

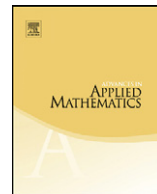


ELSEVIER

Contents lists available at SciVerse ScienceDirect

Advances in Applied Mathematics

www.elsevier.com/locate/yaama



Algebraic/combinatorial proofs of Cayley-type identities for derivatives of determinants and pfaffians

Sergio Caracciolo^a, Alan D. Sokal^{b,*}, Andrea Sportiello^{a,2}

^a Dipartimento di Fisica and INFN, Università degli Studi di Milano, via Celoria 16, I-20133 Milano, Italy

^b Department of Physics, New York University, 4 Washington Place, New York, NY 10003, USA

ARTICLE INFO

Article history:

Received 30 May 2011

Accepted 3 December 2012

Available online 26 January 2013

MSC:

primary 05A19

secondary 05E15, 05E99, 11S90, 13A50,

13N10, 14F10, 15A15, 15A23, 15A24,

15A33, 15A72, 15A75, 16S32, 20G05,

20G20, 32C38, 43A85, 81T18, 82B20

Keywords:

Cayley identity

Capelli identity

Determinant

Pfaffian

Bernstein–Sato polynomial

b-Function

Prehomogeneous vector space

Cayley operator

Omega operator

Omega process

Classical invariant theory

Grassmann algebra

Exterior algebra

Grassmann–Berezin integration

ABSTRACT

The classic Cayley identity states that

$$\det(\partial)(\det X)^s = s(s+1)\cdots(s+n-1)(\det X)^{s-1}$$

where $X = (x_{ij})$ is an $n \times n$ matrix of indeterminates and $\partial = (\partial/\partial x_{ij})$ is the corresponding matrix of partial derivatives. In this paper we present straightforward algebraic/combinatorial proofs of a variety of Cayley-type identities, both old and new. The most powerful of these proofs employ Grassmann algebra (= exterior algebra) and Grassmann–Berezin integration. Among the new identities proven here are a pair of “diagonal-parametrized” Cayley identities, a pair of “Laplacian-parametrized” Cayley identities, and the “product-parametrized” and “border-parametrized” rectangular Cayley identities.

© 2012 Elsevier Inc. All rights reserved.

* Corresponding author.

E-mail addresses: Sergio.Caracciolo@mi.infn.it (S. Caracciolo), sokal@nyu.edu (A.D. Sokal), Andrea.Sportiello@mi.infn.it

(A. Sportiello).

¹ Also at Department of Mathematics, University College London, London WC1E 6BT, England.

² Address after 1 October 2012: LIPN, UMR CNRS 7030, Université Paris-Nord, 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse, France.

1. Introduction

Let $X = (x_{ij})$ be an $n \times n$ matrix of indeterminates, and let $\partial = (\partial/\partial x_{ij})$ be the corresponding matrix of partial derivatives. The following beautiful identity is conventionally³ attributed to Arthur Cayley (1821–1895):

$$\det(\partial)(\det X)^s = s(s + 1) \cdots (s + n - 1)(\det X)^{s-1}. \tag{1.1}$$

[When $n = 1$ this is of course the elementary formula $(d/dx)x^s = sx^{s-1}$.] A generalization of (1.1) to arbitrary minors also holds, and is sometimes⁴ attributed to Alfredo Capelli (1855–1910): if $I, J \subseteq \{1, \dots, n\}$ with $|I| = |J| = k$, then

$$\det(\partial_{IJ})(\det X)^s = s(s + 1) \cdots (s + k - 1)(\det X)^{s-1} \epsilon(I, J)(\det X_{I^c J^c}) \tag{1.2}$$

where $\epsilon(I, J) = (-1)^{\sum_{i \in I} i + \sum_{j \in J} j}$. Analogous identities for symmetric and antisymmetric matrices were proved by Gårding [41] in 1948 and Shimura [89] in 1984, respectively.

Although these identities are essentially algebraic or combinatorial in nature, the simplest proofs currently available in the literature are analytic, exploiting Laplace-type integral representations for $(\det X)^s$ [36,89]. Indeed, most of the existing algebraic/combinatorial proofs [14,34,41,46,93,101,114] are somewhat difficult to follow, partly because of old-fashioned notation.⁵ In this paper we would like to give straightforward algebraic/combinatorial proofs of a variety of Cayley-type identities, some of which are known and others of which are new. The most powerful of these proofs employ Grassmann algebra (= exterior algebra) and Grassmann–Berezin integration.⁶

Nowadays, identities like (1.1) are best understood as calculations of Bernstein–Sato type [9,32,64] for special polynomials. To see what is at issue, let $P(x_1, \dots, x_n) \neq 0$ be a polynomial in n variables with coefficients in a field K of characteristic 0. Then Bernstein [7] proved in 1972 that there exist a polynomial-coefficient partial differential operator $Q(s, x, \partial/\partial x)$ and a polynomial $b(s) \neq 0$ (both with coefficients in K) satisfying

$$Q(s, x, \partial/\partial x)P(x)^s = b(s)P(x)^{s-1}. \tag{1.3}$$

We call any pair (Q, b) satisfying (1.3) a *Bernstein–Sato pair* for P .

The set of all b for which there exists a Q satisfying (1.3) is easily seen to be an ideal in the polynomial ring $K[s]$. By Bernstein’s theorem this ideal is nontrivial, so it is generated by a unique monic polynomial $b(s)$, called the *Bernstein–Sato polynomial* (or *b-function*) of P .⁷ Cayley-type identities thus provide Bernstein–Sato pairs for certain polynomials P arising from determinants.

Bernstein–Sato pairs are especially useful in treating the problem of analytically continuing the distribution P^s_Ω , which can be posed as follows [5,7–9,32,42,43,52,64]: Let $P(x_1, \dots, x_n) \neq 0$ be a polynomial with real coefficients, and let $\Omega \subseteq \mathbb{R}^n$ be an open set such that $P \geq 0$ on Ω and $P = 0$ on $\partial\Omega$. Then, for any complex number s satisfying $\text{Re } s > 0$, the function P^s is well defined on Ω and polynomially bounded, and thus defines a tempered distribution $P^s_\Omega \in \mathcal{S}'(\mathbb{R}^n)$ by the formula

³ But erroneously: see Section 2.6 below.

⁴ Also erroneously: see again Section 2.6.

⁵ Among the exceptions are [14, Theorem 1.3 and Lemma 2.12] and [34, Lemma 2.1].

⁶ We stress that our proofs are *not* “combinatorial” in the narrow sense of exhibiting bijections or exploiting double-counting. Rather, our proofs are based on straightforward algebraic manipulation combined with some elementary arguments of enumerative combinatorics (enumeration of permutations by number of cycles, identities involving binomial coefficients, etc.). For this reason we have opted to describe our methods using the rather awkward adjective “algebraic/combinatorial”.

⁷ In the literature on Bernstein–Sato equations it is customary to shift our s by 1, i.e. write $Q(s, x, \partial/\partial x)P(x)^{s+1} = b(s)P(x)^s$, so that the usual Bernstein–Sato polynomial is our $b(s + 1)$. We choose here the slightly unconventional notation (1.3) because it seems better adapted to the Cayley identity (1.1).

$$\langle P_{\Omega}^s, \varphi \rangle = \int_{\Omega} P(x)^s \varphi(x) dx \quad (1.4)$$

for any test function $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Furthermore, the function $s \mapsto \langle P_{\Omega}^s, \varphi \rangle$ is analytic on the half-plane $\operatorname{Re} s > 0$, with complex derivative given by

$$\frac{d}{ds} \langle P_{\Omega}^s, \varphi \rangle = \int_{\Omega} P(x)^s (\log P(x)) \varphi(x) dx. \quad (1.5)$$

Thus P_{Ω}^s is a distribution-valued analytic function of s on the right half-plane. We want to know whether P_{Ω}^s can be analytically continued to the whole complex plane as a meromorphic function of s . This problem was first posed by I.M. Gel'fand [42] at the 1954 International Congress of Mathematicians. It was answered affirmatively in 1969 independently by Bernstein and S.I. Gel'fand [8] and Atiyah [5], using deep results from algebraic geometry (Hironaka's resolution of singularities [49]). A few years later, Bernstein [7] produced a much simpler proof based on using the Bernstein–Sato equation (1.3) to extend P_{Ω}^s successively to half-planes $\operatorname{Re} s > -1$, $\operatorname{Re} s > -2$, etc. See e.g. [9, Sections 7.1 and 7.3] or [52, Section 5.3] for details.

The special case in which P is a determinant of a symmetric or hermitian matrix (and Ω is e.g. the cone of positive-definite matrices) has been studied by several authors [10,11,36,72–74,80,81]; it plays a central role in the theory of Riesz distributions on Euclidean Jordan algebras (or equivalently on symmetric cones) [36, Chapter VII]. This case is also useful in quantum field theory in studying the analytic continuation of Feynman integrals to “complex space–time dimension” [10,35,91]. In an analogous way, the parametrized symmetric Cayley identity (Theorem 2.11 below) will play a key role in studying the analytic continuation of integrals over products of spheres $S^{N-1} \subset \mathbb{R}^N$ to “complex dimension N ” [23], with the aim of giving a rigorous nonperturbative formulation of the correspondence found in [20,21] between spanning forests and the N -vector model in statistical mechanics at $N = -1$. This latter application was, in fact, our original motivation for studying Cayley-type identities. The original Cayley identity (1.1) was also rediscovered by Creutz [33] and used by him to compute certain invariant integrals over $SU(n)$ that arise in lattice gauge theory.

Many of the polynomials P treated here can also be understood as relative invariants of prehomogeneous vector spaces [52,60]. When applicable, this connection allows the immediate identification of a suitable operator $Q(\partial/\partial x)$ – namely, the dual of P itself – and provides a general proof that the corresponding $b(s)$ satisfies $\deg b = \deg P$ and is indeed (up to a constant factor) the Bernstein–Sato polynomial of P .⁸ Furthermore, this approach sometimes allows the explicit calculation of $b(s)$ by means of microlocal calculus [59,77,78,85,86,97,98,104,107] or other methods [96].⁹

The purpose of the present paper is to give straightforward (and we hope elegant) algebraic/combinatorial proofs of a variety of Cayley-type identities, both old and new. Since our main aim is to illustrate proof techniques that may be useful in other contexts, we shall give, wherever possible, several alternate proofs of each result. One purpose of this paper is, in fact, to make propaganda among mathematicians for the power of Grassmann–Berezin integration as a tool for proving algebraic or combinatorial identities. Among the new results in this paper are the “diagonal-parametrized” Cayley identities (Theorems 2.10 and 2.11), the “Laplacian-parametrized” Cayley identities (Theorems 2.12 and 2.14), and the “product-parametrized” and “border-parametrized” rectangular Cayley identities (Theorems 2.16 and 2.17). We also give an elementary (though rather intricate) proof of the multi-matrix rectangular Cayley identity (Theorem 2.9) that was proven recently by Sugiyama [98].

The plan of this paper is as follows: In Section 2 we state the identities to be proven and briefly discuss their interpretation. In Section 3 we give elementary algebraic/combinatorial proofs of the

⁸ See [52, Corollary 6.1.1 and Theorem 6.1.1], [60, Proposition 2.22] for the first two points, and [52, Theorem 6.3.2] for the third.

⁹ We are grateful to Nero Budur for explaining to us the connection between our results and the theory of prehomogeneous vector spaces.

three basic Cayley-type identities (ordinary, symmetric and antisymmetric). In Section 4 we give very simple proofs of these same identities, based on representing $(\det X)^s$ as a fermionic or bosonic Gaussian integral.¹⁰ In Section 5 we give alternate (and arguably even simpler) proofs, based on representing $\det(\partial)$ as a fermionic Gaussian integral; this method is very powerful and allows us to prove also the (considerably more difficult) “rectangular Cayley identities”. In Section 6 we prove the diagonal-parametrized Cayley identities, in Section 7 we prove the Laplacian-parametrized Cayley identities, and in Section 8 we prove the product-parametrized and border-parametrized rectangular Cayley identities. Finally, in Section 9 we formulate some conjectures concerning the minimality of our Bernstein–Sato pairs. In Appendix A we provide a brief introduction to Grassmann algebra and Grassmann–Berezin (fermionic) integration; we hope that this appendix will prove useful to mathematicians seeking a mathematically rigorous (but conceptually elementary) presentation of this powerful algebraic/combinatorial tool.¹¹ In Appendix B we collect some formulae that will be needed in the proofs.

We have tried hard to write this paper in a “modular” fashion, so that the reader can skip around according to his/her interests without having to read the whole thing. Indeed, after a brief perusal of Section 2, the reader can proceed directly to Section 3, 4 or 5 as desired, consulting Appendices A and B as needed.

2. Statement of main results

Notation. We write $[n] = \{1, \dots, n\}$. Give a matrix A , we denote its transpose by A^T . For an invertible square matrix A , we use the shorthand A^{-T} for $(A^{-1})^T$ (which is also equal to $(A^T)^{-1}$). If $A = (a_{ij})_{i,j=1}^n$ is an $n \times n$ matrix and $I, J \subseteq [n]$, we denote by A_{IJ} the submatrix of A corresponding to the rows I and the columns J , all kept in their original order. We write I^c to denote the complement of I in $[n]$. We define $\epsilon(I) = (-1)^{|I|(|I|-1)/2} (-1)^{\sum_{i \in I} i}$; it is the sign of the permutation that takes the sequence $1 \dots n$ into II^c when the sets I and I^c are each written in increasing order. We also define $\epsilon(I, J) = \epsilon(I)\epsilon(J)$; it is the sign of the permutation that takes II^c into JJ^c . In particular, if $|I| = |J|$, we have $\epsilon(I, J) = (-1)^{\sum_{i \in I} i + \sum_{j \in J} j}$.

