On a formula for moments of the multivariate normal distribution generalizing
Stein’s lemma and Isserlis theorem

Konstantinos Mamis

\textit{Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA}

Abstract

We prove a formula for the evaluation of averages containing a scalar function of a Gaussian random vector multiplied by a product of the random vector components, each one raised at a power. Some powers could be of zeroth-order, and, for averages containing only one vector component to the first power, the formula reduces to Stein’s lemma for the multivariate normal distribution. On the other hand, by setting the said function inside average equal to one, we easily derive Isserlis theorem and its generalizations, regarding higher order moments of a Gaussian random vector.

We provide two proofs of the formula, with the first being a rigorous proof via mathematical induction. The second is a formal, constructive derivation based on treating the average not as an integral, but as the consecutive actions of pseudodifferential operators defined via the moment-generating function of the Gaussian random vector.

Keywords: normal distribution, Stein’s lemma, Isserlis theorem, Wick’s theorem, Hermite polynomials, pseudodifferential operator

2000 MSC: primary 62H10, secondary 60E05, 60C05

1. Introduction and main results

Stein’s lemma, \cite{27, 12}, is a well-known identity of the normal distribution, with applications in statistics, see e.g. \cite{29, 17}. For the case of a scalar random variable $X$ that follows the univariate normal distribution $N(\mu, \sigma^2)$, Stein’s lemma reads

$$E[g(X)X] = \mu E[g(X)] + \sigma^2 E[g'(X)],$$  \hspace{1cm} (1)

where $E[\cdot]$ is the average operator, and prime denotes the first derivative of function $g$. In our recent work \cite{13}, we extended scalar Stein’s lemma (1) for averages containing $X$ at a power $n$:

$$E[g(X)X^n] = \sum_{\ell=0}^{n} \binom{n}{\ell} \mu^{n-\ell} \sum_{k=0}^{\ell/2} H_{\ell,k} \sigma^{2(\ell-k)} E[g^{(\ell-2k)}(X)], \hspace{1cm} n \in \mathbb{N}. \hspace{1cm} (2)$$

In Eq. (2), $g^{(\ell)}$ denotes the $\ell$th derivative of $g$, $\binom{n}{\ell}$ is the binomial coefficient, $\lfloor \cdot \rfloor$ is the floor function, and

$$H_{\ell,k} = \frac{\ell!}{2^k (\ell-2k)!}, \hspace{1cm} k = 0, \ldots, \lfloor \ell/2 \rfloor, \hspace{1cm} (3)$$

are the \textit{signless Hermite coefficients}, i.e., the absolute values of the coefficients appearing in the probabilist’s Hermite polynomial of $\ell$th order; $He_{\ell}(x) = \sum_{k=0}^{\ell/2} (-1)^k H_{\ell,k} x^{\ell-2k}$; \cite[expression 22.3.11]{1}. Results similar to Eq. (2) have also been derived in \cite{19, 20}, by using Rodrigues formula.

For a random vector $X$ following the $N$-variate normal distribution $N(\mu, C)$, the multivariate Stein’s lemma reads

$$E[g(X)X_i] = \mu_i E[g(X)] + \sum_{j=1}^{N} C_{ij} E[\partial_j g(X)], \hspace{1cm} i = 1, \ldots, N. \hspace{1cm} (4)$$
where \( \partial g(X) = \partial g(X)/\partial X_j \) is the first partial derivative of \( g(X) \) with respect to \( X_j \) component. Eq. (4) is easily proven by integration by parts, using the properties of the derivatives of the \( N \)-variate normal distribution, see [18, Eq.347]. The topic of the present work is to generalize Stein’s lemma, providing a formula, in closed form, of expressing the average \( \mathbb{E}[g(X) \prod_{j=1}^{N} X_j^{\mu_j}] \) in terms of the averages of partial derivatives of \( g(X) \).

**Theorem 1** (Generalized multivariate Stein’s lemma). For a smooth enough function \( g : \mathbb{R}^N \to \mathbb{R} \), and under the assumption that all averages involved exist, it holds true that

\[
\mathbb{E}[g(X)X^n] = \sum_{r \leq n} \binom{n}{L} H_{r,L,K} \mathbb{E}[\partial^{L-K} \partial^{K} g(X)].
\]

Eq. (5) is expressed using multi-index \([23, \text{p. } 319]\) and index-matrix \([22, 25]\) notations. By using the multi-index and index-matrix notation, the generalized multivariate Stein’s lemma (5) resembles in form the respective result (2) for the scalar case. Multi-indices are denoted by lowercase boldface letters, and index-matrices are denoted by uppercase boldface letters: \( n = [n_i]_{i=1}^{N} \in \mathbb{N}_{0}^{N}, L = [\ell_{ij}]_{i,j=1}^{N} \in \mathbb{N}_{0}^{N \times N} \), and \( K = [k_{ij}]_{i,j=1}^{N} \in \mathbb{N}_{0}^{N \times N} \). Index-matrix \( L \) is, in general, nonsymmetric, while index-matrix \( K \) is symmetric. Index-matrices \( K_{\text{max}}(L) = [k_{ij}^{\text{max}}]_{i,j=1}^{N} \in \mathbb{N}_{0}^{N \times N}, \tilde{L} = [\tilde{\ell}_{ij}]_{i,j=1}^{N} \in \mathbb{N}_{0}^{N \times N} \) and \( \tilde{K} = [\tilde{k}_{ij}]_{i,j=1}^{N} \in \mathbb{N}_{0}^{N \times N} \) are symmetric, and are defined as:

\[
k_{ii}^{\text{max}} = \lfloor \ell_{ii}/2 \rfloor, \quad \text{and} \quad k_{ij}^{\text{max}} = \min(\ell_{ij}, \ell_{ji}) \quad \text{for} \quad i \neq j,
\]

\[
\ell_{ii} = \ell_{ii}, \quad \text{and} \quad \tilde{\ell}_{ii} = \ell_{ii} + \ell_{ji} \quad \text{for} \quad i \neq j,
\]

\[
k_{ii} = 2k_{ii}, \quad \text{and} \quad \tilde{k}_{ij} = k_{ij} \quad \text{for} \quad i \neq j.
\]

Partial ordering of multi-indices \( m \leq n \) implies that \( m_i \leq n_i \) for all \( i = 1, \ldots, N \). Partial ordering of index-matrices \( K \leq L \) implies that \( k_{ii} \leq \ell_{ii} \) for all \( i, j = 1, \ldots, N \). By \( r(L) \) and \( c(L) \), we denote the row-sum and column-sum vectors of index-matrix \( L \) respectively:

\[
r_{i}(L) = \sum_{j=1}^{N} \ell_{ij}, \quad c_{i}(L) = \sum_{j=1}^{N} \ell_{ji}, \quad i = 1, \ldots, N.
\]

By \( C_L \) we denote the upper triangular part of the symmetric matrix \( C \). Raising a vector to a multi-index power is defined as \( \mu^n = \prod_{i=1}^{N} \mu_i^{n_i} \). Raising an upper triangular matrix to a symmetric index-matrix power is defined as \( C_L^{K} = \prod_{i=1}^{N} \prod_{j=1}^{N} C_{ij}^{k_{ij}} \). The partial derivative of \( g(X) \) of multi-index order \( m \) is defined as \( \partial^m g(X) = \prod_{i=1}^{N} \partial_i^{m_i} g(X) \).

