case by not demanding all the usual properties of determinants. In this paper, a generalization of the determinant is described. It is remarkable that the appropriate generalizations of the classical identities of Laplace and Cauchy–Binet
remain valid and have even simpler proofs. These results, which we describe in detail in this paper, lend insight into the classical case by bringing combinatorial aspects to the forefront and are, in this sense, closely related to the approach of Marcus [6].

The key to defining something like a determinant with noncommuting entries is the need to be able to multiply noncommuting monomials and have the variables arranged in a specified order. This is done by means of a shuffle product. When this product is used on generic variables, the positions of the variables themselves record how they were multiplied. With this bookkeeping, permanents of generic matrices are defined and their Laplace expansion theorems obtained in Section 2. Since generic variables keep track of their multiplication, the usual signs that are introduced in passing from permanents to determinants can be replaced by arbitrary representations, leading to Laplace expansion theorems for generalizations of the Schur matrix functions [7, 8]. This is carried out in Section 3. In Section 4, analogues of the Cauchy–Binet theorem are obtained in this setting.

Shuffle products and related ideas occur in a similar view in ring theory, algebraic topology, in the theory of algorithms, to mention a few, and combinatorics. The latter touches most closely the viewpoint of this paper [2, 3, 4, 5].

2. GENERIC LAPLACE EXPANSIONS

Let $W$ be a finite set. We consider $W$ as noncommuting variables in the free associative $k$-algebra $k\langle W \rangle$ over the field $k$. If $k^W$ denotes the free vector space with $W$ as basis, then $k\langle W \rangle$ may be identified with the full tensor algebra $\otimes k^W$; the monomial $a_1 a_2 \cdots a_p$ of degree $p$ is identified with the homogeneous tensor $a_1 \otimes a_2 \otimes \cdots \otimes a_p$.

Let $m$ be the ordered set $\{1, 2, \ldots, m\}$ of integers. For an (ordered) subset $I \subseteq m$, let $a_I$ be the monomial $\prod_{i \in I} a_i$, where the order of the variables is that of $I$. For $I \subseteq m, \bar{I} = m - I$ is the complement of $I$ with its order as subset.

2.1. Definition (Shuffle product) Let $a = a_1 \cdots a_m$ and $b = b_1 \cdots b_n$ be two monomials in $W$. Let $I \subseteq m + n$ have cardinality $|I| = m$, so $\bar{I} = m + n - I$ has cardinality $n$. Define the $I$-shuffle $a[I] b$ of $a$ and $b$ to be $c = c_1 c_2 \cdots c_{m+n}$ where $c_I = a$ and $c_{\bar{I}} = b$. The shuffle product $a \shuffle b$ of $a$ and $b$ is the sum of all $I$-shuffles $a[I] b$ over all $I \subseteq m + n$ with $|I| = m$. 
2.2. Example Take \( a = a_1 a_2, \ b = b_1 b_2 b_3, \) and \( I = \{2, 4\}. \) Then 
\[ c = c_1 c_2 c_3 c_4 c_5 \] gives that 
\[ c_I = c_2 c_4 \] is 
\[ a_1 a_2 \] and 
\[ c_{\bar{I}} = c_4 c_5 \] is 
\[ b_1 b_2 b_3, \] so 
\[ a \mid b = b_1 a_1 b_2 a_2 b_3. \] And 
\[ a \sqcup b = a_1 a_2 b_1 b_2 b_3 + a_1 b_1 a_2 b_2 b_3 + a_1 b_2 a_3 b_3 + a_1 b_1 a_2 b_2 a_3 + b_1 a_1 b_2 b_3 + b_1 a_1 b_2 a_3 b_3 + b_1 a_2 b_3 a_3 + b_1 b_2 a_3 b_3 + b_1 b_2 a_3 a_3, \]
\[ c = c_1 c_2 c_3 c_4 c_5 \] gives that 
\[ c_I = c_1 c_2 c_5 \] is 
\[ a_1 a_3 \] and 
\[ c_{\bar{I}} = c_3 c_5 \] is 
\[ b_1 b_3, \] so 
\[ u \sqcup h = b, u, b, a, b, a, b, \]
\[ a \sqcup b = a, \bar{a}, b, h, b, a, \]
\[ + a_1 h_1 h_2 b_2 h_3 + a_1 b_1 h_2 b_3 a_3 + b_1 a_1 h_2 a_3 b_3 + b_1 a_1 h_2 a_3 b_3 + b_1 a_1 h_2 a_3 b_3 + b_1 a_1 h_2 a_3 b_3 + b_1 b_1 h_2 a_3 b_3 + b_1 b_1 h_2 a_3 a_3. \]