2.1. Ordinary, symmetric, antisymmetric and hermitian Cayley identities

The basic Cayley-type identity is the following:

Theorem 2.1 (Ordinary Cayley identity). *Let $X = (x_{ij})$ be an $n \times n$ matrix of indeterminates, and let $\partial = (\partial/\partial x_{ij})$ be the corresponding matrix of partial derivatives. Then*

$$\det(\partial)(\det X)^s = s(s + 1) \cdots (s + n - 1)(\det X)^{s-1}. \tag{2.1}$$

More generally, if $I, J \subseteq [n]$ with $|I| = |J| = k$, then

$$\det(\partial_{IJ})(\det X)^s = s(s + 1) \cdots (s + k - 1)(\det X)^{s-1} \epsilon(I, J)(\det X_{I^c J^c}). \tag{2.2}$$

Remark. Since $\det(\partial)$ and $\det(\partial_{IJ})$ are constant-coefficient differential operators, the matrix X can everywhere be replaced by $X + A$ for any fixed matrix A , and the identities (2.1)/(2.2) remain valid.

¹⁰ The proofs based on representing $(\det X)^s$ as a bosonic Gaussian integral are closely related to the existing analytic proofs [36,89]. The proofs using fermionic Gaussian integrals, by contrast, are really algebraic/combinatorial proofs, as fermionic “integration” is a purely algebraic/combinatorial construction (see Appendix A.4 below).

¹¹ Unfortunately, most of the existing presentations of Grassmann–Berezin integration are aimed at physicists (see e.g. [117, Chapter 1] for an excellent treatment) and may not meet mathematicians’ standards of precision and rigor, or else are aimed at applications to differential geometry and hence involve heavier theoretical machinery than is needed for combinatorial applications. One exception is the brief summary given by Abdesselam [2, Section 2]; our presentation can be viewed as an enlargement of his.

We consider next a version of the Cayley identity for *symmetric* matrices $X^{\text{sym}} = (x_{ij})$. What this means is that only the variables $(x_{ij})_{1 \leq i \leq j \leq n}$ are taken as independent indeterminates; then x_{ij} for $i > j$ is regarded as a synonym for x_{ji} .

Theorem 2.2 (*Symmetric Cayley identity*). Let $X^{\text{sym}} = (x_{ij})$ be an $n \times n$ symmetric matrix of indeterminates, and let ∂^{sym} be the matrix whose elements are

$$(\partial^{\text{sym}})_{ij} = \frac{1}{2}(1 + \delta_{ij}) \frac{\partial}{\partial x_{ij}} = \begin{cases} \partial/\partial x_{ii} & \text{if } i = j, \\ \frac{1}{2} \partial/\partial x_{ij} & \text{if } i < j, \\ \frac{1}{2} \partial/\partial x_{ji} & \text{if } i > j. \end{cases} \tag{2.3}$$

Then

$$\det(\partial^{\text{sym}})(\det X^{\text{sym}})^s = s \left(s + \frac{1}{2}\right) \cdots \left(s + \frac{n-1}{2}\right) (\det X^{\text{sym}})^{s-1}. \tag{2.4}$$

More generally, if $I, J \subseteq [n]$ with $|I| = |J| = k$, then

$$\det(\partial_{IJ}^{\text{sym}})(\det X^{\text{sym}})^s = s \left(s + \frac{1}{2}\right) \cdots \left(s + \frac{k-1}{2}\right) (\det X^{\text{sym}})^{s-1} \epsilon(I, J) (\det X_{I^c J^c}^{\text{sym}}). \tag{2.5}$$

Remarks. 1. The matrix X^{sym} can everywhere be replaced by $X^{\text{sym}} + A$ for any fixed *symmetric* matrix A .

2. If we prefer to work over the integers rather than the rationals, it suffices to multiply ∂^{sym} by 2 and correspondingly multiply the right-hand side of (2.4) [resp. (2.5)] by 2^n (resp. 2^k).

Next let us state a version of the Cayley identity for *antisymmetric* matrices $X^{\text{antisym}} = (x_{ij})$.¹² Here only the variables $(x_{ij})_{1 \leq i < j \leq n}$ are taken as independent indeterminates; then x_{ij} for $i > j$ is regarded as a synonym for $-x_{ji}$, and x_{ii} is regarded as a synonym for 0. As befits antisymmetric matrices, the corresponding identity involves pfaffians in place of determinants:

Theorem 2.3 (*Antisymmetric Cayley identity*). Let $X^{\text{antisym}} = (x_{ij})$ be a $2m \times 2m$ antisymmetric matrix of indeterminates, and let $\partial^{\text{antisym}}$ be the corresponding matrix of partial derivatives, i.e.

$$(\partial^{\text{antisym}})_{ij} = \begin{cases} 0 & \text{if } i = j, \\ \partial/\partial x_{ij} & \text{if } i < j, \\ -\partial/\partial x_{ji} & \text{if } i > j. \end{cases} \tag{2.6}$$

Then

$$\text{pf}(\partial^{\text{antisym}})(\text{pf } X^{\text{antisym}})^s = s(s + 2) \cdots (s + 2m - 2) (\text{pf } X^{\text{antisym}})^{s-1}. \tag{2.7}$$

More generally, if $I \subseteq [2m]$ with $|I| = 2k$, then

$$\text{pf}(\partial_I^{\text{antisym}})(\text{pf } X^{\text{antisym}})^s = s(s + 2) \cdots (s + 2k - 2) (\text{pf } X^{\text{antisym}})^{s-1} \epsilon(I) (\text{pf } X_{I^c}^{\text{antisym}}). \tag{2.8}$$

¹² More precisely, *alternating* matrices, i.e. matrices satisfying $x_{ij} = -x_{ji}$ and $x_{ii} = 0$. The latter identity is a consequence of the former whenever the underlying ring of coefficients is an integral domain of characteristic $\neq 2$ (so that $2x = 0$ implies $x = 0$), but not in general otherwise. See e.g. [65, Sections XIII.6 and XV.9]. In this paper we use the term “antisymmetric” to denote $x_{ij} = -x_{ji}$ and $x_{ii} = 0$.

As an immediate corollary we get a result for antisymmetric determinants:

Corollary 2.4 (*Antisymmetric Cayley identity for determinants*). Let $X^{\text{antisym}} = (x_{ij})$ be a $2m \times 2m$ antisymmetric matrix of indeterminates, and let $\partial^{\text{antisym}}$ be the corresponding matrix of partial derivatives. Then

$$\det(\partial^{\text{antisym}})(\det X^{\text{antisym}})^s = (2s - 1)(2s) \cdots (2s + 2m - 2)(\det X^{\text{antisym}})^{s-1}. \tag{2.9}$$

More generally, if $I \subseteq [2m]$ with $|I| = 2k$, then

$$\begin{aligned} \det(\partial_{II}^{\text{antisym}})(\det X^{\text{antisym}})^s \\ = (2s - 1)(2s) \cdots (2s + 2k - 2)(\det X^{\text{antisym}})^{s-1} (\det X_{I^c I^c}^{\text{antisym}}). \end{aligned} \tag{2.10}$$

Please note that in the antisymmetric case we are able at present to handle only principal minors, i.e. we have been unable to find a general formula for $\det(\partial_{IJ}^{\text{antisym}})(\det X^{\text{antisym}})^s$ when $I \neq J$.

Next we state a version of the Cayley identity for “hermitian” matrices $Z^{\text{herm}} = (z_{ij})$.¹³ By this we mean the following: We introduce indeterminates $(x_{ij})_{1 \leq i \leq j \leq n}$ and $(y_{ij})_{1 \leq i < j \leq n}$, and define matrices X^{sym} and Y^{antisym} as before; we then set $Z^{\text{herm}} = X^{\text{sym}} + iY^{\text{antisym}}$ and $\partial^{\text{herm}} = \partial_X^{\text{sym}} - (i/2)\partial_Y^{\text{antisym}}$. In terms of the usual complex derivatives

$$\frac{\partial}{\partial z_{ij}} = \frac{1}{2} \left(\frac{\partial}{\partial x_{ij}} - i \frac{\partial}{\partial y_{ij}} \right), \quad \frac{\partial}{\partial \bar{z}_{ij}} = \frac{1}{2} \left(\frac{\partial}{\partial x_{ij}} + i \frac{\partial}{\partial y_{ij}} \right) \tag{2.11}$$

for $i < j$, this can be written as

$$(\partial^{\text{herm}})_{ij} = \begin{cases} \partial/\partial x_{ii} & \text{if } i = j, \\ \partial/\partial z_{ij} & \text{if } i < j, \\ \partial/\partial \bar{z}_{ji} & \text{if } i > j. \end{cases} \tag{2.12}$$

We then have:

Theorem 2.5 (“Hermitian” Cayley identity). Let Z^{herm} and ∂^{herm} be defined as above. Then

$$\det(\partial^{\text{herm}})(\det Z^{\text{herm}})^s = s(s + 1) \cdots (s + n - 1)(\det Z^{\text{herm}})^{s-1}. \tag{2.13}$$

More generally, if $I, J \subseteq [n]$ with $|I| = |J| = k$, then

$$\det(\partial_{IJ}^{\text{herm}})(\det Z^{\text{herm}})^s = s(s + 1) \cdots (s + k - 1)(\det Z^{\text{herm}})^{s-1} \epsilon(I, J)(\det Z_{I^c J^c}^{\text{herm}}). \tag{2.14}$$

The resemblance of this theorem to the ordinary Cayley identity (2.1)/(2.2) is no accident; indeed, the two identities are immediately interderivable. To see this, it suffices to notice that the action of derivatives on indeterminates is identical in the two cases:

$$\partial_{ij} X_{i'j'} = \delta_{i,i'} \delta_{j,j'}, \tag{2.15a}$$

$$\partial_{ij}^{\text{herm}} Z_{i'j'}^{\text{herm}} = \delta_{i,i'} \delta_{j,j'}, \tag{2.15b}$$

¹³ We put “hermitian” in quotation marks because the variables (x_{ij}) and (y_{ij}) in this identity are neither real nor complex numbers, but are simply indeterminates.

and that this relation completely determines the expressions (2.1)/(2.2) and (2.13)/(2.14). For this reason, we need not consider further the “hermitian” Cayley identity.¹⁴

2.2. Rectangular Cayley identities

Let us now formulate some Cayley-type identities for rectangular matrices of size $m \times n$ with $m \leq n$. These identities are somewhat more complicated than the preceding ones, because the matrices appearing in the determinants (or pfaffians) are quadratic (or of higher order) rather than linear in the indeterminates.

Theorem 2.6 (Two-matrix rectangular Cayley identity). *Let $X = (x_{ij})$ and $Y = (y_{ij})$ be $m \times n$ matrices of indeterminates with $m \leq n$, and let $\partial_X = (\partial/\partial x_{ij})$ and $\partial_Y = (\partial/\partial y_{ij})$ be the corresponding matrices of partial derivatives. Then*

$$\det(\partial_X \partial_Y^T) \det(XY^T)^s = \left(\prod_{j=0}^{m-1} (s+j)(s+n-m+j) \right) \det(XY^T)^{s-1}. \tag{2.16}$$

More generally, if $I, J \subseteq [m]$ with $|I| = |J| = k$, then

$$\begin{aligned} \det[(\partial_X \partial_Y^T)_{IJ}] \det(XY^T)^s &= \left(\prod_{j=0}^{k-1} (s+j)(s+n-m+j) \right) \det(XY^T)^{s-1} \\ &\times \epsilon(I, J) \det[(XY^T)_{J^c J^c}]. \end{aligned} \tag{2.17}$$

If $m = 1$, Theorem 2.6 reduces to the easily-derived formula $(\nabla_{\mathbf{x}} \cdot \nabla_{\mathbf{y}})(\mathbf{x} \cdot \mathbf{y})^s = s(s+n-1)(\mathbf{x} \cdot \mathbf{y})^{s-1}$ for $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{y} = (y_1, \dots, y_n)$. If $m = n$, Theorem 2.6 can be derived by separate applications of Theorem 2.1 to X and Y . In other cases it appears to be new.