Also, we introduce the multinomial coefficient \( \binom{n}{L} \) of a multi-index vector and an index-matrix with respect to the rows of the matrix, called the \( r \)-multinomial coefficient henceforth:

\[
\binom{n}{L} = \frac{n!}{[n-r(L)]!L!}.
\]

The factorial of a multi-index is defined as \( n! = \prod_{i=1}^{N} n_i! \). The factorial of an index-matrix is defined as \( L! = \prod_{i=1}^{N} \prod_{j=1}^{N} \ell_{ij}! \). Last, the coefficients \( H_{r,L,K} \) are defined as

\[
H_{r,L,K} = \frac{L!}{2^{\text{tr}(K)} K_U(L - \tilde{K})!},
\]

where \( \text{tr}(K) = \sum_{i=1}^{N} k_{ii} \) is the trace of index-matrix \( K \), and \( K_U = \prod_{i=1}^{N} \prod_{j=1}^{N} k_{ij}! \).

**Proof.** In Sec. 2, we prove Eq. (5) rigorously, by multidimensional mathematical induction on \( n \in \mathbb{N}_{0}^{N} \). In addition to this proof, and in order to provide the reader with more insight, we present, in Sec. 3, a constructive formal proof of theorem 1. This constructive proof is based on treating the mean value operator not as an integral, but as the sequential action of a number of pseudodifferential operators that are introduced in definition 1 via the moment-generating function of the Gaussian random vector. The action of these pseudodifferential operators is determined by their Taylor series expansions, under the formal assumption that all infinite series involved are summable. \( \square \)
For each of these index-matrices, we calculate, in Table 1, the rest of the quantities appearing in each term of the $L$-sum in the right-hand side of Eq. (5). By substituting the quantities from Table 1 into Eq. (5), and after some regrouping of terms, we obtain Eq. (12). Note that Eq. (12) can also be validated by repetitive applications of Stein’s lemma, Eq. (4), for $\mathbb{E} \left[ g(X_1, X_2, X_3) \right]$.}

**Corollary 1** (Product moment formula for Gaussian vectors). For the Gaussian random vector $X$, the following formula for its product moments holds true:

$$
\mathbb{E} \left[ X^n \right] = \sum_{K \in \mathbb{N}_0^N, r(K) \leq n} d_{n,K} \mu^{n-r(\hat{K})} C^K_{\mu},
$$

where $\hat{K}$ is defined from $K$ using Eq. (8), and

$$
d_{n,K} = \frac{n!}{2^{n(K)}} \frac{K!}{(n-r(\hat{K}))!}.
$$
Summation in the right-hand side of Eq. (13) extends over all index-matrices $K$ for which the respective index-matrix $\tilde{K}$ satisfies the condition $r(\tilde{K}) \leq n$.

**Proof.** Eq. (13) has been proven by Song and Lee in [26], using results from the work of Price [21]. Here, we easily derive Eq. (13) by setting $g(X) = 1$ in formula (5). By setting $g(X) = 1$, all derivatives in the right-hand side of Eq. (5) are zero, except for the zeroth order derivative, that is for $c(L - \tilde{K}) = 0$. By the definition relation (8) of $\tilde{K}$, $c(L - \tilde{K}) = 0$ is achieved for $L = \tilde{K}$. Last, coefficients $a_{n,k}$ are calculated as

$$\binom{n}{K} H_{k,k} = \frac{n!}{(n-r(\tilde{K}))(r(\tilde{K}))!} \frac{\tilde{K}'}{2^m K U!} = a_{n,k}.$$ 

Substitution of the above in Eq. (5) results in Eq. (13). □

**Corollary 2** (Isserlis theorem). From Eq. (13), we derive the formula for the higher order moments of an $N$-dimensional Gaussian random vector $X$ with zero mean value:

$$E\left[\prod_{i=1}^{N} X_i\right] = \begin{cases} 0 & \text{for } N \text{ odd}, \\ \sum_{P \in \mathcal{P}} \prod_{(i,j) \in P} C_{ij} & \text{for } N \text{ even}, \end{cases}$$

with $\mathcal{P}$ being the set of all partitions of $\{1, \ldots, N\}$ into unordered pairs. Eq. (15) has been proven by Isserlis [10], and also known in physics literature as Wick’s theorem [30]. For a review of the related literature see also [28].

**Proof.** For $\mu = 0$ and $n = 1$, the summation in Eq. (13) extends over all $K$ with $r(\tilde{K}) = 1$. Since the diagonal elements of $\tilde{K}$ are even numbers (see definition relation (8)), $\tilde{K}$ in this case has zero diagonal elements, and it is equal to $K$. Also, we calculate that $d_{1,k} = 1$. Thus, Eq. (13) reads

$$E\left[\prod_{i=1}^{N} X_i\right] = \sum_{K \in \mathcal{A}_N} C_{U}^{K},$$

where $\mathcal{A}_N$ is the set of all $N \times N$ matrices that are i) symmetric, ii) have all diagonal elements zero, iii) their elements are either 0 or 1, iv) each row sum equals to one. Thus, matrices $K \in \mathcal{A}_N$ are identified [3, definition 2.1] as the adjacency matrices of unidirected 1-regular graphs between $N$ nodes. By virtue of the handshaking lemma [5, theorem 2.1], the number of nodes $N$ cannot be of the same parity with 1, which is the graph degree. Thus, for $N$ odd, the set $\mathcal{A}_N$ is empty and so the odd moments are zero. For even $N$, each unidirected 1-regular graph between $N$ nodes is equivalent to one partition of set $\{1, \ldots, N\}$ into unordered pairs, and so Eq. (16) is expressed equivalently as the branch of Eq. (15) for even $N$. □

2. **Proof of theorem 1 by mathematical induction**

Theorem 1 is proven by multidimensional mathematical induction on the multi-index of exponents $n$. For this proof, we introduce the multi-index $e^{(j)}$ which has all its components equal to zero, except for the $i$th component which is equal to one. Similarly, we introduce the index-matrix $E^{(j)}$ that has all its elements equal to zero, except for the $ij$ element which is equal to one. Last, index-matrix $E_{\text{sym}}^{(j)}$ is symmetric, having both $ij$ and $ji$ elements equal to one, and the rest of its elements are equal to zero.

**Base case: $n = e^{(j)}, i = 1, \ldots, N$.** For the base case, $E[g(X)X^n] = E[g(X)X_i]$. The index-matrices $L$ with $r(L) \leq e^{(j)}$ are: i) the zero matrix 0, ii) the matrices $E^{(j)}, j = 1, \ldots, N$. Matrix 0 results in the term $\mu X_i g(X)$ in the $L$-sum of Eq. (5). For each matrix $E^{(j)}, j = 1, \ldots, N$, we calculate that $\{E_i^{(j)}\}_{ij} = 1$, $K^{\text{max}}(E^{(j)}) = K = \tilde{K} = 0$, $H_{E^{(j)}} = 1$, $\tilde{L} = E_{\text{sym}}^{(j)}, C_{U}^{E_{\text{sym}}^{(j)}} = C_{ij}$, and $c(E^{(j)}) = e^{(j)}$. Thus, each index-matrix $E^{(j)}$ results in the term $C_{ij} E[\partial_j g(X)]$ in the $L$-sum. Summation of all the said terms results in multivariate Stein’s lemma, Eq. (4).