2.3. Remark The vector space underlying the algebra \( k\langle W \rangle \) together with the shuffle product \( \sqcup \) is an associative algebra called the shuffle algebra of \( W. \) It is generated by all finite shuffle products \( a_1 \sqcup a_2 \sqcup \cdots \sqcup a_r, a_i \in W, \) and is isomorphic to the full symmetric algebra \( S(k^m) \) as well as the algebra \( k[W] \) of commuting polynomials in the variables in the set \( W. \)

The classical determinant expansion theorems of Laplace and Cauchy-Binet do not depend on the entries of the determinant being commutative provided one has the correct notion for the determinant for noncommuting variables and a means of keeping track of the variables. This will be achieved by introducing generic matrix functions and specializing carefully at the appropriate times.

2.4. Definition (Generic permanents) Let \( A = (a_{ij}) \) be an \( m \times m \) matrix with entries in \( W. \) Define 
\[ \text{PER}_r(A) = \sum_{\sigma \in S_m} \prod_{i=1}^m a_{i,\sigma(i)} \]
\[ \text{PER}_c(A) = \sum_{\sigma \in S_m} \prod_{j=1}^m a_{\sigma(j),j} \]
to be the generic row and column permanents of \( A, \) respectively. \( S_m \) is the set of permutations of \( m. \)

2.5. Notation Let \( A = (a_{ij}) \) be an \( m \times n \) matrix. If \( g \in \pi^n, \) the set of all functions from \( m \) to \( n, \) then \( A_{1g} \) is the \( m \times m \) matrix having columns \( g(1), \ldots, g(m) \) from \( A. \) Writing \( A(i, j) = a_{ij}, \) one has \( A_{1g}(i, j) = a_{i,g(j)}. \)

An \( m \times n \) matrix \( A = (a_{ij}) \) with entries \( mn \) distinct elements of \( W \) will be called a generic matrix over \( W. \) In this notation, we have:
\[ \text{PER}_r(A_{1g}) = \sum_{\sigma \in S_m} \prod_{j=1}^m a_{\sigma(j),j}. \]

2.6. Lemma Let \( u_i = \sum_{j=1}^n a_{ij}, i = 1, 2, \ldots, m \) be the row sums of a generic matrix \( A = (a_{ij}). \) Then the shuffle product of the row sums is...
expressible as

\[ u_1 \cup \cdots \cup u_m = \sum_{\sigma \in S_m} u_{\sigma 1} \cdots u_{\sigma m} \]

and

\[ u_1 \cup \cdots \cup u_m = \sum_{\sigma \in S_m} \text{PER}_\sigma (A_{1g}) . \]

**Proof** For distinct generic variables \( x_{ij} \) one has

\[ \prod_{i=1}^{m} \left( \sum_{j=1}^{n} x_{ij} \right) = \sum_{\sigma \in S_m} \prod_{i=1}^{m} x_{\sigma i j} \]

since the left-hand side is the sum of monomials \( x_{1j_1} \cdots x_{nj_n} \) and the \( j \)'s define the \( g \). Replace \( x_{ij} \) by \( a_{i,j} \) and sum on \( \sigma \in S_m \) to obtain

\[ \sum_{\sigma \in S_m} \prod_{i=1}^{m} \left( \sum_{j=1}^{m} a_{i,j} \right) = \sum_{\sigma \in S_m} \prod_{i=1}^{m} a_{\sigma i,j} \]

and this is the sum of monomials \( a_{\sigma 1,j} \cdots a_{\sigma n,j} \) which is evidently

\[ u_1 \cup \cdots \cup u_m . \]

### 2.7. Notation

The following notation will be useful in describing the generic Laplace expansion and Cauchy–Binet theorem.

For an \( m \times n \) matrix \( A \) and \( I \subseteq m \) and \( J \subseteq n \) nonempty, let \( A[I,J] \) be the submatrix of \( A \) having rows indexed by \( I \) and columns indexed by \( J \). Let \( A(I,J) = A[I,J] \), \( I = m - I \) and \( J = n - J \), be the matrix obtained by deleting rows indexed by \( I \) and columns indexed by \( J \). The mixed notation \( A[I,J] \) and \( A[IU] \) have the obvious meanings.