Theorem 2.7 (One-matrix rectangular symmetric Cayley identity). *Let $X = (x_{ij})$ be an $m \times n$ matrix of indeterminates with $m \leq n$, and let $\partial = (\partial/\partial x_{ij})$ be the corresponding matrix of partial derivatives. Then*

$$\det(\partial \partial^T) \det(XX^T)^s = \left(\prod_{j=0}^{m-1} (2s+j)(2s+n-m-1+j) \right) \det(XX^T)^{s-1}. \tag{2.18}$$

More generally, if $I, J \subseteq [m]$ with $|I| = |J| = k$, then

$$\begin{aligned} \det[(\partial \partial^T)_{IJ}] \det(XX^T)^s &= \left(\prod_{j=0}^{k-1} (2s+j)(2s+n-m-1+j) \right) \det(XX^T)^{s-1} \\ &\times \epsilon(I, J) \det[(XX^T)_{J^c J^c}]. \end{aligned} \tag{2.19}$$

If $m = 1$, Theorem 2.7 reduces to the well-known formula $\Delta(\mathbf{x}^2)^s = 2s(2s+n-2)(\mathbf{x}^2)^{s-1}$ for $\mathbf{x} = (x_1, \dots, x_n)$. If $m = n$, Theorem 2.7 can be derived from two applications of Theorem 2.1. The

¹⁴ Let us remark that, by contrast, in the analytic proofs [36,89] it is convenient to consider the hermitian case instead of the ordinary case, as $(\det Z)^s$ for a positive-definite complex hermitian matrix Z has a simple Laplace-type integral representation.

general case has been proven recently by several authors, using analytic methods.¹⁵ We call Theorem 2.7 a “symmetric” identity because the matrices XX^T and $\partial\partial^T$ that appear in it are symmetric by construction.

Finally, here is an analogue of Theorem 2.7 that involves matrices that are *antisymmetric* rather than symmetric by construction; in place of the identity matrix we use the standard $2n \times 2n$ symplectic form

$$J = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & 0 & 1 & \\ & & -1 & 0 & \\ & & & & \ddots \end{pmatrix}. \tag{2.20}$$

Not surprisingly, this formula involves pfaffians rather than determinants:

Theorem 2.8 (One-matrix rectangular antisymmetric Cayley identity). *Let $X = (x_{ij})$ be a $2m \times 2n$ matrix of indeterminates with $m \leq n$, and let $\partial = (\partial/\partial x_{ij})$ be the corresponding matrix of partial derivatives. Then*

$$\text{pf}(\partial J \partial^T) \text{pf}(X J X^T)^s = \left(\prod_{j=0}^{m-1} (s + 2j)(s + 2n - 2m + 1 + 2j) \right) \text{pf}(X J X^T)^{s-1}. \tag{2.21}$$

More generally, if $I \subseteq [2m]$ with $|I| = 2k$, then

$$\begin{aligned} &\text{pf}[(\partial J \partial^T)_{II}] \text{pf}(X J X^T)^s \\ &= \left(\prod_{j=0}^{k-1} (s + 2j)(s + 2n - 2m + 1 + 2j) \right) \text{pf}(X J X^T)^{s-1} \epsilon(I) \text{pf}[(X J X^T)_{I^c I^c}]. \end{aligned} \tag{2.22}$$

If $m = n$, (2.21) reduces to the ordinary Cayley identity (2.1) of size $2m$ since $\text{pf}(A J A^T) = \det A$.

These three rectangular Cayley identities are roughly analogous to the ordinary, symmetric and antisymmetric Cayley identities, respectively, but their proofs are more intricate.

Finally, here is a generalization of Theorems 2.1 and 2.6 to an arbitrary number ℓ of rectangular matrices of arbitrary compatible sizes:

Theorem 2.9 (Multi-matrix rectangular Cayley identity). *Fix integers $\ell \geq 1$ and $n_1, \dots, n_\ell \geq 0$ and write $n_{\ell+1} = n_1$. For $1 \leq \alpha \leq \ell$, let $X^{(\alpha)}$ be an $n_\alpha \times n_{\alpha+1}$ matrix of indeterminates, and let $\partial^{(\alpha)}$ be the corresponding matrix of partial derivatives. Then*

$$\begin{aligned} &\det(\partial^{(1)} \dots \partial^{(\ell)}) \det(X^{(1)} \dots X^{(\ell)})^s \\ &= \left(\prod_{\alpha=1}^{\ell} \prod_{j=0}^{n_\alpha-1} (s + n_\alpha - n_{\alpha+1} + j) \right) \det(X^{(1)} \dots X^{(\ell)})^{s-1}. \end{aligned} \tag{2.23}$$

¹⁵ See Faraut and Korányi [36, Section XVI.4], Khékalo [57,58] and Rubín [83].

More generally, if $I, J \subseteq [n_1]$ with $|I| = |J| = k$, then

$$\det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}] \det(X^{(1)} \dots X^{(\ell)})^s = \left(\prod_{\alpha=1}^{\ell} \prod_{j=0}^{k-1} (s + n_{\alpha} - n_1 + j) \right) \det(X^{(1)} \dots X^{(\ell)})^{s-1} \epsilon(I, J) \det[(X^{(1)} \dots X^{(\ell)})_{J^c J^c}]. \tag{2.24}$$

We expect that there will exist “symmetric” and “antisymmetric” variants of Theorem 2.9, for matrix products of the form $X^{(1)} \dots X^{(\ell)} X^{(\ell)T} \dots X^{(1)T}$ or $X^{(1)} \dots X^{(\ell)} X^{(\ell+1)} X^{(\ell)T} \dots X^{(1)T}$ (with $X^{(\ell+1)}$ square and symmetric) for the symmetric case, and $X^{(1)} \dots X^{(\ell)} J X^{(\ell)T} \dots X^{(1)T}$ or $X^{(1)} \dots X^{(\ell)} X^{(\ell+1)} X^{(\ell)T} \dots X^{(1)T}$ (with $X^{(\ell+1)}$ square and antisymmetric) for the antisymmetric case (and analogous matrices of differential operators). But all good things must come to an end, and for lack of time we have chosen not to pursue this direction.

2.3. Diagonal-parametrized Cayley identities

We would now like to formulate analogues of the ordinary and symmetric Cayley identities in which the diagonal elements of the matrix X are treated as parameters (i.e., not differentiated with respect to) and only the off-diagonal elements are treated as variables.

Theorem 2.10 (Diagonal-parametrized ordinary Cayley identity). *Let $X = (x_{ij})$ be an $n \times n$ matrix of indeterminates, let $\alpha = (\alpha_1, \dots, \alpha_n)$ and $\beta = (\beta_1, \dots, \beta_n)$ be arbitrary numbers (or symbols), and let $D_{\alpha, \beta, s}$ be the matrix of differential operators defined by*

$$(D_{\alpha, \beta, s})_{ij} = \begin{cases} s - \alpha_i \sum_{k \neq i} x_{ik} \frac{\partial}{\partial x_{ik}} - (1 - \alpha_i) \sum_{l \neq i} x_{li} \frac{\partial}{\partial x_{li}} & \text{if } i = j, \\ x_{ii}^{\beta_i} x_{jj}^{1-\beta_j} \frac{\partial}{\partial x_{ij}} & \text{if } i \neq j. \end{cases} \tag{2.25}$$

Then

$$\det(D_{\alpha, \beta, s})(\det X)^s = s(s+1) \dots (s+n-1) \left(\prod_{i=1}^n x_{ii} \right) (\det X)^{s-1}. \tag{2.26}$$

More generally, if $I, J \subseteq [n]$ with $|I| = |J| = k$, we have

$$\det((D_{\alpha, \beta, s})_{IJ})(\det X)^s = s(s+1) \dots (s+k-1) (\det X)^{s-1} \left(\prod_{i \in I} x_{ii}^{\beta_i} \right) \left(\prod_{j \in J} x_{jj}^{1-\beta_j} \right) \epsilon(I, J) (\det X_{J^c J^c}). \tag{2.27}$$

Please note that although the elements of the matrix $D_{\alpha, \beta, s}$ belong to a *noncommutative* algebra of differential operators, this matrix has the special property that each of its elements commutes with all the elements not in its own row or column; therefore, the determinant is well defined without any special ordering prescriptions.

Remarks. 1. Expressions involving $x_{ii}^{\beta_i}$ and $x_{jj}^{1-\beta_j}$ can be understood as follows: we work in the Weyl algebra generated by $\{x_{ij}\}_{i \neq j}$ and $\{\partial_{ij}\}_{i \neq j}$, augmented by s in the usual way (see Section 2.7) as well as by the central elements $y_i = x_{ii}^{\beta_i}$ and $z_i = x_{ii}^{1-\beta_i}$, it being understood that x_{ii} is a shorthand for $y_i z_i$.

2. It is easy to see why (2.26) holds with a right-hand side that is independent of the choice of β . Indeed, the operators $D_{\alpha,\beta,s}$ and $D_{\alpha,\beta',s}$ are related by the similarity transformation

$$D_{\alpha,\beta',s} = \text{diag}(x_{ii}^{\beta'_i - \beta_i}) D_{\alpha,\beta,s} \text{diag}(x_{ii}^{\beta_i - \beta'_i}) \tag{2.28}$$

where the quantities $x_{ii}^{\pm(\beta_i - \beta'_i)}$ commute with all entries in all matrices, so that $\det(D_{\alpha,\beta,s}) = \det(D_{\alpha,\beta',s})$. Similar reasoning explains why the right-hand side of (2.27) depends on β in the way it does.

3. It should be stressed that (2.26) provides a *non-minimal* Bernstein–Sato pair. In fact, a lower-order Bernstein–Sato pair can be obtained from (2.27) by taking $I = J = [n] \setminus \{i_0\}$ for any fixed $i_0 \in [n]$:

$$\det((D_{\alpha,\beta,s})_{\{i_0\}^c \{i_0\}^c}) (\det X)^s = s(s+1) \cdots (s+n-2) \left(\prod_{i=1}^n x_{ii} \right) (\det X)^{s-1}. \tag{2.29}$$

Theorem 2.11 (*Diagonal-parametrized symmetric Cayley identity*). Let $X^{\text{sym}} = (x_{ij})$ be an $n \times n$ symmetric matrix of indeterminates, let $\beta = (\beta_1, \dots, \beta_n)$ be arbitrary numbers (or symbols), and let $D_{\beta,s}^{\text{sym}}$ be the matrix of differential operators defined by

$$(D_{\beta,s}^{\text{sym}})_{ij} = \begin{cases} s - \frac{1}{2} \sum_{k>i} x_{ik} \frac{\partial}{\partial x_{ik}} - \frac{1}{2} \sum_{l<i} x_{li} \frac{\partial}{\partial x_{li}} & \text{if } i = j, \\ \frac{1}{2} x_{ii}^{\beta_i} x_{jj}^{1-\beta_j} \frac{\partial}{\partial x_{ij}} & \text{if } i < j, \\ \frac{1}{2} x_{ii}^{\beta_i} x_{jj}^{1-\beta_j} \frac{\partial}{\partial x_{ji}} & \text{if } i > j. \end{cases} \tag{2.30}$$

Then

$$\det(D_{\beta,s}^{\text{sym}}) (\det X^{\text{sym}})^s = s \left(s + \frac{1}{2} \right) \cdots \left(s + \frac{n-1}{2} \right) \left(\prod_{i=1}^n x_{ii} \right) (\det X^{\text{sym}})^{s-1}. \tag{2.31}$$

More generally, if $I, J \subseteq [n]$ with $|I| = |J| = k$, we have

$$\begin{aligned} & \det((D_{\beta,s}^{\text{sym}})_{IJ}) (\det X^{\text{sym}})^s \\ &= s \left(s + \frac{1}{2} \right) \cdots \left(s + \frac{k-1}{2} \right) (\det X^{\text{sym}})^{s-1} \left(\prod_{i \in I} x_{ii}^{\beta_i} \right) \left(\prod_{j \in J} x_{jj}^{1-\beta_j} \right) \epsilon(I, J) (\det X_{I^c J^c}^{\text{sym}}). \end{aligned} \tag{2.32}$$

Please note that in the symmetric case we are forced to take $\alpha_i = \frac{1}{2}$ for all i . Note also that (2.31) provides a non-minimal Bernstein–Sato pair, and that a lower-order pair can be obtained from (2.32) by taking $I = J = [n] \setminus \{i_0\}$ for any fixed $i_0 \in [n]$:

$$\det((D_{\beta,s}^{\text{sym}})_{\{i_0\}^c \{i_0\}^c}) (\det X^{\text{sym}})^s = s \left(s + \frac{1}{2} \right) \cdots \left(s + \frac{n-2}{2} \right) \left(\prod_{i=1}^n x_{ii} \right) (\det X^{\text{sym}})^{s-1}. \tag{2.33}$$

2.4. Laplacian-parametrized Cayley identities

In the preceding subsection we treated the off-diagonal elements $\{x_{ij}\}_{i \neq j}$ as indeterminates and the diagonal elements x_{ii} as parameters. Here we again treat the off-diagonal elements as indeterminates, but now we use the row sums $t_i = \sum_{j=1}^n x_{ij}$ as the parameters. (This way of writing a matrix arises in the matrix-tree theorem [2,27,69].) In other words, we define the row-Laplacian matrix with off-diagonal elements $\{x_{ij}\}_{i \neq j}$ and row sums 0,

$$(X^{\text{row-Lap}})_{ij} = \begin{cases} x_{ij} & \text{if } i \neq j, \\ -\sum_{k \neq i} x_{ik} & \text{if } i = j \end{cases} \tag{2.34}$$

and the diagonal matrix $T = \text{diag}(t_i)$, and we then study $\det(T + X^{\text{row-Lap}})$ where the t_i are treated as parameters. We shall need the matrix of differential operators

$$(\partial^{\text{row-Lap}})_{ij} = \begin{cases} \partial/\partial x_{ij} & \text{if } i \neq j, \\ 0 & \text{if } i = j. \end{cases} \tag{2.35}$$