**Inductive hypothesis:** Eq. (5) holds true for $n$.
Inductive step: Prove that Eq. (5) holds true for $n + e^{(i)}$, $i = 1, \ldots, N$. By using the inductive hypothesis, we have

$$E[g(X)X^{n+e^{(i)}}] = E[(g(X)X)X^n] = \sum_{r|L| \leq n} \binom{n}{L} \mu^{n-r(L)} \sum_{K \subseteq K^{\text{sym}}(L)} H_{L,K} C_{U}^{L-K} E[\partial^{(L-K)} g(X)X].$$ (17)

By the general Leibniz rule [1, expression 3.3.8], we calculate the derivative

$$\partial^{(L-K)}_i (g(X)X) = \sum_{p=0}^{\frac{e^{(i)}}{K}} \binom{c_i (L-K)}{p} \left( \partial^{(L-K)-p} g(X) \right) \partial^p X_i.$$

(18)

Since $\partial^p X_i = X_i$, $\partial^p X_i = 1$, and $\partial^p X_i = 0$ for $p \geq 2$, Eq. (18) is simplified into

$$\partial^{(L-K)}_i (g(X)X) = X_i \partial^{(L-K)}_i g(X) + c_i (L-K) \partial^{(L-K)-1} g(X) = X_i \partial^{(L-K)}_i g(X) + \sum_{j=1}^{N} [\ell_{ji} - (1 + \delta_{ij})k_{ij}] \partial^{(L-K)-1} g(X),$$

(19)

where $\delta_{ij}$ is Kronecker’s delta. By using Eq. (19), we rewrite Eq. (17) as

$$E[(g(X)X)X^n] = A + \sum_{j=1}^{N} B_j,$$

(20)

with

$$A = \sum_{r|L| \leq n} \binom{n}{L} \mu^{n-r(L)} \sum_{K \subseteq K^{\text{sym}}(L)} H_{L,K} C_{U}^{L-K} E[X_i \partial^{(L-K)} g(X)].$$

(21)

and

$$B_j = \sum_{r|L| \leq n} \binom{n}{L} \mu^{n-r(L)} \sum_{K \subseteq K^{\text{sym}}(L-E^{(i)})} [\ell_{ji} - (1 + \delta_{ij})k_{ij}] H_{L,K} C_{U}^{L-K} E\left[\partial^{(L-K)-1} \prod_{p=1}^{N} \partial^{(L-K)} g(X)\right].$$

(22)

In Eq. (22), the term in $K$-sum is zero for $\ell_{ji} = 2k_{ji}$, or $\ell_{ji} = k_{ij}$ for $i \neq j$. In order to exclude zero terms, we update Eq. (22) to

$$B_j = \sum_{r|L| \leq n} \binom{n}{L} \mu^{n-r(L)} \sum_{K \subseteq K^{\text{sym}}(L-E^{(i)})} [\ell_{ji} - (1 + \delta_{ij})k_{ij}] H_{L,K} C_{U}^{L-K} E\left[\partial^{(L-K)-1} \prod_{p=1}^{N} \partial^{(L-K)} g(X)\right].$$

(23)

By performing the change of index-matrix $K' = K + E^{(i)}$, we recast Eq. (23) into

$$B_j = \sum_{r|L| \leq n} \binom{n}{L} \mu^{n-r(L)} \sum_{E^{(i)}_{\text{sym}} \subseteq K \subseteq K^{\text{sym}}(L-E^{(i)})+E^{(i)}_{\text{sym}}} [\ell_{ji} - (1 + \delta_{ij})(k_{ij} - 1)] H_{L,K-E^{(i)}_{\text{sym}}} C_{U}^{L-K+E^{(i)}_{\text{sym}}} E\left[\partial^{(L-K)-1} \prod_{p=1}^{N} \partial^{(L-K)} g(X)\right].$$

(24)

Since $\{(\ell_{ji} - 1)/2\} + 1 = \{(\ell_{ji} + 1)/2\}$, and $\min(\ell_{ji}, \ell_{ji} - 1) = \min(\ell_{ji} + 1, \ell_{ji})$ for $i \neq j$, it holds true that $K^{\text{max}}(L - E^{(i)}) + E^{(i)}_{\text{sym}} = K^{\text{max}}(L + E^{(i)})$. Thus, Eq. (24) is expressed equivalently as

$$B_j = \sum_{r|L| \leq n} \binom{n}{L} \mu^{n-r(L)} \sum_{E^{(i)}_{\text{sym}} \subseteq K^{\text{max}}(L+E^{(i)})} [\ell_{ji} - (1 + \delta_{ij})(k_{ij} - 1)] H_{L,K+E^{(i)}_{\text{sym}}} C_{U}^{L-K+E^{(i)}_{\text{sym}}} E\left[\partial^{(L-K)-1} \prod_{p=1}^{N} \partial^{(L-K)} g(X)\right].$$

(25)
By applying Stein’s lemma (4) at the average appearing in the right-hand side of Eq. (21), we obtain

\[ A = A_0 + \sum_{j=1}^{N} A_j \]  

(26)

with

\[
A_0 = \sum_{r,L \in n} \binom{n}{L} \mu^{r-r(L)} \sum_{K \leq K_{\text{max}}(L)} H_{L,K} C_U^{L-K} \mathbb{E} \left[ \varphi^{(L-K)} g(X) \right] 
\]  

(27)

and

\[
A_j = \sum_{r,L \in n} \binom{n}{L} \mu^{r-r(L)} \sum_{K \leq K_{\text{max}}(L)} H_{L,K} C_U^{L-K+E_{ij}^{(j)}} \mathbb{E} \left[ \varphi^{(L-K)} g(X) \right]. 
\]  

(28)

Thus, under Eqs. (20) and (26), the average is expressed as

\[
\mathbb{E} \left[ (g(X)X)^a \right] = A_0 + \sum_{j=1}^{N} (A_j + B_j). 
\]  

(29)

In order to evaluate each \( A_j + B_j \) further, we prove the following lemma.

**Lemma 1** (Recurrence relation for \( H_{L,K} \)). For \( L \in \mathbb{N}^{n \times N} \), \( K \in \mathbb{N}^{n \times N} \) symmetric with \( 0 \leq K \leq K_{\text{max}}(L + E^{(j)}) \), and under the convention that \( H_{L,K} = 0 \) for \( K < 0 \) or \( K > K_{\text{max}}(L) \), it holds true that

\[
H_{L+L',K+E_{ij}^{(j)}} = H_{L,K} + [\ell_{ji} - (1 + \delta_{ij})(k_{ij} - 1)]H_{L,K-E_{ij}^{(j)}}. 
\]  

(30)

**Proof.** See Appendix A. \( \square \)

Thus

\[
A_j + B_j = \sum_{r,L \in n} \binom{n}{L} \mu^{r-r(L)} \sum_{K \leq K_{\text{max}}(L + E^{(j)})} H_{L,E^{(j)}} C_U^{L-K+E_{ij}^{(j)}} \mathbb{E} \left[ \varphi^{(L-K)} g(X) \right]. 
\]  

(31)

and by performing the index-matrix change \( L' = L + E^{(j)} \)

\[
A_j + B_j = \sum_{r,L' \in n} \binom{n}{L'} \mu^{r-r(L')} \sum_{K \leq K_{\text{max}}(L')} H_{L',K} C_U^{L-K} \mathbb{E} \left[ \varphi^{(L-K)} g(X) \right]. 
\]  

(32)

By also considering \( A_0 \) from Eq. (27). Eq. (29) reads

\[
\mathbb{E} \left[ (g(X)X)^a \right] = \sum_{r,L \in n} \binom{n}{L} \mu^{r-r(L)} \sum_{K \leq K_{\text{max}}(L)} H_{L,K} C_U^{L-K} \mathbb{E} \left[ \varphi^{(L-K)} g(X) \right] + \sum_{r,L \in n} \binom{n}{L} \mu^{r-r(L)} \sum_{K \leq K_{\text{max}}(L)} H_{L,K} C_U^{L-K} \mathbb{E} \left[ \varphi^{(L-K)} g(X) \right]. 
\]  

(33)

The inductive proof of Eq. (5) is completed by the following lemma.