For \( I \) and \( J \) subsets of \( m \), let \( S_m(I,J) \) be the set of all permutations of \( n \) which map \( I \) to \( J \):

\[ S_m(I,J) = \{ \sigma \in S_m : \sigma(I) = J \} . \]

**Note**

(a) A permutation \( \sigma \) of \( S_m(I,J) \) is completely determined by its restrictions \( \sigma[I] \) and \( \sigma[I]^c \) to \( I \) and its complement \( \bar{I} \). Thus if \( |I| = p \),

\[ |S_m(I,J)| = p! \cdot (m - p)! . \]

(b) If \( I_0 \) is a fixed nonempty subset of \( m \), then \( S_m \) is the disjoint union

\[ S_m = \bigcup_{J \subseteq m} S_m(I_0,J) . \]
Likewise for fixed nonempty \( J_0 \subseteq m \),
\[
S_m = \bigcup_{I \subseteq m} S_m(I, J_0).
\]

These disjoint unions are reflected in
\[
m! = p! (m - p)!.
\]

2.8. THEOREM (Laplace expansion for generic permanents) Let \( A = (a_{ij}) \) be a generic \( m \times m \) matrix (entries are distinct elements of \( \mathbb{W} \)).

(1) For \( I_0 \subseteq m \) fixed and \( |I| = |I_0| \),
\[
\text{PER}_\xi(A) = \sum_{J \subseteq m} \text{PER}_\xi(A[I_0|J]) |J_0| \text{PER}_\xi(A(I_0|J)).
\]

(2) For \( J_0 \subseteq m \) fixed and \( |J| = |J_0| \),
\[
\text{PER}_\xi(A) = \sum_{I \subseteq m} \text{PER}_\xi(A[I|J_0]) |I| \text{PER}_\xi(A(I|J_0)).
\]

(3) For \( J_0 \subseteq m \) fixed and \( |J| = |J_0| \),
\[
\text{PER}_\xi(A) = \sum_{I \subseteq m} \text{PER}_\xi(A[I|J]) |I| \text{PER}_\xi(A(I|J)).
\]

(4) For \( I_0 \subseteq m \) fixed and \( |I| = |I_0| \),
\[
\text{PER}_\xi(A) = \sum_{J \subseteq m} \text{PER}_\xi(A[|J_0|J]) |J| \text{PER}_\xi(A(I_0|J)).
\]

Proof Consider the first formula (1). By definition of the \( I_0 \)-shuffle, one has for any \( \sigma \in S_m \):
\[
\prod_{i=1}^m a_{\sigma(i)} = \prod_{i \in I_0} a_{\sigma(i)} \prod_{i \in I \setminus I_0} a_{\sigma(i)}.
\]

Now, if \( \sigma \in S_m(I_0, J) \), then \( \prod_{i \in J} a_{\sigma(i)} \) is a summand for \( \text{PER}_\xi(A[I_0|J]) \) and likewise the second factor for \( \text{PER}_\xi(A(I_0|J)) \). By note (a) above, we have
\[
\sum_{\sigma \in S_m(I_0, J)} \prod_{i=1}^m a_{\sigma(i)} = \text{PER}_\xi(A[I_0|J]) |I_0| \text{PER}_\xi(A(I_0|J)).
\]

By note (b) above, summing over \( J \subseteq m \) gives \( \text{PER}_\xi(A) \) proving formula (1). Formulas (2), (3), and (4) are proved in a similar fashion.

Classically, formula (1) would be called the expansion by rows indexed by \( I_0 \).
Let $x_1, \ldots, x_m$ be noncommuting variables each of which commutes with every element $W$. Thus $X = \begin{pmatrix} x_1 & 0 \\ \vdots & \ddots \\ 0 & x_m \end{pmatrix}$ is a generic diagonal matrix. In a generic $m \times n$ matrix $A = (a_{ij})$, we may replace $a_{ij}$ by $x_i a_{ij} = a_{ij} x_i$ in any formula. For example,

$$\text{PER}_S(A X) = \sum_{\sigma \in \Sigma} \prod_{i=1}^m a_{i,\sigma(i)} \prod_{i=1}^m x_{\sigma(i)}.$$ 

The effect is twofold: $\prod x_{\sigma(i)}$ keeps a record of the permutations $\sigma$ in the summands, and it also plays a role like that of the signs distinguishing a determinant from a permanent. Observe also $\text{PER}_S(A X)$ is a polynomial in the set of noncommuting variables $\{x_1, \ldots, x_m\}$. The permanent $\text{PER}_S(A X)$ will be called the generic row matrix function.