Theorem 2.12 (Laplacian-parametrized ordinary Cayley identity). *Let $X^{\text{row-Lap}}$, T and $\partial^{\text{row-Lap}}$ be $n \times n$ matrices defined as above. Let U be the $n \times n$ matrix with all entries equal to 1. Then*

$$\begin{aligned} & [\det(U + \partial^{\text{row-Lap}}) - \det(\partial^{\text{row-Lap}})](\det(T + X^{\text{row-Lap}}))^s \\ &= \left(\sum_i t_i \right) s(s+1) \cdots (s+n-2) (\det(T + X^{\text{row-Lap}}))^{s-1}. \end{aligned} \tag{2.36}$$

Let us recall that, by the matrix-tree theorem [2,27,69], $\det(T + X^{\text{row-Lap}})$ is the generating polynomial of rooted directed spanning forests on the vertex set $[n]$, with weight $\prod_{i \in R} t_i \prod_{\vec{ij} \in E(F)} (-x_{ij})$ for a (rooted directed) forest with roots at the vertices $i \in R$ and edges $\vec{ij} \in E(F)$, directed towards the roots. In particular, the term linear in t_i – whose coefficient is the principal minor of order $n - 1$, $\det(X^{\text{row-Lap}})_{\{i\}^c \{i\}^c}$ – enumerates the directed spanning trees rooted at i . Taking this limit in (2.36), we obtain:

Corollary 2.13 (Cayley identity for the directed-spanning-tree polynomial). *For each $i \in [n]$, we have*

$$\begin{aligned} & [\det(U + \partial^{\text{row-Lap}}) - \det(\partial^{\text{row-Lap}})](\det(X^{\text{row-Lap}})_{\{i\}^c \{i\}^c})^s \\ &= s(s+1) \cdots (s+n-2) (\det(X^{\text{row-Lap}})_{\{i\}^c \{i\}^c})^{s-1}. \end{aligned} \tag{2.37}$$

We also have an analogous identity for symmetric Laplacian-parametrized matrices. What this means is that we introduce indeterminates $\{x_{ij}\}_{1 \leq i < j \leq n}$ and regard x_{ij} for $i > j$ as a synonym for x_{ji} ; we then define the symmetric Laplacian matrix

$$(X^{\text{sym-Lap}})_{ij} = \begin{cases} x_{ij} & \text{if } i < j, \\ x_{ji} & \text{if } i > j, \\ -\sum_{k \neq i} x_{ik} & \text{if } i = j \end{cases} \tag{2.38}$$

and the corresponding matrix of partial derivatives

$$(\partial^{\text{sym-Lap}})_{ij} = \begin{cases} \partial/\partial x_{ij} & \text{if } i < j, \\ \partial/\partial x_{ji} & \text{if } i > j, \\ 0 & \text{if } i = j. \end{cases} \tag{2.39}$$

We then have:

Theorem 2.14 (Laplacian-parametrized symmetric Cayley identity). Let $X^{\text{sym-Lap}}$, T and $\partial^{\text{sym-Lap}}$ be $n \times n$ matrices defined as above. Let U be the $n \times n$ matrix with all entries equal to 1. Then

$$\begin{aligned} & [\det(U + \partial^{\text{sym-Lap}}) - \det(\partial^{\text{sym-Lap}})](\det(T + X^{\text{sym-Lap}}))^s \\ &= \left(\sum_i t_i \right) 2s(2s + 1) \cdots (2s + n - 2) (\det(T + X^{\text{sym-Lap}}))^{s-1}. \end{aligned} \tag{2.40}$$

Similarly, $\det(T + X^{\text{sym-Lap}})$ is the generating polynomial of rooted (undirected) spanning forests on the vertex set $[n]$, with weight $\prod_{i \in R} t_i \prod_{ij \in E(F)} (-x_{ij})$ for a (rooted undirected) forest with roots at the vertices $i \in R$ and edges $ij \in E(F)$. In particular, the term linear in t_i – whose coefficient is $\det(X^{\text{sym-Lap}})_{\{i\}^c \{i\}^c}$ – is independent of i and enumerates the spanning trees. We thus obtain:

Corollary 2.15 (Cayley identity for the spanning-tree polynomial). For each $i \in [n]$, we have

$$\begin{aligned} & [\det(U + \partial^{\text{sym-Lap}}) - \det(\partial^{\text{sym-Lap}})](\det(X^{\text{sym-Lap}})_{\{i\}^c \{i\}^c})^s \\ &= 2s(2s + 1) \cdots (2s + n - 2) (\det(X^{\text{sym-Lap}})_{\{i\}^c \{i\}^c})^{s-1}. \end{aligned} \tag{2.41}$$

See [90] for many interesting additional properties of the spanning-tree polynomial.

2.5. Product-parametrized and border-parametrized rectangular Cayley identities

Here we will present two curious Cayley identities for rectangular matrices that are much simpler than the identities presented in Section 2.2, because the indeterminates occur linearly rather than quadratically in the argument of the determinant.

So let $X = (x_{ij})$ be an $m \times n$ matrix of indeterminates with $m \leq n$, and let $\partial = (\partial/\partial x_{ij})$ be the corresponding matrix of partial derivatives. One easy way to obtain square ($m \times m$) matrices from X and ∂ is to right-multiply them by $n \times m$ matrices A and B , respectively. We then have the following Cayley-type identity, in which A and B occur only as parameters:

Theorem 2.16 (Product-parametrized rectangular Cayley identity). Let $X = (x_{ij})$ be an $m \times n$ matrix of indeterminates with $m \leq n$, and let $\partial = (\partial/\partial x_{ij})$ be the corresponding matrix of partial derivatives. Let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times m$ matrices of constants. Then

$$\det(\partial B)(\det XA)^s = \det(A^T B) s(s + 1) \cdots (s + m - 1) (\det XA)^{s-1}. \tag{2.42}$$

More generally, if $I, J \subseteq [m]$ with $|I| = |J| = k$, then

$$\det((\partial B)_{IJ})(\det XA)^s = \det(MA) s(s + 1) \cdots (s + k - 1) (\det XA)^{s-1}. \tag{2.43}$$

where M is an $m \times n$ matrix defined as follows: if $I = \{i_1, \dots, i_k\}$ and $J = \{j_1, \dots, j_k\}$ in increasing order, then

$$M_{\alpha\beta} = \begin{cases} x_{\alpha\beta} & \text{if } \alpha \notin I, \\ b_{\beta j_h} & \text{if } \alpha = i_h. \end{cases} \tag{2.44}$$

Note that if $n = m$, then (2.42) reduces to the ordinary Cayley identity (2.1) multiplied on both sides by $(\det B)(\det A)^s$, while (2.43) specialized to $A = B = I_m$ gives the all-minors identity (2.2). In the general case $m \leq n$, (2.43) reduces to (2.42) when $I = J = [m]$ (since we then have $M = B^T$), while (2.43) reduces to the trivial identity $(\det XA)^s = (\det XA)^s$ when $I = J = \emptyset$ (since we then have $M = X$).

A second easy way to complete X and ∂ to square $(n \times n)$ matrices is to adjoin $n - m$ rows of constants at the bottom:

$$\hat{X} = \begin{pmatrix} X \\ A \end{pmatrix}, \quad \hat{\partial} = \begin{pmatrix} \partial \\ B \end{pmatrix} \quad (2.45)$$

where A and B are $(n - m) \times n$ matrices of constants. We can think of \hat{X} and $\hat{\partial}$ as “bordered” matrices obtained by filling out X and ∂ . We then have the following Cayley-type identity, in which A and B again occur only as parameters:

Theorem 2.17 (*Border-parametrized rectangular Cayley identity*). *Let $X = (x_{ij})$ be an $m \times n$ matrix of indeterminates with $m \leq n$, let $\partial = (\partial/\partial x_{ij})$ be the corresponding matrix of partial derivatives, and let A and B be $(n - m) \times n$ matrices of constants. Define \hat{X} and $\hat{\partial}$ as in (2.45). Then*

$$\det(\hat{\partial})(\det \hat{X})^s = \det(AB^T)s(s + 1) \cdots (s + m - 1)(\det \hat{X})^{s-1}. \quad (2.46)$$

When $m = n$ this formula reduces to the ordinary Cayley identity (2.1), if we make the convention that the determinant of an empty matrix is 1.

Please note that, by Laplace expansion, $\det \hat{X}$ is a linear combination of $m \times m$ minors of X (and likewise for $\det \hat{\partial}$). Indeed, when $n - m = 1$ one can obtain *all* such linear combinations in this way, by suitable choice of the row vector A ; for $n - m \geq 2$ one obtains in general a subset of such linear combinations. It is striking that the form of the identity – and in particular the polynomial $b(s)$ occurring in it – does not depend on the choice of the matrices A and B .

In Section 8 we will prove the “product-parametrized” and “border-parametrized” identities and then explain the close relationship between them.

2.6. Historical remarks

As noted in the Introduction, the identity (1.1) is conventionally attributed to Arthur Cayley (1821–1895); the generalization (1.2) to arbitrary minors is sometimes attributed to Alfredo Capelli (1855–1910). The trouble is, neither (1.1) nor (1.2) occurs anywhere – as far as we can tell – in the *Collected Papers of Cayley* [26]. Nor are we able to find these formulae in any of the relevant works of Capelli [15–19]. The operator $\Omega = \det(\partial)$ was indeed introduced by Cayley on the second page of his famous 1846 paper on invariants [25]; it became known as Cayley’s Ω -process and went on to play an important role in classical invariant theory (see e.g. [34,39,75,88,111]). But we strongly doubt that Cayley ever knew (1.1).

A detailed history of (1.1) and (1.2) will be presented elsewhere [3]. Suffice it to say that the special case for 2×2 matrices appears already in the 1872 book of Alfred Clebsch (1833–1872) on the invariant theory of binary forms [29, p. 20]. But even for $n = 3$, the first unambiguous statement of which we are aware appears in an 1890 paper of Giulio Vivanti (1859–1949) [106].¹⁶ And amazingly, in this

¹⁶ This paper is also cited in Muir’s massive annotated bibliography of work on the theory of determinants [70, vol. 4, p. 479]. Indeed, it was thanks to Muir that we discovered Vivanti’s paper. Malek Abdesselam has also drawn our attention to the papers of Clebsch (1861) [28, pp. 7–14] and Gordan (1872) [45, pp. 107–116], where formulae closely related to (1.1) can be found.

very first paper, Vivanti proves not only the basic “Cayley” identity (1.1) but also the generalization (1.2) for minors, for completely general n and k albeit only in the case $I = J$ (i.e., principal minors). In fact, his inductive method of proof works only because he is handling the “all-principal-minors” version; it would not work for the “simple” identity (1.1) alone.

Proofs (by direct computation) of the “Cayley” identity (1.1) for $n = 3$ can be found in the early-twentieth-century books of Grace and Young (1903) [46], Weitzenböck (1923) [110] and Turnbull (1928) [101]. Weitzenböck also states (1.1) without proof for general n , saying that it is obtained “by completely analogous calculation” [110, p. 16]; similarly, Turnbull states both (1.1) and (1.2) for general n and leaves them as exercises for the reader [101, pp. 114–116]. (Unfortunately, Turnbull’s old-fashioned notation is very difficult to follow.) We are not convinced that the extension from $n = 3$ to general n is quite so trivial as these authors imply. We will, in any case, provide an elementary algebraic/combinatorial proof of (1.1)/(1.2) in Section 3.1 below.

The symmetric analogues (2.4)/(2.5) are due to Gårding in 1948 [41]; see also [36, Proposition VII.1.4]. The antisymmetric analogue (2.7) is due to Shimura [89] in 1984; see again [36, Proposition VII.1.4], where the quaternionic hermitian determinant (Moore determinant [4]) is equivalent to a pfaffian, and see also [61, Corollary 3.13].

Shimura [89] also gives generalizations of all these formulae in which $\det(\partial)$ is replaced by other homogeneous differential operators. Similarly, Rubenthaler and Schiffmann [82, especially Section 5] and Faraut and Korányi [36, Proposition VII.1.6] give generalizations in which both $\det(\partial)$ and $\det X$ are replaced by suitable products of leading principal minors. All these proofs are analytic in nature. Elegant algebraic/combinatorial proofs of identities in which $\det X$ is replaced by a product of minors have been given by Canfield, Williamson and Evans [14, Theorem 1.3 and Lemma 2.12], [114, Theorem 4.1].¹⁷

The one-matrix rectangular symmetric Cayley identity (2.18) has been proven recently by Faraut and Korányi [36, Section XVI.4], Khékalo [57,58] and Rubin [83], using analytic methods.