**Lemma 2** (Addition of \( r \)-multinomial coefficients). It holds true that

\[
\binom{n + e^{(j)}}{L} = \binom{n}{L} + \sum_{j=1}^{N} \binom{n}{L - E^{(j)}}, \text{ for } r(L) \leq n, \]

(34)

\[
\binom{n + e^{(j)}}{L} = \sum_{j=1}^{N} \binom{n}{L - E^{(j)}}, \text{ for } r(L) = n + e^{(j)}. \]

(35)
Appendix B

35

Appendix C

5

1

33

\[ g(x_1, x_2) = \sum_{\ell_1, \ell_2} G_{\ell_1, \ell_2, k} x_1^{\ell_1} x_2^{\ell_2} \]

where \( G_{\ell_1, \ell_2, k} \)'s are identified as the absolute values of the coefficients appearing in the 2-dimensional Hermite polynomials introduced by Itô [9, 8]; \( \text{He}_{\ell_1, \ell_2}(x_1, x_2) = \sum_{k=0}^{\min[\ell_1, \ell_2]} (-1)^k G_{\ell_1, \ell_2, k} x_1^{\ell_1-k} x_2^{\ell_2-k} \).
Proof. See Appendix E. \[\Box\]

By expressing the average \(E\left[g(X)\left(\prod_{i=1}^{N} X_i^n\right)\right]\) via definition 1, we understand that, for its evaluation, it suffices to sequentially apply operators \(T_{ji}, T_{ij}, i, j = 1, \ldots, N, j > i\) at the product \(g(X)\left(\prod_{i=1}^{N} X_i^n\right)\), and set \(x = \mu\) afterwards. After algebraic manipulations and using the operator properties of remark 1, we obtain

\[
E\left[g(X)\left(\prod_{i=1}^{N} X_i^n\right)\right] = \sum_{m_1, \ldots, m_N} \prod_{i=1}^{N} \left(\prod_{j=1}^{m_i} \ell_j, \ldots, \ell_n\right) \times \prod_{i=1}^{N} \sum_{k=0}^{m_i} H_{\ell_i k} \sigma_i^{2(\ell_i - k_i)} \prod_{j=1}^{N} \prod_{k=0}^{\min(\ell_j, k_j)} \sum_{k_{ij}} G_{\ell_{ij} k_{ij} k_{ji}} C_{(ij) t - t_{ij} - k_{ij}} \right] E\left[\prod_{i=1}^{N} \partial_{x_i}^n g(X)\right],
\]

where \(\left(m_1, \ell_1, \ldots, m_N\right)\) is the multinomial coefficient with \(N + 1\) factors, and the orders \(a_i\) of partial derivatives are

\[
a_i = \sum_{j=1}^{N} \left(\ell_{ji} - (1 + \delta_{ji}) k_{ij}\right), \quad \text{with} \quad k_{jj} = k_{ji}.
\]

Sum \(\sum_{m_1, \ldots, m_N, \ell_1, \ldots, N}\) is over all combinations of nonnegative integers \(\{m_1, \ell_1, \ldots, \ell_N\}_{i=1}^{N}\) with \(m_i + \sum_{j=1}^{N} \ell_{ij} = n_i, i = 1, \ldots, N\). By recasting Eq. (44) into multi-index and index matrix notation, we obtain Eq. (5).

4. Conclusions and future works

In the present work, we derived formula (5) generalizing Stein’s lemma for the evaluation of \(E\left[g(X)\prod_{i=1}^{N} X_i^n\right]\), where \(X\) is an \(N\)-dimensional Gaussian random vector. By our generalizing formula, the said average is expressed in terms of the averages of partial derivatives of \(g(X)\), as well as the mean value vector and autocovariance matrix of \(X\). Furthermore, by setting \(g(X) = 1\), generalizing formula (5) results in Isserlis theorem [10] and Song & Lee formula [26] for Gaussian product moments \(E\left[\prod_{i=1}^{N} X_i^n\right]\).

A direction for future works is the generalization of the infinite-dimensional analog of Stein’s lemma, called the Novikov-Furutsu theorem (see [24, Sec. 11.5, (2)]. In the infinite-dimensional case, \(X\) is a Gaussian random process of time argument \(t\), whose mean value is the function \(\mu(t)\), and its two-time autocovariance function is \(C(t_1, t_2)\). Thus, for \(g\) being a function of \(X\) over the time interval \([t_0, t]\), Novikov-Furutsu theorem reads:

\[
E\left[g[X]X(t)\right] = \mu(t)E\left[g(X)\right] + \int_{t_0}^{t} C(t, s)E\left[\frac{\delta g[X]}{\delta s}\right] ds,
\]

where \(\delta g[X]/\delta s(X)\) is the Volterra functional derivative of \(g[X]\) with respect to a local perturbation of process \(X\) centered at time \(s\) (see e.g. [2, Appendix A] for more on Volterra calculus). Novikov-Furutsu theorem, Eq. (46), is the main tool in deriving evolution equations, that resemble the classical Fokker-Planck equation, for the response probability density of dynamical systems under Gaussian random excitation, see e.g. [7, Eq.(3.19)], [16, 14]. Recently [15, Ch. 3], we extended Novikov-Furutsu theorem for averages that contain the Gaussian argument at various times; \(E\left[g[X]\prod_{i=1}^{N} X(t_i)\right]\). As we have already shown in [2], the introduction and use of averaged shift operators is very helpful in constructing generalizations of the Novikov-Furutsu theorem.

Acknowledgments

Part of this work was conducted while the author was a postdoctoral researcher in the North Carolina State University, Department of Mathematics.
Appendix A. Proof of lemma 1

By using the definition relation (11), we easily calculate that

\[ H_{L,E^{0/1}} = H_{L,0} = 1. \]  

(A.1)

Since, by convention, \( H_{L,E^{0/1}} = 0 \), Eq. (A.1) coincides with recurrence relation (30) for \( K = 0 \). For \( E^{0/1}_{\text{sym}} = K \leq K_{\text{max}}(L) \), we have the following two cases.

First case: \( i = j \)

\[
H_{L,K} + (\ell_{ii} - 2k_{ii} + 2)H_{L,K,E^{0/1}_{\text{sym}}} = \\
\frac{L!}{2 \sum_{p=1}^{N} k_{pp} \prod_{p \neq q(i) \neq (j), (j)} \prod_{p \neq q(i) \neq (j)}^N (\ell_{pq} - \tilde{k}_{pq})! \\
\left[ \frac{1}{2^{\ell_{ii} - 2k_{ii}}!(\ell_{ii} - 2k_{ii})!} + \frac{\ell_{ii} - 2k_{ii} + 2}{2^{\ell_{ii} - 2k_{ii} + 2}!(\ell_{ii} - 2k_{ii} + 2)!} \right] \\
\quad = L! + \left( E^{0/1}(L) \right)! \\
\quad = H_{L,E^{0/1}}. 
\]

(A.2)

Second case: \( i \neq j \)

\[
H_{L,K} + (\ell_{ji} - k_{ij} + 1)H_{L,K,E^{0/1}_{\text{sym}}} = \\
\frac{L!}{2 \sum_{p=1}^{N} k_{pp} \prod_{p \neq q(i) \neq (j)} \prod_{p \neq q(i) \neq (j)}^N (\ell_{pq} - \tilde{k}_{pq})!} \\
\times \left[ \frac{1}{k_{ij}(\ell_{ji} - k_{ij})!(\ell_{ji} - k_{ij})!} + \frac{\ell_{ji} - k_{ij} + 1}{(\ell_{ji} - k_{ij} + 1)(\ell_{ji} - k_{ij} + 1)!} \right] \\
\quad = L! + (L + E^{0/1}(L) - K)! \\
\quad = H_{L,E^{0/1}}. 
\]

(A.3)

Last, we have to prove Eq. (30) for \( K = K_{\text{max}}(L + E^{0/1}(L)). \) Again, we distinguish two cases.