Likewise,

$$\text{PER}_S(X A) = \sum_{\sigma \in \Sigma} \prod_{j=1}^n a_{\sigma(j),j} \prod_{j=1}^n x_{\sigma(j)}$$

is the generic column matrix function.

It is well known that the classical Laplace expansion theorem is a consequence of the associativity of the exterior product in a Grassmanian algebra [6]. In a similar fashion, the Laplace expansion theorem for generic permanents and its consequences for generic matrix functions, is a consequence of the associativity of the shuffle product in the shuffle algebra. We now explain the connection in this latter setting.

The key is several applications of Lemma 2.6, but we require a technical observation. Let $I_0$ be a fixed subset of $m$. For each $J \subseteq m$ with $|J| = |I_0|$, let $b_J$ denote the order preserving bijection from $I_0$ to $J$. If $h \in \mathcal{P}^n$, then $h \circ b_J \in \mathcal{P}^n$. Thus $h \circ b_J$ sends $i_1$ to $h(j_1)$, where $i_1 < i_2 < \cdots < i_p$ and $j_1 < j_2 < \cdots < j_p$ are the ordered elements of $I_0$ and $J$, respectively. Likewise, $b_J$ is the order-preserving bijection from $I_0 = m - I_0$ to $J$ and $h \circ b_J \in \mathcal{P}^n$.

The observation: Given $I_0$ and $J$, $|J| = |I_0|$, each $h \in \mathcal{P}^n$ gives rise to the pair of functions $h \circ b_J \in \mathcal{P}^n$ and $h \circ b_J \in \mathcal{P}^n$, and, for fixed $J$, every pair of functions from $\mathcal{P}^n$ and $\mathcal{P}^n$ is so obtained:

$$\{(f, g) : f \in \mathcal{P}^n, g \in \mathcal{P}^n\} = \{h \circ b_J, h \circ b_J) : h \in \mathcal{P}^n\}.$$ 

Indeed, this is a “change of indices” from $I_0$ to $J$. Recalling Notation 2.7, we have
2.9. **Theorem**  Let \( A \) be an \( m \times n \) generic matrix. Then for each fixed subset \( I_0 \subseteq m \).

\[
\sum_{h \in \mathcal{P}} \operatorname{PER}_s(A_{1,h}) = \sum_{h \in \mathcal{P}} \left( \sum_{j \in \mathcal{P}} \operatorname{PER}_s(A_{1,h \cup [I_0[n]]}) \right) \frac{1}{|I_0|} \operatorname{PER}_s(A_{1,h \cup [I_0[n]]})
\]

**Proof**  Let \( u_i = \sum_{j=1}^n a_{ij} \) be the row sums of \( A \). For any \( \sigma \in S_m \),

\[
u_{1,1,1, \ldots ,1} u_m = \nu_{1,1,1, \ldots ,1} u_m
\]

by (1) of Lemma 2.6. Thus by associativity,

\[
u_{1,1,1, \ldots ,1} u_m = \left( \underbrace{\nu_{1,1,1, \ldots ,1} u_1}_{1\in \mathcal{P}_I} \right) \ldots \left( \underbrace{\nu_{1,1,1, \ldots ,1} u_i}_{i\in \mathcal{P}_I} \right)
\]

where \( \nu_{1,1,1, \ldots ,1} \) is the shuffle product in the order given by \( I_0 \) and likewise for \( I_0 \). By (2) of Lemma 2.8, applied once to the left side and twice to the right side of the last equation, gives

\[
\sum_{h \in \mathcal{P}} \operatorname{PER}_s(A_{1,h}) = \left( \sum_{h \in \mathcal{P}} \operatorname{PER}_s(A_{1,h \cup [I_0[n]]}) \right) \frac{1}{|I_0|} \left( \sum_{\sigma \in S_m} \operatorname{PER}_s(A_{1,\sigma(I_0[n])}) \right)
\]

The desired assertion now follows from the observation preceding the theorem.