Many of the polynomials P treated here can also be understood as relative invariants of prehomogeneous vector spaces [52,60]; and the corresponding b -functions have been computed in that context, mostly by means of microlocal calculus [59,85,98]. See [60, Appendix] for a table of all irreducible reduced prehomogeneous vector spaces and their relative invariants and b -functions. All but one of the “generic” cases in this table (i.e., those involving matrices with one or more arbitrary dimensions) correspond to identities treated here: namely, the ordinary, symmetric and antisymmetric Cayley identities [(2.1), (2.4) and (2.7)] correspond to cases (1), (2) and (3), respectively, while the one-matrix rectangular symmetric and antisymmetric Cayley identities [(2.18) and (2.21)] correspond to cases (15) and (13). The b -functions listed in that table of course agree with ours.¹⁸ (The final “generic” case (30) involves matrices X of order $3 \times 2n$ with $P(X) = \text{tr}(XJX^T)^2$, which falls outside our methods since it is neither a determinant nor a pfaffian.) Finally, the most difficult result obtained in the present paper – namely, the multi-matrix rectangular Cayley identity (2.23) – has very recently been proven independently by Sugiyama [98, Theorem 0.1] in the context of prehomogeneous vector spaces associated to equioriented quivers of type A; his proof uses a decomposition formula found earlier by himself and F. Sato [86]. Indeed, Sugiyama proved an even more general result [98, Theorem 3.4], applying to quivers of type A with arbitrary orientation.¹⁹

It is worth stressing that the Cayley identity (1.1) – though not, as far as we can tell, the all-minors version (1.2) – is an immediate consequence of a deeper identity due to Capelli [16–18], in which the operator $H = (\det X)(\det \partial)$ is represented as a noncommutative determinant involving the $\mathfrak{gl}(n)$ generators $X^T \partial$: see e.g. [105, p. 53], [51, pp. 569–570] or [22, Appendix] for the easy deduction of Cayley from Capelli. Likewise, the symmetric Cayley identity (2.4) follows from a symmetric Capelli-type identity due to Turnbull [102] (see also [108]), and the antisymmetric Cayley identity

¹⁷ See also Turnbull [103] for a similar result, but expressed in difficult-to-follow notation.

¹⁸ After making the translation of conventions $s \rightarrow s + 1$: see footnote 7 above. See also F. Sato and Sugiyama [86, Section 3.1 and Lemma 4.2] for an alternate approach to (2.1), (2.7) and (2.18).

¹⁹ We are grateful to Nero Budur for explaining to us the connection between our results and the theory of prehomogeneous vector spaces, and for drawing our attention to the work of Sugiyama [98].

(2.7) follows from an antisymmetric Capelli-type identity due independently to Kostant and Sahi [63] and to Howe and Umeda [51] (see also [61]).

The one-matrix rectangular symmetric Cayley identity (2.18) for $m = 1$ follows from a Capelli-type identity given in [111, pp. 291–293] and [105, p. 61]. For $m = n$ it of course follows from the ordinary Capelli identity. For $2 \leq m \leq n - 1$ we do not know any Capelli-type identity.

Proofs of the Capelli-type identities based on group-representation theory have been given by Howe and Umeda [50,51]. Combinatorial proofs of the Capelli and Turnbull identities have been given by Foata and Zeilberger [37]. We have recently given very simple algebraic proofs of these same identities as well as some generalizations [22]. See also Weyl [111, pp. 39–42] and Fulton–Harris [39, Appendix F.3] for more traditional proofs. Further information on Capelli-type identities can be found in [51,105].

2.7. Some algebraic preliminaries

A few words are needed about how the identities (2.1) ff. – or more generally, Bernstein-type identities of the form

$$Q(s, x, \partial/\partial x)P(x)^s = b(s)P(x)^{s-1} \tag{2.47}$$

where $x = (x_1, \dots, x_n)$ – are to be interpreted. On the one hand, they can be interpreted as analytic identities for functions of real or complex variables x_1, \dots, x_n , where s is a real or complex number; here $P(x)^s$ denotes any fixed branch on any open subset of \mathbb{R}^n or \mathbb{C}^n where it is well defined. Alternatively, these formulae can be regarded as purely algebraic identities, in several different ways:

- (1) For integer $s \geq 1$, as an identity in the ring $R[x_1, \dots, x_n]$ of polynomials in the indeterminates x_1, \dots, x_n with coefficients in some commutative ring R (for instance, R could be $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ or \mathbb{C}).²⁰
- (2) For any integer s (positive or negative), as an identity in the field $K(x_1, \dots, x_n)$ of rational fractions in the indeterminates x_1, \dots, x_n with coefficients in some field K (for instance, K could be \mathbb{Q}, \mathbb{R} or \mathbb{C}).²¹
- (3) For s interpreted symbolically, as an identity in a module defined as follows [32, pp. 93–94], [64, p. 96 ff.]: Let K be a field of characteristic 0, let x_1, \dots, x_n and s be indeterminates, and let $A_n(K)[s]$ be the K -algebra generated by $x_1, \dots, x_n, \partial/\partial x_1, \dots, \partial/\partial x_n$ and s with the usual commutation relations. (That is, it is the algebra of differential operators with respect to x_1, \dots, x_n in which the coefficients are polynomials in x_1, \dots, x_n, s with coefficients in K .) Now fix a nonzero polynomial $P \in K[x_1, \dots, x_n]$, and let $K[x, s, P^{-1}]$ denote the ring of rational fractions in the indeterminates x_1, \dots, x_n, s whose denominators are powers of P . (It is a subring of the field $K(x_1, \dots, x_n, s)$ of all rational fractions in x_1, \dots, x_n, s .) Let $K[x, s, P^{-1}]P^s$ be the free $K[x, s, P^{-1}]$ -module consisting of objects of the form fP^s where $f \in K[x, s, P^{-1}]$; here P^s is treated as a formal symbol. We can define formal differentiation by

$$\frac{\partial}{\partial x_i}(fP^s) = \left(\frac{\partial f}{\partial x_i} + sf \frac{\partial P}{\partial x_i} P^{-1} \right) P^s \tag{2.48}$$

²⁰ When $b(s)$ contains fractions [e.g. (2.4)/(2.5) and (2.31)/(2.32)] we should assume that the coefficient ring R contains those fractions [i.e. in this case $\frac{1}{2}$].

²¹ When $b(s)$ contains fractions [e.g. (2.4)/(2.5) and (2.31)/(2.32)] we should assume that the coefficient field K contains those fractions [i.e. in this case $\frac{1}{2}$]. Usually we will take K to be a field of characteristic 0, so that K contains the rationals \mathbb{Q} as a subfield.

where $\partial f/\partial x_i$ is the standard formal derivative of a rational fraction. This differentiation is easily extended to an action of $A_n(K)[s]$ on $K[x, s, P^{-1}]P^s$, making the latter into a left $A_n(K)[s]$ -module.²²

Let us now show that all these interpretations are equivalent.

We begin by recalling some elementary facts. Let $p(x_1, \dots, x_n)$ be a polynomial with coefficients in some commutative ring R , and let d_i be the degree of p with respect to the variable x_i . Suppose that there exist sets $X_1, \dots, X_n \subseteq R$ with $|X_i| > d_i$ for all i , such that $p(x_1, \dots, x_n) = 0$ whenever $(x_1, \dots, x_n) \in X_1 \times \dots \times X_n$. Then p must be the zero polynomial, i.e. all its coefficients are zero. Note in particular that if the sets X_1, \dots, X_n are infinite, then this reasoning applies to polynomials of arbitrary degree. As a special case of this, if $R = \mathbb{R}$ or \mathbb{C} and $p(x_1, \dots, x_n) = 0$ for all $x = (x_1, \dots, x_n)$ lying in some nonempty open set U of \mathbb{R}^n or \mathbb{C}^n , then p must be the zero polynomial.

Now let K be a field of characteristic 0, and let $P(x)$, $Q(s, x, \partial/\partial x)$ and $b(s)$ be polynomials with coefficients in K [as always we use the shorthand $x = (x_1, \dots, x_n)$]. Then elementary algebraic manipulations allow us to write

$$Q(s, x, \partial/\partial x)P(x)^s - b(s)P(x)^{s-1} = R(s, x)P(x)^{s-m} \tag{2.49}$$

for some polynomial $R(s, x)$ and some integer $m \geq 0$. Combining this fact with the preceding observations, we obtain immediately the following two propositions:

Proposition 2.18 (Equivalence theorem for symbolic s). *Let K be a field of characteristic 0, and let P, Q, b and R be as before. Then the following are equivalent:*

- (a) (2.47) holds as an algebraic identity in the module $K[x, s, P^{-1}]P^s$, symbolically in s .
- (b) $R(s, x) = 0$ in the polynomial ring $K[s, x]$.

Furthermore, if K is infinite, then (a)–(b) are equivalent to:

- (c) For infinitely many $s \in K$, there exist infinite sets $X_1, \dots, X_n \subseteq K$ (possibly depending on s), such that $R(s, x_1, \dots, x_n) = 0$ for $(x_1, \dots, x_n) \in X_1 \times \dots \times X_n$ and the given value of s .

Finally, if $K = \mathbb{R}$ or \mathbb{C} , then (a)–(c) are equivalent to:

- (d) For infinitely many $s \in K$, there exist a nonempty open set $U \subseteq K^n$ (possibly depending on s) and a branch of $P(x)^s$ defined on U such that (2.47) holds for all $x \in U$.
- (e) For every $s \in K$, every nonempty open set $U \subseteq K^n$ and every branch of $P(x)^s$ defined on U , (2.47) holds for all $x \in U$.

²² More generally, we can proceed as follows: Let R be an integral domain and let A be an abelian group. We then define R^A to be the commutative ring with identity generated by the symbols x^a ($x \in R, a \in A$) subject to the relations $x^a x^b = x^{a+b}$, $x^a y^a = (xy)^a$ and $x^0 = 1$. In particular, if A contains the integers as a subgroup, then we can consider R as a subring of R^A by identifying $x \in R$ with $x^1 \in R^A$.

Now suppose that R is a polynomial ring $S[x_1, \dots, x_n]$ where S is an integral domain of characteristic 0, and that A is a subgroup of the additive group of S (where $1 \in \mathbb{Z} \subseteq A$ is identified with $1 \in S$). Then we can define an action of the differential operators $\partial/\partial x_i$ on R^A by

$$\frac{\partial}{\partial x_i}(P^a) = a \frac{\partial P}{\partial x_i} P^{a-1}$$

together with the usual product rule. This makes R^A into a left $A_n(S)$ -module [where $A_n(S)$ is the Weyl algebra in n variables over S].

In order to handle Bernstein-type identities, we will introduce an indeterminate s and take $A = \mathbb{Z} + s\mathbb{Z}$ and $S = K[s]$ for some field K . Then we will work within the submodule of R^A consisting of elements of the form fP^{s+a} for $f \in K[x_1, \dots, x_n, s]$, $a \in \mathbb{Z}$ and some fixed $P \in K[x_1, \dots, x_n]$. This submodule is isomorphic to $K[x, s, P^{-1}]P^s$.

Proposition 2.19 (Equivalence theorem for fixed s). Let K be a field of characteristic 0, let P, Q, b and R be as before, and fix some element $s \in K$. Then the following are equivalent:

(b) $R(s, x) = 0$ in the polynomial ring $K[x]$ (for the given value of s).

Furthermore, if K is infinite, then (a)–(b) are equivalent to:

(c) There exist infinite sets $X_1, \dots, X_n \subseteq K$ such that $R(s, x_1, \dots, x_n) = 0$ for $(x_1, \dots, x_n) \in X_1 \times \dots \times X_n$ (for the given value of s).

Finally, if $K = \mathbb{R}$ or \mathbb{C} , then (a)–(c) are equivalent to:

(d) There exist a nonempty open set $U \subseteq K^n$ and a branch of $P(x)^s$ defined on U such that (2.47) holds for all $x \in U$ (for the given value of s).

(e) For every nonempty open set $U \subseteq K^n$ and every branch of $P(x)^s$ defined on U , (2.47) holds for all $x \in U$ (for the given value of s).

In particular, Proposition 2.18 shows that it suffices to prove (2.47) for infinitely many positive or negative integers s , using the elementary interpretation (1) or (2) above; it then holds automatically for arbitrary s as an identity in the module $K[x, s, P^{-1}]P^s$ and as an analytic identity. Likewise, it suffices to specialize x_1, \dots, x_n to real or complex variables and to prove (2.47) for some nonempty open set in \mathbb{R}^n or \mathbb{C}^n . In what follows, we shall repeatedly take advantage of these simplifications. (Many previous authors – especially the earlier ones – have done so as well, but without making Proposition 2.18 explicit.)

3. Elementary proofs of Cayley-type identities

In this section we give proofs of the three main Cayley-type identities for square matrices (Theorems 2.1–2.3) that use nothing but elementary properties of determinants (notably, Jacobi’s identity for cofactors) along with the elementary formulae for the derivative of a product or a power.