First case: \( i = j \). If \( \ell_{ii} \) is even, \( \lfloor (\ell_{ii} + 1)/2 \rfloor = \lfloor \ell_{ii}/2 \rfloor \), and thus \( K_{\text{max}}(L + E^{0/1}(L)) = K_{\text{max}}(L) \). So, it remains to prove Eq. (30) for odd \( \ell_{ii} = 2a + 1 \) and for \( k_{ii} = k_{\text{max}}(L + E^{0/1}(L)) = \lfloor (\ell_{ii} + 1)/2 \rfloor \). In this case, we calculate

\[
\frac{H_{L,E^{0/1}}(1)}{H_{L,K_{\text{max}}(L + E^{0/1}(L))}^{(1)}} = \frac{2^{(2a+1)!}}{2^{(2a+1)!}} = 1. 
\]

(A.4)

Second case: \( i \neq j \). If \( \ell_{ij} + 1 > \ell_{ji}, K_{\text{max}}(L + E^{0/1}(L)) = K_{\text{max}}(L) \). Thus, it remains to prove Eq. (30) for \( \ell_{ij} + 1 < \ell_{ji} \), and for \( k_{ij} = k_{\text{max}}(L + E^{0/1}(L)) = \ell_{ij} + 1 \). In this case, we calculate

\[
\frac{H_{L,E^{0/1}}(1)}{H_{L,K_{\text{max}}(L + E^{0/1}(L))}^{(1)}} = \frac{(\ell_{ij} + 1)!}{(\ell_{ij} + 1)![(\ell_{ij} + 1)!(\ell_{ij} + 1)!]} = \ell_{ji} - \ell_{ij}. 
\]

(A.5)

Since, by convention, \( H_{L,K_{\text{max}}(L + E^{0/1}(L))} = 0 \) for this case, Eq. (A.5) is the specification of recurrence relation (30).

Second case: \( i \neq j \). If \( \ell_{ij} + 1 > \ell_{ji}, K_{\text{max}}(L + E^{0/1}(L)) = K_{\text{max}}(L) \). Thus, it remains to prove Eq. (30) for \( \ell_{ij} + 1 < \ell_{ji} \), and for \( k_{ij} = k_{\text{max}}(L + E^{0/1}(L)) = \ell_{ij} + 1 \). In this case, we calculate

\[
\frac{H_{L,E^{0/1}}(1)}{H_{L,K_{\text{max}}(L + E^{0/1}(L))}^{(1)}} = \frac{(\ell_{ij} + 1)!}{(\ell_{ij} + 1)![(\ell_{ij} + 1)!(\ell_{ij} + 1)!]} = \ell_{ji} - \ell_{ij}. 
\]

(A.6)

Since, by convention, \( H_{L,K_{\text{max}}(L + E^{0/1}(L))} = 0 \) for this case, Eq. (A.6) is the specification of recurrence relation (30).
Appendix B. Proof of lemma 2

First case: \( r(L) \leq n \). Using the definition relation (10) for \( r \)-multinomial coefficients, we have

\[
\binom{n}{L} + \sum_{j=1}^{n} \binom{n}{L} \binom{n-r(L)}{j} = \frac{n!}{[n-r(L)]!L!} + \sum_{j=1}^{n-r} \frac{n!}{[n-r(L)-E^{(j)}]!L!} = \\
= \frac{n!}{[n-r(L)]!L!} + \sum_{j=1}^{n} \frac{n!}{[n+e^{(0)}-r(L)]!L!] \left| n_j + 1 - r(L) + \sum_{j=1}^{N} \ell_{ij} \right| = \frac{[n+e^{(0)}]!}{[n+e^{(0)}-r(L)]!L!} = (n+e^{(0)})_r \tag{B.1}
\]

Second case: \( r(L) = n+e^{(0)} \). Using the definition relation (10), we have

\[
\sum_{j=1}^{n} (L-E^{(j)}) = \sum_{j=1}^{n} \frac{n!}{[n-r(L)-E^{(j)}]!L!} = \sum_{j=1}^{n} \frac{n!}{[n+e^{(0)}-r(L)]!L!] = \frac{n!(n+1)}{[n+e^{(0)}-r(L)]!L!} = \frac{[n+e^{(0)}]!}{[n+e^{(0)}-r(L)]!L!} = (n+e^{(0)})_r. \tag{B.2}
\]

Appendix C. Formal derivation of definition 1

The Taylor expansion of a \( C^\infty \left( \mathbb{R}^N \rightarrow \mathbb{R} \right) \) function \( g \) around \( x_0 \) is expressed via the shift pseudodifferential operator in exponential form (see e.g. [6, Sec. 1.1]) as

\[
g(x) = \left( 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \sum_{n=1}^{\infty} [n_1! \ldots n_m!] \sum_{i=1}^{n} \hat{x}_i \cdot \hat{x}_i \cdots \hat{x}_i \partial_{x} \cdots \partial_{x} \right) g(x_0) = \exp \left( \sum_{m=1}^{\infty} \frac{1}{m!} \left( \sum_{i=1}^{n} \hat{x}_i \partial_{x} \right)^m \right) g(x_0) = \exp \left( \sum_{i=1}^{n} \hat{x}_i \partial_{x} \right) g(x_0). \tag{C.1}
\]

where \( \hat{x} = x - x_0 \) is called the shift argument. By substituting the random vector \( X \) as the argument of function \( g \), choosing \( x_0 = \mu \) where \( \mu \) is the mean value of \( X \), and taking the average in both sides of Eq. (C.1) results into

\[
\mathbb{E}[g(X)] = \mathbb{E} \left[ \exp \left( \sum_{i=1}^{n} \hat{x}_i \partial_{x} \right) \right] g(\mu) = M_{\hat{X}}(\nabla) g(\mu), \tag{C.2}
\]

where \( \hat{X} := X - \mu \) (the centered random vector) and \( \nabla = [\partial_1, \ldots, \partial_N]^T \) (the del vector). In Eq. (C.2), \( M_{\hat{X}}(u) \) is identified as the moment-generating function of \( \hat{X} \); \( M_{\hat{X}}(u) = \mathbb{E} \left[ \exp \left( u^T \hat{X} \right) \right] \). For the Gaussian vector \( X \) with autocovariance matrix \( C \), the moment-generating function for the corresponding centered Gaussian random vector \( \hat{X} \) takes the form \( M_{\hat{X}}(u) = \exp \left( u^T Cu / 2 \right) \). Substitution of Gaussian \( M_{\hat{X}}(u) \) into Eq. (C.2) results in

\[
\mathbb{E}[g(X)] = \exp \left( \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} \partial_i \partial_j \right) g(\mu). \tag{C.3}
\]

and by using the symmetry property of autocovariance matrix \( C \):

\[
\mathbb{E}[g(X)] = \exp \left( \frac{\sigma^2}{2} \partial_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} C_{ij} \partial_i \partial_j \right) g(\mu) = \left[ \prod_{i=1}^{n} \exp \left( \frac{\sigma^2}{2} \partial_i^2 \right) \right] \left[ \prod_{i=1}^{n} \prod_{j=1}^{n} \exp \left( C_{ij} \partial_i \partial_j \right) \right] g(\mu). \tag{C.4}
\]