Since the matrix \( A = (a_{ij}) \) of Theorem 2.9 is generic, the monomials in

\[
\operatorname{PER}_s(A_{1,h}) = \sum_{\sigma \in S_m} \prod_{i=1}^m a_{i \sigma(i)}
\]

are linearly independent over the field \( k \) and we may extract the summands for each \( h \) in the theorem to obtain a formula for them.

2.10. **Corollary**  For any \( h \in \mathcal{P}^m \) and \( I_0 \subseteq m \),

\[
\operatorname{PER}_s(A_{1,h}) = \sum_{j \in \mathcal{P}} \operatorname{PER}_s(A_{1,h \cup [I_0[n]]}) \left( \frac{1}{|I_0|} \right) \operatorname{PER}_s(A_{1,h \cup [I_0[n]]})
\]

If \( A \) is an \( m \times m \) square matrix and \( h \) is the identity map, then Corollary 2.10 is just Theorem 2.8(4).

Both 2.9 and 2.10 immediately give expansions for the generic column matrix functions expressing \( \operatorname{PER}_s(XA_{1,h}) \) and its sums over \( h \in \mathcal{P}^m \) in
terms of sums of

\[ \text{PER}_x((X A)_{1,h}^{-, j} | \langle I_0 \rangle_{|J|}) \mid \langle J_0 \rangle_{|J|} \text{ PER}_x((X A)_{1,h}^{-, j} | \langle I_0 \rangle_{|J|}) \]

over \( J \subseteq m \) and \( h \in \mathfrak{n}^m \). Theorem 2.9 and Corollary 2.10 have obvious analogs for row permanents \( \text{PER}_x \).

3. GENERIC SCHUR MATRIX FUNCTIONS

Classically, one obtains the determinant from the permanent by introducing a plus or minus sign before a summand \( a_{1,i_1} \cdots a_{m,i_m} \), the sign being the value of the alternating character of the permutation \( \sigma \) in \( S_m \). This choice of signs was extended by Schur [7] to include representations of \( S_m \) and such an extension may be incorporated into the generic case also.

3.1. Definition (Generic Schur matrix functions) Let \( M \) be a representation of the group \( S_m \) of perturbations of \( m \). For any \( m \times m \) matrix \( A = (a_{ij}) \) with entries in \( W \), define

\[
\text{PER}_x^R(A) = \sum_{\sigma \in S_m} \prod_{i=1}^{m} a_{i, \sigma(i)} M(\sigma)
\]

\[
\text{PER}_x^C(A) = \sum_{\sigma \in S_m} \prod_{j=1}^{m} a_{\sigma(j), j} M(\sigma)
\]

to be the generic row and column Schur matrix functions of \( A \) with respect to \( M \).

When \( M \) is a matrix representation of \( S_m \), these Schur matrix functions are matrices with entries in \( k(W) \).

To obtain a Laplace expansion theorem for Schur matrix functions, we need several preliminary observations.

3.2. Let \( I \subseteq m \) be of cardinality \( |I| = p \).

(a) Let \( c_I \) be the order-preserving bijection from \( p \) to \( I \). If \( x \in S_p \), define \( \hat{x} \) by \( \hat{x} |I| = c_I \circ x \circ c_I^{-1} \). Thus \( \hat{x} \in S_m(I, I) = S_m(I, I) \) and is just carrying out the permutation \( x \) on the \( p \) elements of \( I \) and fixing pointwise the elements of \( I \).
(b) In a like manner, let $\tilde{c}_t$ be the order-preserving bijection from $\tilde{T} = m - p$ to $\tilde{T}$. If $\beta \in S_{m-p}$, define $\tilde{\beta}$ by \[
abla \tilde{T} = \tilde{c}_t \circ \beta \circ \tilde{c}_t^{-1}.
\] Then $\tilde{\beta} \in S_m(\tilde{T}, \tilde{T}) = S_m(I, I)$ is the permutation $\tilde{\beta}$ carried out on the $m - p$ elements of $\tilde{T}$.

Let $I$ and $J$ be two subsets of $m$ of the same cardinality $p$. Recall that $S_m(I, J)$ is the set of $\sigma \in S_m$ mapping $I$ to $J$.