The general situation we have to handle in all three cases is as follows: Let Γ be a finite index set, let $(E_\gamma)_{\gamma \in \Gamma}$ be given $n \times n$ matrices with elements in some field K , and let $(x_\gamma)_{\gamma \in \Gamma}$ be indeterminates. Now define the matrix $X = \sum_{\gamma \in \Gamma} x_\gamma E_\gamma$. If $A = (A_\gamma)_{\gamma \in \Gamma}$ is a K -valued vector, we write $\partial_A = \sum_{\gamma \in \Gamma} A_\gamma \partial / \partial x_\gamma$ and $E_A = \sum_{\gamma \in \Gamma} A_\gamma E_\gamma$, so that $\partial_A X = E_A$. We need a formula for successive derivatives of $(\det X)^s$:

Lemma 3.1. Let $(E_\gamma)_{\gamma \in \Gamma}$ be $n \times n$ matrices, let $(x_\gamma)_{\gamma \in \Gamma}$ be indeterminates, and define $X = \sum_{\gamma \in \Gamma} x_\gamma E_\gamma$. For any sequence A_1, \dots, A_k of K -valued vectors, we have

$$\begin{aligned} & \left(\prod_{i=1}^k \partial_{A_i} \right) (\det X)^s \\ &= (-1)^k (\det X)^s \sum_{\tau \in \mathcal{S}_k} (-s)^{\#(\text{cycles of } \tau)} \prod_{\substack{\text{cycles of } \tau \\ C=(\alpha_1, \dots, \alpha_\ell)}} \text{tr}(X^{-1} E_{A_{\alpha_1}} \dots X^{-1} E_{A_{\alpha_\ell}}) \end{aligned} \tag{3.1a}$$

$$= (\det X)^s \sum_{\tau \in \mathcal{S}_k} \text{sgn}(\tau) s^{\#(\text{cycles of } \tau)} \prod_{\substack{\text{cycles of } \tau \\ C=(\alpha_1, \dots, \alpha_\ell)}} \text{tr}(X^{-1} E_{A_{\alpha_1}} \dots X^{-1} E_{A_{\alpha_\ell}}). \tag{3.1b}$$

Proof. By induction on k . The case $k = 1$ follows from Cramer’s rule or alternatively from the relation

$$(\det X)^s = \exp(s \text{tr} \log X). \tag{3.2}$$

For the inductive step we shall need the identity

$$\partial_A(X^{-1}) = -X^{-1}(\partial_A X)X^{-1}, \tag{3.3}$$

which follows from $\partial_A(X^{-1}X) = 0$. Assume now that the theorem is valid for k and let us apply $\partial_{A_{k+1}}$. When this derivative hits $(\det X)^s$, it creates a new cycle $(k + 1)$, with prefactor $s = (-1)(-s)$; these terms correspond to permutations $\tau \in \mathcal{S}_{k+1}$ in which the element $k + 1$ is fixed. Alternatively, the derivative can hit one of the X^{-1} factors in one of the traces; by (3.3) this inserts the element $k + 1$ into one of the existing cycles at an arbitrary position, and produces an extra factor -1 ; these terms correspond to permutations $\tau \in \mathcal{S}_{k+1}$ in which the element $k + 1$ is not fixed. \square

For $1 \leq i, j \leq n$, let E^{ij} be the matrix with a 1 in position ij and zeros elsewhere, i.e.

$$(E^{ij})_{i'j'} = \delta_{i,i'}\delta_{j,j'}. \tag{3.4}$$

We will express the matrices E_γ in each of our three cases in terms of the E^{ij} .

3.1. Ordinary Cayley identity

Proof of Theorem 2.1. In this case the index set Γ is simply $[n] \times [n]$, and we write $X = \sum_{i,j=1}^n x_{ij} E^{ij}$. Now let $I = \{i_1, \dots, i_k\}$ with $i_1 < \dots < i_k$ and $J = \{j_1, \dots, j_k\}$ with $j_1 < \dots < j_k$, so that

$$\det(\partial_{IJ}) = \sum_{\sigma \in \mathcal{S}_k} \text{sgn}(\sigma) \prod_{r=1}^k \frac{\partial}{\partial x_{i_r j_{\sigma(r)}}}. \tag{3.5}$$

For each fixed $\sigma \in \mathcal{S}_k$, we apply Lemma 3.1 with $\partial_{A_r} = \partial / \partial x_{i_r j_{\sigma(r)}}$. In the traces we have $E_{A_{\alpha_r}} = E^{i_{\alpha_r} j_{\sigma(\alpha_r)}}$ and hence

$$\text{tr}(X^{-1} E^{i_{\alpha_1} j_{\sigma(\alpha_1)}} \dots X^{-1} E^{i_{\alpha_\ell} j_{\sigma(\alpha_\ell)}}) = X_{j_{\sigma(\alpha_\ell)} i_{\alpha_1}}^{-1} X_{j_{\sigma(\alpha_1)} i_{\alpha_2}}^{-1} \dots X_{j_{\sigma(\alpha_{\ell-1})} i_{\alpha_\ell}}^{-1} \tag{3.6a}$$

$$= \prod_{r=1}^{\ell} X_{i_{\tau(\alpha_r)} j_{\sigma(\alpha_r)}}^{-T}, \tag{3.6b}$$

where $X^{-T} \equiv (X^{-1})^T$ and we have used the fact that, for τ as in Lemma 3.1, $\tau(\alpha_i) = \alpha_{i+1}$ for $i = 1, \dots, \ell - 1$ and $\tau(\alpha_\ell) = \alpha_1$. We can now combine all the different traces into a single product. We obtain

$$\det(\partial_{IJ})(\det X)^s = (\det X)^s \sum_{\sigma \in \mathcal{S}_k} \text{sgn}(\sigma) \sum_{\tau \in \mathcal{S}_k} \text{sgn}(\tau) s^{\#(\text{cycles of } \tau)} \prod_{r=1}^k (X^{-T})_{i_{\tau(r)} j_{\sigma(r)}}. \tag{3.7}$$

Let us now define the permutation $\pi = \sigma \circ \tau^{-1}$ and change variables from (σ, τ) to (τ, π) , using $\text{sgn}(\sigma) \text{sgn}(\tau) = \text{sgn}(\pi)$. The product over r can be written equivalently as a product over $t = \tau(r)$. We have

$$\sum_{\pi \in \mathcal{S}_k} \text{sgn}(\pi) \prod_{t=1}^k (X^{-T})_{i_t j_{\pi(t)}} = \det((X^{-T})_{IJ}) = (\det X)^{-1} \epsilon(I, J) (\det X_{J^c J^c}) \tag{3.8}$$

by Jacobi’s identity [Lemma A.1(e)], while

$$\sum_{\tau \in \mathcal{S}_k} s^{\#(\text{cycles of } \tau)} = s(s + 1) \cdots (s + k - 1) \tag{3.9}$$

(see e.g. [47, p. 263, Eq. (6.11)] or [92, Proposition 1.3.4] for this well-known equality). \square

3.2. Symmetric Cayley identity

Proof of Theorem 2.2. In this case the index set Γ consists of ordered pairs $(i, j) \in [n] \times [n]$ with $i \leq j$, and we write $X^{\text{sym}} = \sum_{i < j} x_{ij}(E^{ij} + E^{ji}) + \sum_i x_{ii}E^{ii}$. Then

$$\partial_{ij}^{\text{sym}} X^{\text{sym}} = \frac{1}{2}(E^{ij} + E^{ji}) \tag{3.10}$$

in all three cases ($i < j$, $i > j$ and $i = j$). Now let us apply $\det(\partial_{ij}^{\text{sym}})$ to $(\det X^{\text{sym}})^s$ and compare to what we had in the previous proof. On the one hand we have a factor 2^{-k} coming from the k derivatives. On the other hand, each E_A is now a sum of two terms E^{ij} and E^{ji} ; in each cycle $(\alpha_1, \dots, \alpha_l)$ of τ , the argument of the trace becomes

$$X^{-1}(E^{i\alpha_1 j\sigma(\alpha_1)} + E^{j\sigma(\alpha_1) i\alpha_1})X^{-1}(E^{i\alpha_2 j\sigma(\alpha_2)} + E^{j\sigma(\alpha_2) i\alpha_2}) \cdots \tag{3.11}$$

(to lighten the notation we have written X instead of X^{sym}). We therefore need to sum over all the 2^l ways of choosing E^{ij} or E^{ji} in each factor within the given trace (hence 2^k choices overall). Performing the trace, we will obtain terms of the form

$$X_{j\sigma(\alpha_r) i\alpha_{r+1}}^{-1}, X_{i\alpha_r i\alpha_{r+1}}^{-1}, X_{j\sigma(\alpha_r) j\sigma(\alpha_{r+1})}^{-1}, X_{i\alpha_r j\sigma(\alpha_{r+1})}^{-1}. \tag{3.12}$$

Let us now fix one of the 2^k choices and sum over the permutation σ . If one or more of the factors is of the form $X_{j\sigma(\alpha_r) j\sigma(\alpha_{r+1})}^{-1}$, then the sum over σ will vanish because the exchange between $\sigma(\alpha_r)$ and $\sigma(\alpha_{r+1})$ takes a -1 from $\text{sgn}(\sigma)$. Therefore, in each cycle of τ there are only two nonvanishing contributions, corresponding to the two ways of coherently orienting the cycle. One of these has X^{-T} (as in the previous proof) and the other has X^{-1} , but these are in fact equal since X is symmetric. We thus have, compared to the previous proof, an extra factor $2^{\#(\text{cycles of } \tau)}$. We now change variables, as before, in the sum over permutations. The sum over π gives (3.8) exactly as before, while the sum over τ now gives

$$2^{-k} \sum_{\tau \in \mathcal{S}_k} (2s)^{\#(\text{cycles of } \tau)} = s \left(s + \frac{1}{2}\right) \cdots \left(s + \frac{k-1}{2}\right) \tag{3.13}$$

by (3.9). \square

3.3. Antisymmetric Cayley identity

Let us recall the definition of the pfaffian of a $2n \times 2n$ antisymmetric matrix:

$$\text{pf } A = \frac{1}{2^n n!} \sum_{\sigma \in \mathcal{S}_{2n}} \text{sgn}(\sigma) a_{\sigma(1)\sigma(2)} \cdots a_{\sigma(2n-1)\sigma(2n)}. \tag{3.14}$$

Proof of Theorem 2.3. In this case the index set Γ consists of ordered pairs $(i, j) \in [n] \times [n]$ with $i < j$, and we write $X^{\text{antisym}} = \sum_{i < j} x_{ij}(E^{ij} - E^{ji})$. Then

$$\partial_{ij}^{\text{antisym}} X^{\text{antisym}} = E^{ij} - E^{ji}. \tag{3.15}$$

Now let $I = \{i_1, \dots, i_{2k}\}$ with $i_1 < \dots < i_{2k}$, and let us apply $\text{pf}(\partial_{II}^{\text{antisym}})$ to $(\text{pf} X^{\text{antisym}})^s = (\det X^{\text{antisym}})^{s/2}$; using the representation (3.14) for $\text{pf}(\partial_{II}^{\text{antisym}})$, we obtain

$$\text{pf}(\partial_{II}^{\text{antisym}})(\text{pf} X^{\text{antisym}})^s = \frac{1}{2^k k!} \sum_{\sigma \in S_{2k}} \text{sgn}(\sigma) \left(\prod_{r=1}^k \frac{\partial}{\partial x_{i_{\sigma(2r-1)} i_{\sigma(2r)}}} \right) (\det X^{\text{antisym}})^{s/2}. \tag{3.16}$$

Now apply Lemma 3.1 as before. In each cycle $(\alpha_1, \dots, \alpha_\ell)$ of τ , the argument of the trace becomes

$$X^{-1}(E^{i_{\sigma(2\alpha_1-1)} i_{\sigma(2\alpha_1)}} - E^{i_{\sigma(2\alpha_1)} i_{\sigma(2\alpha_1-1)}}) X^{-1}(E^{i_{\sigma(2\alpha_2-1)} i_{\sigma(2\alpha_2)}} - E^{i_{\sigma(2\alpha_2)} i_{\sigma(2\alpha_2-1)}}) \dots \tag{3.17}$$

Once again we have 2^k choices in the E factors; but here these choices correspond simply to pre-multiplying σ by one of the 2^k permutations that leave fixed the pairs $\{1, 2\}, \dots, \{2k-1, 2k\}$; therefore, after summing over σ we simply get a factor 2^k . Let us now introduce the permutation σ^τ defined by

$$\sigma^\tau(2r-1) = \sigma(2\tau(r)-1), \tag{3.18a}$$

$$\sigma^\tau(2r) = \sigma(2r). \tag{3.18b}$$

For each $\tau \in S_k$, the map $\sigma \mapsto \sigma^\tau$ is an automorphism of S_{2k} and satisfies $\text{sgn}(\sigma^\tau) = \text{sgn}(\sigma) \text{sgn}(\tau)$. We have

$$\begin{aligned} & \text{pf}(\partial_{II}^{\text{antisym}})(\text{pf} X^{\text{antisym}})^s \\ &= 2^k \frac{1}{2^k k!} \sum_{\sigma \in S_{2k}} \text{sgn}(\sigma) \sum_{\tau \in S_k} \text{sgn}(\tau) (s/2)^{\#(\text{cycles of } \tau)} \prod_{r=1}^k (X^{-T})_{i_{\sigma(2\tau(r)-1)} i_{\sigma(2r)}}. \end{aligned} \tag{3.19}$$

We now define $\pi = \sigma^\tau$ and change variables from (σ, τ) to (τ, π) as before. The sum over π gives

$$\frac{1}{2^k k!} \sum_{\pi \in S_{2k}} \text{sgn}(\pi) \prod_{r=1}^k (X^{-T})_{i_{\pi(2r-1)} i_{\pi(2r)}} = \text{pf}((X^{-T})_{II}) = (\text{pf} X)^{-1} \in (I)(\text{pf} X_{I^c I^c}) \tag{3.20}$$

by the pfaffian version of Jacobi's identity [cf. (A.21)], while the sum over τ gives

$$2^k \sum_{\tau \in S_k} (s/2)^{\#(\text{cycles of } \tau)} = s(s+2) \dots (s+2k-2) \tag{3.21}$$

by (3.9). \square

4. Proofs of Cayley-type identities by Grassmann/Gaussian representation of $(\det X)^s$

Let us now give simple proofs of Theorems 2.1–2.3, based on representing $(\det X)^s$ as a fermionic or bosonic Gaussian integral. A brief introduction to fermionic and bosonic Gaussian integration can be found in Appendix A.