Eq. (C.4) coincides with Eq. (36).
Appendix D. Proof of lemma 3

Expressing $T_u[g(x)x_n^p]$ via Eq. (39) we have

$$T_u[g(x)x_n^p] = \sum_{m=0}^{\infty} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m}(g(x)x_n^p).$$

(D.1)

The derivatives appearing in the right-hand side of Eq. (D.1) are further evaluated using the general Leibniz rule:

$$\partial_i^{2m}(g(x)x_n^p) = \binom{n_i}{\ell} \sum_{m=0}^{\infty} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m-\ell}g(x) \partial_i^{\ell}x_n^p.$$  

(D.2)

Since $\partial_i^\ell x_n^p = (n_i!(n_i-\ell)!)x_n^{p-\ell}$ for $n_i \geq \ell$ and zero for $n_i < \ell$, Eq. (D.2) is updated to

$$\partial_i^{2m}(g(x)x_n^p) = \sum_{m=0}^{\infty} \binom{n_i}{\ell} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m-\ell}g(x) \partial_i^{\ell}x_n^p,$$

where $(2m)^! = (2m)(2m-1)\cdots(2m-\ell+1)$ is the falling factorial. Substitution of Eq. (D.3) into Eq. (D.1) results in

$$T_u[g(x)x_n^p] = \sum_{m=0}^{\infty} \sum_{\ell=0}^{\min(n_i,2m)} \binom{n_i}{\ell} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m-\ell}g(x) \partial_i^{\ell}x_n^p.$$  

(D.4)

In Eq. (D.4), the $m$ and $\ell$-summations are interchanged using formula (G.1), resulting into

$$T_u[g(x)x_n^p] = \sum_{\ell=0}^{\min(m,\ell)} \sum_{\ell=0}^{\infty} \binom{n_i}{\ell} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m-\ell}g(x) \partial_i^{\ell}x_n^p.$$  

(D.5)

Following [4, Sec. 8.4], see also [13, Eq.(30)], $(2m)^!$ is expressed in terms of $m^2$ as

$$(2m)^! = \sum_{p=\lceil\ell/2\rceil}^{\min(m,\ell)} C(\ell, p; 2) m^p!$$

(D.6)

where $C(\ell, p; 2)$ are the generalized factorial coefficients with parameter 2, and $\lceil \cdot \rceil$ is the ceiling function. Using also the fact $m^p! = 1/(m-p)!$, Eq. (D.5) is expressed as

$$T_u[g(x)x_n^p] = \sum_{\ell=0}^{\min(m,\ell)} \sum_{\ell=0}^{\infty} \binom{n_i}{\ell} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m-\ell}g(x) \partial_i^{\ell}x_n^p.$$  

(D.7)

By also interchanging the $m$ and $p$-summations using formula (G.7), we have

$$T_u[g(x)x_n^p] = \sum_{\ell=0}^{\min(m,\ell)} \sum_{\ell=0}^{\infty} \binom{n_i}{\ell} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m-\ell}g(x) \partial_i^{\ell}x_n^p.$$  

(D.8)

An index change in the $m$-sum, and the use of Eq. (39), results in

$$T_u[g(x)x_n^p] = \sum_{\ell=0}^{\min(m,\ell)} \sum_{\ell=0}^{\infty} \binom{n_i}{\ell} \frac{\sigma_i^{2m}}{2^m m!} \partial_i^{2m-\ell}g(x) \partial_i^{\ell}x_n^p.$$  

(D.9)

Last, we perform the change of index $k = \ell - p$ to obtain

$$T_u[g(x)x_n^p] = \sum_{k=0}^{\min(m,\ell)} \sum_{k=0}^{\infty} \frac{\binom{n_i}{(\ell-k)/2}}{2^{\ell-k} k!} \sigma_i^{2m+2\ell-2k} \partial_i^{2m}g(x).$$  

(D.10)

As we have showed in the recent work [13, lemma 2], $H_{\ell,k} = C(\ell, \ell - k; 2)/2^{\ell-k}$, for $\ell \in \mathbb{N}, k = 0, \ldots, [\ell/2]$. Under this, Eq. (D.10) coincides with Eq. (41).
Appendix E. Proof of lemma 4

Expressing $T_{ij}[g(x)x_i^{n_i}x_j^{n_j}]$ via Eq. (42) we have

$$T_{ij}[g(x)x_i^{n_i}x_j^{n_j}] = \sum_{m=0}^{\infty} C_m^{n_i} \frac{\partial^m}{\partial x_i^m} g(x)x_i^{n_i}x_j^{n_j}. \quad (E.1)$$

As in Appendix D, the derivatives $\partial^m_i$, $\partial^m_j$ in the right-hand side of Eq. (E.1) can be evaluated further using general Leibniz rule, resulting in

$$T_{ij}[g(x)x_i^{n_i}x_j^{n_j}] = \sum_{m=0}^{\infty} \sum_{\ell_1=0}^{\text{min}[n_1,m]} \sum_{\ell_2=0}^{\text{min}[n_2,m]} \sum_{\ell_3=0}^{\text{min}[n_3,m]} \frac{\partial^m_{\ell_1} \partial^m_{\ell_2} \partial^m_{\ell_3}}{m!} g(x). \quad (E.2)$$

By rearranging the summations in Eq. (E.2) using formula (G.10), we obtain

$$T_{ij}[g(x)x_i^{n_i}x_j^{n_j}] = \sum_{m=0}^{\infty} \sum_{\ell_1=0}^{\text{min}[n_1,m]} \sum_{\ell_2=0}^{\text{min}[n_2,m]} \sum_{\ell_3=0}^{\text{min}[n_3,m]} \sum_{\ell_4=0}^{\text{min}[n_4,m]} \frac{\partial^m_{\ell_1} \partial^m_{\ell_2} \partial^m_{\ell_3} \partial^m_{\ell_4}}{m!} g(x). \quad (E.3)$$

In order to evaluate the right-hand side of Eq. (E.3) further, the product of the two falling factorials $m^\ell_1 n^\ell_2$ has to be expressed in terms of falling factorials of $m$. This is performed by the following lemma.