(a) Define $\psi$ by \[
\psi I = c_I \circ \tilde{c}_t^{-1}.
\] Then $\psi$ is the unique map of $S_m(I, J) = S_m(\tilde{T}, \tilde{T})$ which preserves order from $I$ to $J$ and from $\tilde{T}$ to $\tilde{T}$.

(b) Note that every element of $S_m(I, J) = S_m(\tilde{T}, \tilde{T})$ is uniquely $\psi \tilde{\beta} = \psi \tilde{\beta}$ with $x \in S_p$ and $\beta \in S_{m-p}$.

**Note** The dependency of $\tilde{z}$ and $\tilde{T}$ on $I$ and of $\psi$ on $I$ and $J$ may be denoted by subscripts: $\tilde{z}_I, \tilde{T}_I, \psi_{i,j}$. These subscripts will frequently be omitted.

(c) If $M$ is a representation of $S_m$, then $M(\tilde{T}) = M(c_I \circ \tilde{c}_t^{-1})$ gives a representation of $S_p$: $x$ is sent to $M(\tilde{T})$. If $A[I,J]$ is the $p \times p$ submatrix of $A$ obtained by selecting rows from $I$ and columns from $J$, then

\[
\text{PER}^M(A[I,J]) = \sum_{\tilde{T}} \prod_{x \in \tilde{T}} a_{x \tilde{T}} M(\tilde{T})\]

A similar observation holds for $M(\tilde{T}) = M(c_I \circ \tilde{c}_t^{-1})$ and $A[I,J]$ = $A[I,J]$.

3.3. Theorem (Laplace expansion theorem for generic Schur matrix functions) Let $A = (a_{ij})$ be a generic $m \times m$ matrix and $M$ a representation of $S_m$. For generic row Schur matrix functions one has:

(1) For $I_0 \subseteq m$ fixed and $|J| = |I_0|$,

\[
\text{PER}^M(A) = \sum_{I \subseteq m} M(\psi_{I,J_0}) \text{PER}^M(\tilde{A}[I,J]_{I_0}) \text{PER}^M(\tilde{A}[I,J])
\]

(2) For $J_0 \subseteq m$ fixed and $|I| = |I_0|$,

\[
\text{PER}^M(A) = \sum_{J \subseteq m} M(\psi_{I_0,J}) \text{PER}^M(\tilde{A}[I,J_{0}]_{I}) \text{PER}^M(\tilde{A}[I,J_{0}])
\]

And similar formulas for generic column Schur matrix functions.
Proof. This proof is an extension of that of 2.8. By the definition of the $I_o$-shuffle and 3.2(b),

$$
\prod_{i=1}^{m} a_{i,\sigma(i)} = \prod_{i=1}^{m} a_{i,\sigma(i)} \prod_{i=1}^{m} a_{i,\sigma(i)},
$$

where $\sigma = \psi \hat{\sigma}$. Multiply by $M(\sigma) = M(\psi)M(\hat{\sigma})M(\hat{\beta})$ and sum on $\sigma \in S_m(I_o, J)$ to obtain

$$
\sum_{\sigma \in S_m(I_o, J)} a_{i,\sigma(i)} M(\sigma) = M(\psi) \sum_{\sigma \in S_m(I_o, J)} \prod_{i=1}^{m} a_{i,\sigma(i)} M(\hat{\sigma}) \prod_{j \in I_o, i \in I_o} a_{i,\sigma(i)} M(\hat{\beta}).
$$

By note (c) above, the right-hand side is

$$
M(\psi_{I_o,J}) \text{PER}_{S_m}(A[I_o,J])[I_o] \text{PER}_{S_m}(A[I_o,J]).
$$

The proof is concluded by summing over $J \subseteq m$ by note 2.7(b). The other formulas are proved in a similar fashion.

3.4. Remark. Let $A$ be a generic $m \times m$ matrix, $X$ a generic diagonal matrix whose entries commute with those of $A$, and $M$ a representation of $S_m$. Since $x_1 \cdots x_m$ uniquely determines $\sigma$, we may replace $x_1 \cdots x_m$ in $\text{PER}_{S_m}(AX)$ by $M(\sigma)$ to obtain $\text{PER}_{S_m}(A)$. Theorem 3.3 is a consequence of Theorem 2.8 for matrices $AX$ and $XA$ once one describes the multiplication rules for the sequences of $x$'s.