4.1. Ordinary Cayley identity

Proof of Theorem 2.1. Assume that s is a positive integer, and let us introduce Grassmann variables $\psi_i^{(\alpha)}, \bar{\psi}_i^{(\alpha)}$ for $i = 1, \dots, n$ and $\alpha = 1, \dots, s$. We can then write

$$(\det X)^s = \int \mathcal{D}(\psi, \bar{\psi}) e^{\bar{\psi} X \psi}, \tag{4.1}$$

where we have used the shorthand

$$\bar{\psi} X \psi \equiv \sum_{\alpha=1}^s \sum_{i,j=1}^n \bar{\psi}_i^{(\alpha)} X_{ij} \psi_j^{(\alpha)}. \tag{4.2}$$

Now let $I = \{i_1, \dots, i_k\}$ with $i_1 < \dots < i_k$ and $J = \{j_1, \dots, j_k\}$ with $j_1 < \dots < j_k$, so that

$$\det(\partial_{IJ}) = \sum_{\sigma \in S_k} \text{sgn}(\sigma) \prod_{r=1}^k \frac{\partial}{\partial X_{i_r j_{\sigma(r)}}}. \tag{4.3}$$

Applying this to (4.1), we obtain

$$\det(\partial_{IJ})(\det X)^s = \int \mathcal{D}(\psi, \bar{\psi}) \sum_{\sigma \in S_k} \text{sgn}(\sigma) \sum_{\alpha_1, \dots, \alpha_k=1}^s \left(\prod_{r=1}^k \bar{\psi}_{i_r}^{(\alpha_r)} \psi_{j_{\sigma(r)}}^{(\alpha_r)} \right) e^{\bar{\psi} X \psi}. \tag{4.4}$$

When X is an invertible real or complex matrix, Wick’s theorem for “complex” fermions (Theorem A.16) gives

$$\int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{r=1}^k \bar{\psi}_{i_r}^{(\alpha_r)} \psi_{j_{\sigma(r)}}^{(\alpha_r)} \right) e^{\bar{\psi} X \psi} = (\det X)^s \sum_{\tau \in S_k} \text{sgn}(\tau) \prod_{r=1}^k (X^{-T})_{i_r j_{\sigma(\tau(r))}} \delta_{\alpha_r \alpha_{\tau(r)}}. \tag{4.5}$$

Let us now define the permutation $\pi = \sigma \circ \tau$ and change variables from (σ, τ) to (τ, π) , using $\text{sgn}(\sigma) \text{sgn}(\tau) = \text{sgn}(\pi)$. We then have

$$\sum_{\pi \in S_k} \text{sgn}(\pi) \prod_{r=1}^k (X^{-T})_{i_r j_{\pi(r)}} = \det((X^{-T})_{IJ}) = (\det X)^{-1} \epsilon(I, J) (\det X_{J^c I^c}) \tag{4.6}$$

by Jacobi’s identity, while

$$\sum_{\tau \in S_k} \sum_{\alpha_1, \dots, \alpha_k=1}^s \prod_{r=1}^k \delta_{\alpha_r \alpha_{\tau(r)}} = \sum_{\tau \in S_k} s^{\#(\text{cycles of } \tau)} = s(s+1) \dots (s+k-1) \tag{4.7}$$

by (3.9). This proves (2.2) for integer $s \geq 1$, when X is an invertible real or complex matrix. By Proposition 2.18, this is sufficient to prove the identity. \square

Alternate proof of Theorem 2.1. Instead of using “complex” fermions, we can use complex bosons. So assume that $s = -m$ where m is a positive integer, and that X is a complex matrix whose hermitian part is positive-definite; and let us introduce bosonic variables $\varphi_i^{(\alpha)}, \bar{\varphi}_i^{(\alpha)}$ for $i = 1, \dots, n$ and $\alpha = 1, \dots, m$. We then have

$$(\det X)^{-m} = \int \mathcal{D}(\varphi, \bar{\varphi}) e^{-\bar{\varphi} X \varphi}. \tag{4.8}$$

By the same method as before, we obtain

$$\det(\partial_{IJ})(\det X)^{-m} = (-1)^k \int \mathcal{D}(\varphi, \bar{\varphi}) \sum_{\sigma \in \mathcal{S}_k} \text{sgn}(\sigma) \sum_{\alpha_1, \dots, \alpha_k=1}^m \left(\prod_{r=1}^k \bar{\varphi}_{i_r}^{(\alpha_r)} \varphi_{j_{\sigma(r)}}^{(\alpha_r)} \right) e^{-\bar{\varphi} X \varphi}. \tag{4.9}$$

Wick’s theorem for complex bosons (Theorem A.4) now gives

$$\int \mathcal{D}(\varphi, \bar{\varphi}) \left(\prod_{r=1}^k \bar{\varphi}_{i_r}^{(\alpha_r)} \varphi_{j_{\sigma(r)}}^{(\alpha_r)} \right) e^{-\bar{\varphi} X \varphi} = (\det X)^{-m} \sum_{\tau \in \mathcal{S}_k} \prod_{r=1}^k (X^{-T})_{i_r j_{\sigma(\tau(r))}} \delta_{\alpha_r \alpha_{\tau(r)}}. \tag{4.10}$$

Once again we define the permutation $\pi = \sigma \circ \tau$ and change variables, using now $\text{sgn}(\sigma) = \text{sgn}(\tau) \text{sgn}(\pi)$. We again have

$$\sum_{\pi \in \mathcal{S}_k} \text{sgn}(\pi) \prod_{r=1}^k (X^{-T})_{i_r j_{\pi(r)}} = \det((X^{-T})_{IJ}) = (\det X)^{-1} \epsilon(I, J) (\det X_{I^c J^c}), \tag{4.11}$$

while now

$$\begin{aligned} \sum_{\tau \in \mathcal{S}_k} \text{sgn}(\tau) \sum_{\alpha_1, \dots, \alpha_k=1}^m \prod_{r=1}^k \delta_{\alpha_r \alpha_{\tau(r)}} &= \sum_{\tau \in \mathcal{S}_k} \text{sgn}(\tau) m^{\#(\text{cycles of } \tau)} \\ &= \sum_{\tau \in \mathcal{S}_k} (-1)^{\#(\text{odd cycles of } \tau)} (-m)^{\#(\text{cycles of } \tau)} \\ &= (-1)^k s(s+1) \cdots (s+k-1) \end{aligned} \tag{4.12}$$

by (3.9) since $s = -m$. This proves (2.2) for integer $s \leq -1$, when X is a complex matrix whose hermitian part is positive-definite; we conclude by Proposition 2.18 as before. \square

4.2. Symmetric Cayley identity

Proof of Theorem 2.2. In this case we use real bosons. So assume that $s = -m/2$ where m is a positive integer, and that X^{sym} is a real symmetric positive-definite matrix; and let us introduce bosonic variables $\varphi_i^{(\alpha)}$ for $i = 1, \dots, n$ and $\alpha = 1, \dots, m$. We then have

$$(\det X^{\text{sym}})^{-m/2} = \int \mathcal{D}\varphi e^{-\frac{1}{2} \varphi X^{\text{sym}} \varphi}. \tag{4.13}$$

The operator $\partial_{ij}^{\text{sym}} = \frac{1}{2}(1 + \delta_{ij}) \partial / \partial x_{ij}$ is exactly what is needed to bring down a factor $-\frac{1}{2}\varphi_i\varphi_j$ when acting on $e^{-\frac{1}{2}\varphi X^{\text{sym}}\varphi}$. (To lighten the notation, let us henceforth write X instead of X^{sym} .) We therefore have

$$\det(\partial_{IJ}^{\text{sym}})(\det X)^{-m/2} = \left(-\frac{1}{2}\right)^k \int \mathcal{D}\varphi \sum_{\sigma \in \mathcal{S}_k} \text{sgn}(\sigma) \sum_{\alpha_1, \dots, \alpha_k=1}^m \left(\prod_{r=1}^k \varphi_{i_r}^{(\alpha_r)} \varphi_{j_{\sigma(r)}}^{(\alpha_r)}\right) e^{-\frac{1}{2}\varphi X \varphi}. \tag{4.14}$$

Wick’s theorem for real bosons (Theorem A.3) applied to $\int \mathcal{D}\varphi \prod_{r=1}^k \varphi_{i_r}^{(\alpha_r)} \varphi_{j_{\sigma(r)}}^{(\alpha_r)}$ now gives rise to two types of contractions: those that only pair i ’s with j ’s, and those that pair at least one i with another i (and hence also a j with another j). The pairings of the first class are given by a sum over permutations τ , and yield

$$(\det X)^{-m/2} \sum_{\tau \in \mathcal{S}_k} \prod_{r=1}^k (X^{-1})_{i_r j_{\sigma(\tau(r))}} \delta_{\alpha_r \alpha_{\tau(r)}}. \tag{4.15}$$

Changing variables as in the alternate proof of Theorem 2.1, the sum over π again yields $(\det X)^{-1} \epsilon(I, J)(\det X_{I^c J^c})$, while the sum over τ yields

$$\sum_{\tau \in \mathcal{S}_k} (-1)^{\#\text{(odd cycles of } \tau)} (-m)^{\#\text{(cycles of } \tau)} = (-1)^k (2s)(2s + 1) \cdots (2s + k - 1) \tag{4.16}$$

by (3.9) since $s = -m/2$. Inserting these results into (4.14), we obtain (2.5).

Let us now show that for each pairing of the second class, the sum over σ yields zero. By hypothesis, at least one j is paired with another j , say j_p with j_q . Now let $\pi_{pq} \in \mathcal{S}_k$ be the permutation that interchanges p and q while leaving all other elements fixed. Then $\sigma \mapsto \sigma \circ \pi_{pq}$ is a sign-reversing involution for the sum in question.

Once again we argue that because the equality (2.5) holds for infinitely many s and for X in a nonempty open set, it must hold symbolically in s . \square

4.3. Antisymmetric Cayley identity

Proof of Theorem 2.3. In this case we use “real” fermions. So assume that s is a positive integer, and let us introduce “real” Grassmann variables $\theta_i^{(\alpha)}$ for $i = 1, \dots, 2m$ and $\alpha = 1, \dots, s$. We can then write

$$(\text{pf } X^{\text{antisym}})^s = \int \mathcal{D}\theta e^{\frac{1}{2}\theta X^{\text{antisym}}\theta}. \tag{4.17}$$

Now let $I = \{i_1, \dots, i_{2k}\}$ with $i_1 < \dots < i_{2k}$, so that

$$\text{pf}(\partial_{II}^{\text{antisym}}) = \frac{1}{2^k k!} \sum_{\sigma \in \mathcal{S}_{2k}} \text{sgn}(\sigma) \prod_{r=1}^k \frac{\partial}{\partial x_{i_{\sigma(2r-1)} i_{\sigma(2r)}}}. \tag{4.18}$$

Applying this to (4.17), we obtain

$$\begin{aligned} & \text{pf}(\partial_{II}^{\text{antisym}})(\text{pf} X^{\text{antisym}})^s \\ &= \int \mathcal{D}\theta \frac{1}{2^k k!} \sum_{\sigma \in \mathcal{S}_{2k}} \text{sgn}(\sigma) \sum_{\alpha_1, \dots, \alpha_k=1}^s \left(\prod_{r=1}^k \theta_{i_{\sigma(2r-1)}^{(\alpha_r)}} \theta_{i_{\sigma(2r)}^{(\alpha_r)}} \right) e^{\frac{1}{2}\theta X^{\text{antisym}}\theta}. \end{aligned} \tag{4.19}$$

(To lighten the notation, let us henceforth write X instead of X^{antisym} .) When X is an invertible real or complex matrix, Wick’s theorem for “real” fermions (Theorem A.15) gives

$$\begin{aligned} & \int \mathcal{D}\theta \left(\prod_{r=1}^k \theta_{i_{\sigma(2r-1)}^{(\alpha_r)}} \theta_{i_{\sigma(2r)}^{(\alpha_r)}} \right) e^{\frac{1}{2}\theta X\theta} \\ &= (\text{pf} X)^s \frac{1}{2^k k!} \sum_{\tau \in \mathcal{S}_{2k}} \text{sgn}(\tau) \prod_{r=1}^k (X^{-T})_{i_{\sigma(\tau(2r-1))} i_{\sigma(\tau(2r))}} \delta_{\alpha_{[\tau(2r-1)]} \alpha_{[\tau(2r)]}}. \end{aligned} \tag{4.20}$$