**Lemma 5** (Product of two falling factorials of $m$). It holds true that

$$m^\ell_1 n^\ell_2 = \sum_{k=0}^{\min[\ell_1,\ell_2]} G_{\ell_1,\ell_2,k}m^{\ell_1+\ell_2-k}. \quad (E.4)$$

**Proof.** See Appendix F. \(\square\)

Since, by the definition of falling factorial, $m^{\ell_1+\ell_2-k}$ is zero for $\ell_1 + \ell_2 - k > m$, Eq. (E.4) is updated to

$$m^\ell_1 n^\ell_2 = \sum_{k=\text{max}(0,\ell_1+\ell_2-m)}^{\min[\ell_1,\ell_2]} G_{\ell_1,\ell_2,k}m^{\ell_1+\ell_2-k}. \quad (E.5)$$

Substitution of Eq. (E.5) into Eq. (E.3), and use of $m^{\ell_1+\ell_2-k}/m! = 1/(m - \ell_1 - \ell_2 + k)!$ results in

$$T_{ij}[g(x)x_i^{n_i}x_j^{n_j}] = \sum_{m=0}^{\infty} \sum_{\ell_1=0}^{\text{min}[n_1,m]} \sum_{\ell_2=0}^{\text{min}[n_2,m]} \sum_{\ell_3=0}^{\text{min}[n_3,m]} \sum_{\ell_4=0}^{\text{min}[n_4,m]} \sum_{k=\text{max}(0,\ell_1+\ell_2-m)}^{\min[\ell_1,\ell_2]} \frac{G_{\ell_1,\ell_2,k}m^{\ell_1+\ell_2-k}}{(m - \ell_1 - \ell_2 + k)!} \partial^m_{\ell_1} \partial^m_{\ell_2} \partial^m_{\ell_3} \partial^m_{\ell_4} g(x). \quad (E.6)$$

By interchanging $m$ and $k$-summations using formula (G.15), we have

$$T_{ij}[g(x)x_i^{n_i}x_j^{n_j}] = \sum_{m=0}^{\infty} \sum_{\ell_1=0}^{\text{min}[n_1,m]} \sum_{\ell_2=0}^{\text{min}[n_2,m]} \sum_{\ell_3=0}^{\text{min}[n_3,m]} \sum_{\ell_4=0}^{\text{min}[n_4,m]} \sum_{k=0}^{\text{min}[\ell_1,\ell_2]} \frac{G_{\ell_1,\ell_2,k}m^{\ell_1+\ell_2-k}}{m!} \partial^m_{\ell_1} \partial^m_{\ell_2} \partial^m_{\ell_3} \partial^m_{\ell_4} g(x). \quad (E.7)$$

An index change in the $m$-sum results in

$$T_{ij}[g(x)x_i^{n_i}x_j^{n_j}] = \sum_{m=0}^{\infty} \sum_{\ell_1=0}^{\text{min}[n_1,m]} \sum_{\ell_2=0}^{\text{min}[n_2,m]} \sum_{\ell_3=0}^{\text{min}[n_3,m]} \sum_{\ell_4=0}^{\text{min}[n_4,m]} \sum_{k=0}^{\text{min}[\ell_1,\ell_2]} \frac{G_{\ell_1,\ell_2,k}m^{\ell_1+\ell_2-k}}{m!} \partial^m_{\ell_1} \partial^m_{\ell_2} \partial^m_{\ell_3} \partial^m_{\ell_4} g(x). \quad (E.8)$$

Eq. (E.8), by virtue of Eq. (40), coincides with Eq. (42).
Appendix F. Proof of lemma 5

Our starting point for proving Eq. (E.4) is Vandermonde’s identity (see e.g. [4, example 3.6])

\[
\binom{m}{\ell_i} = \sum_{k=0}^{\ell_i} \binom{\ell_i}{k} \binom{m-\ell_i}{\ell_i-k}
\]

(F.1)

By multiplying both sides of Eq. (F.1) by \(\binom{m}{\ell_j}\), and after some algebraic manipulations, we obtain

\[
\binom{m}{\ell_j} \binom{m}{\ell_i} = \sum_{k=0}^{\ell_i} \binom{\ell_i}{k} \binom{m-\ell_j}{\ell_j-k} = \sum_{k=0}^{\ell_i} \binom{\ell_i + \ell_j - k}{k, \ell_i - k, \ell_j - k} \binom{m}{\ell_i + \ell_j - k}.
\]

(F.2)

where \(\binom{\ell_i + \ell_j - k}{k, \ell_i - k, \ell_j - k}\) is the multinomial coefficient with three factors. Since \(\binom{\ell_i + \ell_j - k}{k, \ell_i - k, \ell_j - k} = 0\) for \(k > \ell_i\) or \(k > \ell_j\), upper limit of \(k\)-sum in Eq. (F.2) is updated to \(\min(\ell_i, \ell_j)\). By also using the fact that \(m! = (\begin{array}{c} m \\ n \end{array})!\), see [4, Eq. (3.11)], Eq. (F.2) results in

\[
m^\ell_i = \sum_{k=0}^{\min(\ell_i, \ell_j)} G_{\ell_i, \ell_j, k} m^{\ell_i + \ell_j - k}
\]

(F.3)

where

\[G_{\ell_i, \ell_j, k} = \binom{\ell_i + \ell_j - k}{k, \ell_i - k, \ell_j - k} \frac{(\ell_i, \ell_j)!}{(\ell_i + \ell_j - k)!} \leq \binom{\ell_i}{\ell_j} k!.
\]

(F.4)

Eq. (F.3) coincides with Eq. (E.4), and Eq. (F.4) coincides with the definition relation (43) of \(G_{\ell_i, \ell_j, k}\), completing thus the proof of lemma 5.

Appendix G. Summation rearrangement formulas and their proofs

In this Appendix, we prove the formulas (G.1), (G.7), (G.10), (G.15) rearranging multiple summations, that are employed in Appendix D, Appendix E for the proofs of lemmata 3, 4 respectively.

\[
\sum_{m=0}^{\infty} \sum_{\ell=0}^{\infty} A_{m,\ell} = \sum_{n=0}^{\infty} \sum_{\ell=0}^{\infty} A_{m,\ell}.
\]

(G.1)

Proof. We distinguish the cases of even and odd \(n\). For \(n = 2p\), the left-hand side of Eq. (G.1) is split into

\[
\sum_{m=0}^{\infty} \sum_{\ell=0}^{2p} A_{m,\ell} = \sum_{m=0}^{2p} \sum_{\ell=0}^{2m} A_{m,\ell} + \sum_{m=p+1}^{2p} \sum_{\ell=0}^{\infty} A_{m,\ell}.
\]

(G.2)

In the right-hand side of Eq. (G.2), the sums of the second term are interchanged. The double summation of the first term is over the triangle \(0 \leq m \leq p, 0 \leq \ell \leq 2m\) which is rearranged into \(0 \leq \ell \leq 2p, \ell/2 \leq m \leq p\) and since \(m, \ell\) are integers; \(0 \leq \ell \leq 2p, [\ell/2] \leq m \leq p\). Thus:

\[
\sum_{m=0}^{\infty} \sum_{\ell=0}^{\infty} A_{m,\ell} = \sum_{m=0}^{2p} \sum_{\ell=0}^{p} A_{m,\ell} + \sum_{m=p+1}^{2p} \sum_{\ell=0}^{\infty} A_{m,\ell} = \sum_{m=0}^{2p} \sum_{\ell=0}^{p} A_{m,\ell} + \sum_{m=p+1}^{2p} \sum_{\ell=0}^{\infty} A_{m,\ell}.
\]

(G.3)

For \(n = 2p + 1\), the left-hand side of Eq. (G.1) is split into

\[
\sum_{m=0}^{\infty} \sum_{\ell=0}^{\infty} A_{m,\ell} = \sum_{m=0}^{p} \sum_{\ell=0}^{2m} A_{m,\ell} + \sum_{m=p+1}^{\infty} \sum_{\ell=0}^{2m} A_{m,\ell} = \sum_{m=0}^{p} \sum_{\ell=0}^{2m} A_{m,\ell} + \sum_{m=p+1}^{\infty} \sum_{\ell=0}^{2m} A_{m,\ell} + \sum_{m=p+1}^{\infty} \sum_{\ell=0}^{2p+1} A_{m,\ell}.
\]

(G.4)
In the rightmost side of Eq. (G.4), the double summations are rearranged as for Eq. (G.2), resulting into

\[
\sum_{m=0}^{\infty} \sum_{\ell=0}^{\min(2p+1,2m)} A_{m,\ell} = \sum_{\ell=0}^{2p} \sum_{m=\ell/2}^{\infty} A_{m,\ell} + \sum_{m=p+1}^{\infty} A_{m,2p+1}. \tag{G.5}
\]

Since \( p + 1 = \lfloor (2p + 1)/2 \rfloor \), we identify the single sum in the right-hand side of Eq. (G.5) as the \( \ell = 2p + 1 \) term of the double sum:

\[
\sum_{m=0}^{\infty} \sum_{\ell=0}^{\min(2p+1,2m)} A_{m,\ell} = \sum_{\ell=0}^{\infty} \sum_{m=\ell/2}^{\infty} A_{m,\ell}. \tag{G.6}
\]

Thus, the proof of Eq. (G.1) for both even and odd \( n \) is completed.