In fact, the four formulas for

$$
\text{PER}_{S_m}(AX) = \sum_{\sigma \in S_m} \prod_{i=1}^{m} a_{i,\sigma(i)} \prod_{i=1}^{m} x_{\sigma(i)} M(\sigma)
$$

encompass both sets of formulas.

3.5. Example. Let $e$ be the representation of $S_m$ which is the alternating character which assigns to $\sigma \in S_m$ its sign $e(\sigma) = \pm 1$ according as the permutation is even or odd. If $I = \{i_1 < \cdots < i_p\}$ and $J = \{j_1 < \cdots < j_q\}$ in $m$, then

$$
e(\psi_{I,J}) = (-1)^{\sum_{i,j} x_i + x_j}.
$$

In this case, Theorem 3.3(1) applied to matrices $A$ is the Laplace expansion of

$$
\det A = \text{PER}_{S_m}(A)
$$

by unions and complementary unions with its well-known choice of signs [6]. In the classical case, the entries of $A$ are commuting variables.
4. GENERIC CAUCHY–BINET THEOREMS

The classical expansion theorem of Cauchy–Binet [6] has analogs for generic permanents and Schur matrix functions that are of the same style as previous analogies. We begin with a preliminary observation.

The group $S_m$ of permutations of $m$ acts on $\mathcal{P}^m$ by means of the argument of $f \in \mathcal{P}^m$: $\sigma \cdot f = f \circ \sigma^{-1}$. The stability subgroup of an element $f \in \mathcal{P}^m$ has cardinality

$$\nu(f) = \prod_{j \in \text{fix}(f)} |f^{-1}(j)|!$$

and each orbit of $S_m$ contains a unique element of $\mathcal{P}^m$, the set of nondecreasing functions from $m$ to $n$. Thus a sum of quantities $Q(f)$ indexed by $f \in \mathcal{P}^m$ may be written as a sum over the disjoint orbits of $S_m$:

$$\sum_{f \in \mathcal{P}^m} Q(f) = \sum_{f \in \mathcal{P}^m / S_m} \frac{1}{\nu(f)} \sum_{\phi \in S_m} Q(f \circ \phi^{-1}).$$

4.1. THEOREM (Generic Cauchy–Binet) Let $A = (a_{ij})$ and $B = (b_{ij})$ be generic $m \times n$ and $n \times m$ matrices, respectively. Then

$$\text{PER}_m(AB) = \sum_{h \in \mathcal{P}^m} \frac{1}{\nu(h)} \sum_{\phi \in S_m} \prod_{i=1}^m A_{1,h_{i-1}}(i,i) \text{PER}_s(B_{h_{i-1},1})$$

and

$$\text{PER}_n(AB) = \sum_{h \in \mathcal{P}^m} \frac{1}{\nu(h)} \sum_{\phi \in S_m} \prod_{j=1}^n \text{PER}_s(A_{1,h_{j-1}})B_{h_{j-1},1}(j,j).$$

Proof. As in the proof of Lemma 2.6, for distinct generic variables $x_a$ one has

$$\prod_{i=1}^m \sum_{j=1}^n x_{a_{ij}} = \sum_{h \in \mathcal{P}^m} \prod_{i=1}^m x_{a_{hi}}.$$

For fixed $\sigma \in S_m$, let $x_{\sigma} = a_{i,\sigma(i)}b_{\sigma(i),i}$. Then

$$\prod_{i=1}^m \sum_{\sigma(j) = i} a_{i,\sigma(j)}b_{\sigma(j),i} = \sum_{h \in \mathcal{P}^m} \prod_{i=1}^m a_{i,h_i}b_{h_i,i}$$

or

$$\prod_{i=1}^m (AB)(i,\sigma(i)) = \sum_{h \in \mathcal{P}^m} \prod_{h_{\sigma(i)}} A(i,h_i)B(h_i,\sigma(i)).$$
Recall from 2.5, $A_{14}(i, j) = a_{i, b_j}$. Sum on $\sigma \in S_m$ to obtain

$$\text{PER}_S(AB) = \sum_{h \in G} \prod_{i=1}^m A(i, h_i) \text{PER}_S(B_{h_i}).$$

By the preliminary observation, we have the first formula. The second formula is proved in a similar fashion.

4.2. Remark  As in Section 3, a representation $M$ may be included in the Cauchy-Binet theorem. If, in the first formula of the proof of 4.1 we set $x_{i, h} = a_{i, h_i} M(\sigma)$, the succeeding steps lead to $\text{PER}_S$. This gives as corollary of the proof the formula

$$\text{PER}_S(AB) = \frac{1}{\mu(\sigma) \sigma \in \rho S_m} \sum_{h \in \rho} \prod_{i=1}^m A(i, h_i) \text{PER}_S(B_{h_i})$$

and the analogous ones for $\text{PER}_S(AB)$.

Suppose now that the entries of an $m \times m$ square matrix $C$ commute. Then we have an equality of products

$$\prod_{i=1}^m C_{i, j} = \prod_{i=1}^m C_{i, j}$$

for any $\sigma \in S_m$. We may apply this observation to the $m \times m$ matrix $B_{h_1}$ obtained from an $n \times m$ matrix $B$ and $h \in \rho^m$.

4.3. Lemma  Let $B$ be a generic $n \times m$ matrix of commuting variables. Then

$$\text{PER}_S(\phi^{-1})(B_{h_1}) = M(\phi^{-1}) \text{PER}_S(B_{h_1})$$

for $h \in \rho^m$ and $\phi \in S_m$.

Proof  As $\sigma$ ranges over $S_m$ so does $\phi^{-1} \sigma$, so for $B = (b_{i, j})$ we have

$$\sum_{\sigma \in S_m} \prod_{i=1}^m b_{h_1(i, \sigma)} M(\sigma) = \sum_{\sigma \in S_m} \prod_{i=1}^m b_{h_1(\phi^{-1}(i, \sigma)}, M(\phi^{-1}(\sigma)).$$

Since $M$ is a representation, the preceding observation gives the formula.

This now yields a Cauchy-Binet theorem where the entries of one matrix commute among themselves but not necessarily with the entries of the other matrix.
4.4. Theorem Let $A$ and $B$ be generic $m \times n$ and $n \times m$ matrices, respectively with the entries of $B$ commuting among themselves, and let $M$ be a representation of $S_m$. Then

$$\text{PER}_S^M(AB) = \sum_{h \in S_n} \frac{1}{v(h)} \text{PER}_S^M(A_{1h}) \text{PER}_S^M(B_{h1}).$$

Proof By 4.2 and the 4.3, $\text{PER}_S^M(AB)$ is

$$\sum_{h \in S_n} \frac{1}{v(h)} \sum_{a \in S_n} \prod_{i=1}^m A_{1h}(i, i)M(h^{-1}) \text{PER}_S^M(B_{h1}).$$

Replacing $\phi^{-1}$ by $\phi$ we recognize the inner sum as $\text{PER}_S^M(A_{1h})$.

The analogous formula for $\text{PER}_S^M(AB)$ holds when the entries of $A$ commute, but not necessarily with those of $B$.

4.5. Example Let the entries of the $m \times m$ square matrix $A$ commute.

(a) If the representation $M$ is the trivial character assigning the scalar 1 to every $\sigma \in S_m$, then

$$\text{per } A = \text{PER}_S^M(A)$$

is the classical permanent of $A$. By commutativity, this row permanent is the same as the column permanent $\text{PER}_S^M(A)$. By 4.3, $\text{per } A$ is a symmetric function of the rows or columns.

(b) If the representation $M$ is the alternating character $\varepsilon$ as in 3.5, then

$$\text{det } A = \text{PER}_S^M(A)$$

is the classical determinant of $A$. This row determinant is the same as the column determinant $\text{PER}_S^M(A)$ by commutativity. By 4.3, $\text{det } A$ is an alternating function of its rows or columns.

Let $\mathcal{R}^n$ denote the set of all strictly increasing functions from $m$ to $n$.

4.6. Corollary (Classical Cauchy-Binet) Let $A$ and $B$ be $m \times n$ and $n \times m$ matrices with all entries commuting among each other. Then

(a) $\text{per } AB = \sum_{h \in \mathcal{R}^n} \frac{1}{v(h)} \text{per } A_{1h} \text{ per } B_{h1}$.

(b) $\text{det } AB = \sum_{h \in \mathcal{R}^n} \text{det } A_{1h} \text{ det } B_{h1}$. 


Proof This is an immediate consequence of 4.4 and 4.5 once we observe for (b) that $\det A_{1k} = 0$ by repeated columns unless $h \in \mathcal{H}^n$ is strictly increasing, and in this case $v(h) = 1$.

References