**Proof.** The left-hand side of Eq. (G.7) is split into

\[
\sum_{m=n}^{\infty} \sum_{k=n}^{\min(\ell,m)} A_{m,k} = \sum_{m=n}^{\ell} \sum_{k=n}^{m} A_{m,k} + \sum_{m=\ell+1}^{\infty} \sum_{k=n}^{\ell} A_{m,k}. \tag{G.8}
\]

In the right-hand side of Eq. (G.8), the sums of the second term are interchanged. The double summation of the first term is over the triangle \( n \leq m \leq \ell, n \leq k \leq m \) which is rearranged into \( n \leq k \leq \ell, k \leq m \leq \ell \). Thus:

\[
\sum_{m=n}^{\infty} \sum_{k=n}^{\min(\ell,m)} A_{m,k} = \sum_{k=n}^{\ell} \sum_{m=k}^{\ell} A_{m,k} + \sum_{k=n}^{\ell} \sum_{m=\ell+1}^{\infty} A_{m,k}, \tag{G.9}
\]

which completes the proof of Eq. (G.7).

\[
\sum_{m=0}^{\infty} \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} A_{m,\ell_1,\ell_2} = \sum_{m=0}^{n_1} \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2} + \sum_{m=n_1+1}^{n_2} \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2}. \tag{G.10}
\]

**Proof.** Without loss of generality, we assume that \( n_1 < n_2 \). Then, the left-hand side of Eq. (G.10) is split into

\[
\sum_{m=0}^{\infty} \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} A_{m,\ell_1,\ell_2} = \sum_{m=0}^{n_1} \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2} + \sum_{m=n_1+1}^{n_2} \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2}. \tag{G.11}
\]

In the first term of the right-hand side of Eq. (G.11), the summation is over \( 0 \leq m \leq n_1, 0 \leq \ell_1 \leq m, 0 \leq \ell_2 \leq m \), which can be rearranged into \( 0 \leq \ell_1 \leq n_1, 0 \leq \ell_2 \leq n_1 \), \( \max(\ell_1,\ell_2) \leq m \leq n_1 \). In the second term, the summation over \( n_1 + 1 \leq m \leq n_2, 0 \leq \ell_2 \leq m \) is rearranged into \( 0 \leq \ell_2 \leq n_2, \max(n_1 + 1, \ell_2) \leq m \leq n_2 \). Thus, Eq. (G.11) is expressed equivalently as

\[
\sum_{m=0}^{\infty} \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} A_{m,\ell_1,\ell_2} = \sum_{m=0}^{n_1} \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2} + \sum_{m=n_1+1}^{n_2} \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2}. \tag{G.12}
\]
The second term in the right-hand side of Eq. (G.12) is regrouped as

\[ \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} \sum_{m=0}^{m_{\max}} A_{m,\ell_1,\ell_2} = \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} A_{m,\ell_1,\ell_2} + \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} \sum_{m=0}^{m_{\max}} A_{m,\ell_1,\ell_2} = \]

\[ = \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} \sum_{m=0}^{m_{\max}} A_{m,\ell_1,\ell_2} = \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} \sum_{m=0}^{m_{\max}} A_{m,\ell_1,\ell_2}, \quad (G.13) \]

Substitution of Eq. (G.13) into Eq. (G.12) results into

\[ \sum_{m=0}^{\infty} \min(n_{1,m}) \min(n_{2,m}) \sum_{\ell_1=0}^{n_1} \sum_{\ell_2=0}^{n_2} A_{m,\ell_1,\ell_2} = \sum_{m=0}^{\infty} \min(n_{1,m}) \sum_{\ell_1=0}^{n_1} \sum_{m=0}^{m_{\max}} \sum_{\ell_2=0}^{n_2} A_{m,\ell_1,\ell_2}, \quad (G.14) \]

which coincides with Eq. (G.10).

**Proof.** Without the loss of generality, we assume that \( \ell_1 < \ell_2 \). Thus, Eq. (G.15) is simplified into

\[ \sum_{m=0}^{\infty} \min(n_{1,m}) \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2} = \sum_{m=0}^{\infty} \sum_{\ell_1=0}^{\ell_1} \sum_{\ell_2=0}^{\ell_2} A_{m,\ell_1,\ell_2}, \quad (G.16) \]

Then, the left-hand side of Eq. (G.16) is split into

\[ \sum_{m=\ell_2}^{\infty} \sum_{k=0}^{\ell_1} A_{m,k} = \sum_{m=\ell_2}^{\infty} \sum_{k=0}^{\ell_1} A_{m,k}, \quad (G.17) \]

In the right-hand side of Eq. (G.17), the sums of the second term are interchanged. The double summation of the first term is over the triangle \( \ell_2 \leq m \leq \ell_1 + \ell_2, \ell_1 + \ell_2 - m \leq k \leq \ell_1 \) which is rearranged into \( 0 \leq k \leq \ell_1, \ell_1 + \ell_2 - k \leq m \leq \ell_1 + \ell_2 \). Thus:

\[ \sum_{m=\ell_2}^{\infty} \sum_{k=0}^{\ell_1} A_{m,k} = \sum_{m=\ell_2}^{\infty} \sum_{k=0}^{\ell_1} A_{m,k} + \sum_{m=\ell_2}^{\infty} \sum_{k=0}^{\ell_1} A_{m,k}, \quad (G.18) \]

which coincides with Eq. (G.16).  

**References**

[1] Abramowitz M and Stegun I. A.: *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10th edition. Dover Publications, (1964, 1983 reprint).
[2] Athanassoulis G.A. and Mamis K.I.: Extensions of the Novikov-Furutsu theorem, obtained by using Volterra functional calculus. *Physica Scripta* **94**(11), (2019), 115217. doi:10.1088/1402-4896/ab10b5
[3] Biggs, N: *Algebraic Graph Theory*, Cambridge Mathematical Library (2nd ed.), Cambridge University Press, Cambridge UK, (1993).
[4] Charalambides, Ch. A.: *Enumerative combinatorics*. Chapman & Hall/CRC, Boca Raton, (2002).
[5] Chartrand, G.: *Introductory Graph Theory*, Dover Publications, (1969).
[6] Glaeske, H.-J., Puhl, M. and Skornik, K.A.: *Operational Calculus and Related Topics*. Chapman & Hall/CRC, Boca Raton, (2006).
[7] Hanggi, P. and Jung, P.: Colored Noise in Dynamical Systems, in *Advances in Chemical Physics* **89** (I. Prigogine and S.A. Rice eds.), John Wiley & Sons, New York, (1994), 239-326. doi:10.1002/9780470141489.ch4

15