Once again we define the permutation $\pi = \sigma \circ \tau$ and change variables from (σ, τ) to (τ, π) , using $\text{sgn}(\sigma)\text{sgn}(\tau) = \text{sgn}(\pi)$. The sum over π gives

$$\frac{1}{2^k k!} \sum_{\pi \in \mathcal{S}_{2k}} \text{sgn}(\pi) \prod_{r=1}^k (X^{-T})_{i_{\pi(2r-1)} i_{\pi(2r)}} = \text{pf}((X^{-T})_{II}) = (\text{pf} X)^{-1} \epsilon(I) (\text{pf} X_{I^c I^c}) \tag{4.21}$$

by the pfaffian version of Jacobi’s identity [cf. (A.21)]. The remaining sums give

$$\frac{1}{2^k k!} \sum_{\tau \in \mathcal{S}_{2k}} \sum_{\alpha_1, \dots, \alpha_k=1}^s \prod_{r=1}^k \delta_{\alpha_{[\tau(2r-1)]} \alpha_{[\tau(2r)]}} = \sum_{M \in \mathcal{M}_{2k}} s^{\#(\text{cycles of } M \cup M_0)}, \tag{4.22}$$

where the latter sum runs over all perfect matchings M of $2k$ elements, and M_0 is some fixed perfect matching [in our case $(12)(34) \dots (2k - 12k)$]; we observe that if M and M_0 are thought of as edge sets on $[2k]$, then $M \cup M_0$ is the edge set of a graph in which each vertex has degree 2, and so is a union of cycles, showing that the right-hand side is well defined. Let us now show that this latter sum equals $s(s + 2) \dots (s + 2k - 2)$:

Let e_1, \dots, e_k be the edges of M_0 . Then each matching $M \in \mathcal{M}_{2k}$ induces a decomposition of the set $\{e_1, \dots, e_k\}$ into cycles, according to how those edges are traversed in $M \cup M_0$; in other words, M induces a permutation of $[k]$. Moreover, for each cycle C in this decomposition, there are $2^{|C|-1}$ ways of connecting up the vertices. Thus each permutation π of $[k]$ arises from $\prod_{C \in \pi} 2^{|C|-1} = 2^{k - \#(\text{cycles of } \pi)}$ different matchings M . We therefore have

$$\sum_{M \in \mathcal{M}_{2k}} s^{\#(\text{cycles of } M \cup M_0)} = 2^k \sum_{\pi \in \mathcal{S}_k} (s/2)^{\#(\text{cycles of } \pi)} \tag{4.23a}$$

$$= s(s + 2) \dots (s + 2k - 2). \tag{4.23b}$$

Once again we invoke Proposition 2.18 to conclude.²³ □

²³ We thank Alex Scott for help in cleaning up our proof of (4.23).

5. Proofs of Cayley-type identities by Grassmann representation of $\det(\partial)$

In this section we give an alternate Grassmann-based approach to proving Cayley-type identities: now it is the differential operator $\det(\partial)$ that is represented as a fermionic Gaussian integral. This technique is in our opinion very powerful: not only does it give the slickest proofs of the three main Cayley-type identities for square matrices; it also gives the only direct algebraic/combinatorial proofs (thus far) of the (considerably more difficult) rectangular Cayley identities.

The basic fact we will need is that an operator $\exp(a \cdot \partial)$ generates translation by a . More precisely, let R be a commutative ring containing the rationals; then for any polynomial $P(z_1, \dots, z_n)$ with coefficients in R and any constants $a_1, \dots, a_n \in R$, we have the *translation formula*

$$\exp\left(\sum_i a_i \frac{\partial}{\partial z_i}\right) P(z) = P(z + a). \tag{5.1}$$

(Here \exp is defined by Taylor series; note that all but finitely many terms will annihilate P .) Indeed, the identity (5.1) is nothing other than Taylor’s theorem for polynomials P . In particular we will use (5.1) when the commutative ring R consists of the even elements of some Grassmann algebra. Moreover, in our applications the elements a_i will be nilpotent, so that the Taylor series for $\exp(a \cdot \partial)$ is in fact finite.

We will also need a formula for the change of a determinant under a low-rank perturbation: see Appendix B.2.

Let us begin by explaining the general structure of all these proofs. We introduce a Grassmann integral representation for the differential operator $\det(\partial)$ and let it act on $(\det X)^s$. After a change of variables in the Grassmann integral, we obtain the desired quantity $(\det X)^{s-1}$ [or its generalization for minors I, J] multiplied by a purely combinatorial factor that is independent of the matrix X . We then proceed to calculate this combinatorial factor, which turns out to be an explicit polynomial in s .

Unfortunately, the “all-minors” versions of these proofs (i.e. those for $|I| = |J| = k < n$) are slightly more complicated than the “basic” versions (i.e. those for $I = J = [n]$). We have therefore structured our presentation so as to give the “basic” proof first, and then indicate the modifications needed to handle the “all-minors” case.

5.1. Ordinary Cayley identity

Proof of Theorem 2.1. We introduce Grassmann variables $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) and use the representation

$$\det(\partial) = \int \mathcal{D}_n(\eta, \bar{\eta}) e^{\bar{\eta}^T \partial \eta}, \tag{5.2}$$

where the subscript on \mathcal{D} serves to remind us of the length of the vectors in question, and we have employed the shorthand notation

$$\bar{\eta}^T \partial \eta \equiv \sum_{i,j=1}^n \bar{\eta}_i \frac{\partial}{\partial x_{ij}} \eta_j = \sum_{i,j=1}^n \bar{\eta}_i \eta_j \frac{\partial}{\partial x_{ij}}. \tag{5.3}$$

[Indeed, (5.2) is simply a special case of the fermionic Gaussian integral (A.74) where the coefficient ring R is the ring $\mathbb{Q}[\partial]$ of polynomials in the differential operators $\partial/\partial x_{ij}$.] Applying (5.2) to $(\det X)^s$ where s is a positive integer and using the translation formula (5.1), we obtain²⁴

²⁴ We have assumed here that s is a positive integer, because we have proven (5.1) only for polynomials P . Alternatively, we could avoid this assumption by proving (5.1) also for more general functions (e.g. powers of polynomials) when all the a_i are nilpotent.

$$\det(\partial)(\det X)^s = \int \mathcal{D}_n(\eta, \bar{\eta}) \det(X + \bar{\eta}\eta^T)^s \tag{5.4a}$$

$$= (\det X)^s \int \mathcal{D}_n(\eta, \bar{\eta}) \det(I + X^{-1}\bar{\eta}\eta^T)^s. \tag{5.4b}$$

(We assume here that X is an invertible real or complex matrix.) Let us now change variables from $(\eta, \bar{\eta})$ to $(\eta', \bar{\eta}') \equiv (\eta, X^{-1}\bar{\eta})$; we pick up a Jacobian $\det(X^{-1}) = (\det X)^{-1}$ and thus have

$$\det(\partial)(\det X)^s = (\det X)^{s-1} \int \mathcal{D}_n(\eta', \bar{\eta}') \det(I + \bar{\eta}'\eta'^T)^s. \tag{5.5}$$

Formula (5.5) expresses $\det(\partial)(\det X)^s$ as the desired quantity $(\det X)^{s-1}$ times a purely combinatorial factor

$$P(s, n) \equiv \int \mathcal{D}_n(\eta, \bar{\eta}) \det(I + \bar{\eta}\eta^T)^s, \tag{5.6}$$

which we now proceed to calculate. The matrix $I + \bar{\eta}\eta^T$ is a rank-1 perturbation of the identity matrix; by Lemma B.11 we have

$$\det(I + \bar{\eta}\eta^T)^s = (1 - \bar{\eta}^T\eta)^{-s} \tag{5.7a}$$

$$= \sum_{\ell=0}^{\infty} (-1)^\ell \binom{-s}{\ell} (\bar{\eta}^T\eta)^\ell \tag{5.7b}$$

where

$$\bar{\eta}^T\eta \equiv \sum_{i=1}^n \bar{\eta}_i\eta_i. \tag{5.8}$$

Since

$$\int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T\eta)^\ell = n! \delta_{\ell,n}, \tag{5.9}$$

it follows that

$$P(s, n) = (-1)^n \binom{-s}{n} n! \tag{5.10a}$$

$$= s(s+1) \cdots (s+n-1). \tag{5.10b}$$

This proves (2.1) when X is an invertible real or complex matrix and s is a positive integer; the general validity of the identity then follows from Proposition 2.18.²⁵

²⁵ Alternatively, one can observe (see the Remark after Theorem 2.1) that the Cayley identity (2.1) is equivalent to

$$\det(\partial) \det(X + A)^s = s(s+1) \cdots (s+n-1) \det(X + A)^{s-1}$$

for any fixed matrix A . This latter identity can be proven exactly as above, with X^{-1} replaced by $(X + A)^{-1}$, and it works whenever $X + A$ is invertible. Taking, for instance, $A = -cI$, we prove the identity for all real or complex matrices X whose spectrum does not contain the point c . Putting together these identities for all $c \in \mathbb{C}$, we cover all matrices X .

Let us now indicate the modifications needed to prove (2.2) for a minor I, J , where $I, J \subseteq [n]$ with $|I| = |J| = k$. We begin by introducing a Grassmann representation for $\det(\partial_{IJ})$:

$$\det(\partial_{IJ}) = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} \bar{\eta} \eta \right) e^{\bar{\eta}^T \partial \eta} \tag{5.11}$$

where

$$\left(\prod_{I^c, J^c} \bar{\eta} \eta \right) \equiv \bar{\eta}_{i_1} \eta_{j_1} \cdots \bar{\eta}_{i_{n-k}} \eta_{j_{n-k}} \tag{5.12}$$

and i_1, \dots, i_{n-k} (resp. j_1, \dots, j_{n-k}) are the elements of I^c (resp. J^c) in increasing order. Applying the translation formula (5.1) as before, we obtain

$$\det(\partial_{IJ})(\det X)^s = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} \bar{\eta} \eta \right) \det(X + \bar{\eta} \eta^T)^s \tag{5.13a}$$

$$= \epsilon(I, J)(\det X)^s \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} \bar{\eta} \eta \right) \det(I + X^{-1} \bar{\eta} \eta^T)^s. \tag{5.13b}$$

Once again we change variables from $(\eta, \bar{\eta})$ to $(\eta', \bar{\eta}') \equiv (\eta, X^{-1} \bar{\eta})$, picking up a Jacobian $\det(X^{-1}) = (\det X)^{-1}$; and we use the identity (5.7), obtaining

$$(\det \partial_{IJ})(\det X)^s = \epsilon(I, J)(\det X)^{s-1} \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} (X \bar{\eta} \eta) \right) \sum_{\ell=0}^{\infty} (-1)^\ell \binom{-s}{\ell} (\bar{\eta}^T \eta)^\ell \tag{5.14a}$$

$$= \epsilon(I, J)(\det X)^{s-1} \sum_{r_1, \dots, r_{n-k} \in [n]} \left(\prod_{p=1}^{n-k} X_{i_p, r_p} \right) \sum_{\ell=0}^{\infty} (-1)^\ell \binom{-s}{\ell} \times \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{p=1}^{n-k} \bar{\eta}_{r_p} \eta_{j_p} \right) (\bar{\eta}^T \eta)^\ell. \tag{5.14b}$$

The rules of Grassmann integration constrain the integral to be zero unless $\ell = k$ and $(r_1, \dots, r_{n-k}) = (j_{\sigma(1)}, \dots, j_{\sigma(n-k)})$ for some permutation $\sigma \in \mathcal{S}_{n-k}$. We therefore have

$$(\det \partial_{IJ})(\det X)^s = \epsilon(I, J)(\det X)^{s-1} \sum_{\sigma \in \mathcal{S}_{n-k}} \text{sgn}(\sigma) \left(\prod_{p=1}^{n-k} X_{i_p, j_{\sigma(p)}} \right) (-1)^k \binom{-s}{k} \times \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{p=1}^{n-k} \bar{\eta}_{j_p} \eta_{j_p} \right) k! \left(\prod_{i \in J} \bar{\eta}_i \eta_i \right) \tag{5.15a}$$

$$= \epsilon(I, J)(\det X)^{s-1} s(s+1) \cdots (s+k-1) \sum_{\sigma \in \mathcal{S}_{n-k}} \text{sgn}(\sigma) \prod_{p=1}^{n-k} X_{i_p, j_{\sigma(p)}} \tag{5.15b}$$

$$= \epsilon(I, J)(\det X)^{s-1} s(s+1) \cdots (s+k-1) \det(X|_{I^c J^c}). \tag{5.15c}$$

This proves (2.2). \square

5.2. Two useful lemmas for the all-minors case

Let us now pause to abstract the type of reasoning that was just used in proving the all-minors identity (2.2), as similar reasoning will be needed to prove the all-minors versions of the symmetric and rectangular Cayley identities and the all-principal-minors versions of the antisymmetric identities. The reader who is interested mainly in the case $I = J = [n]$ can skip this subsection on a first reading.

The key results of this subsection will be a pair of general formulae: Lemma 5.1 for “complex” fermions and Lemma 5.3 for “real” fermions. Important special cases of these formulae (which also have easier direct proofs) will be stated in Corollaries 5.2 and 5.4, respectively. These corollaries are, in fact, all we need to handle the all-minors symmetric Cayley identity (Section 5.3) and the all-principal-minors antisymmetric Cayley identity (Section 5.4). However, the rectangular Cayley identities (Sections 5.5–5.9) will need the full strength of Lemmas 5.1 and 5.3 to handle the all-minors case.

Let $I, J \subseteq [n]$ with $|I| = |J| = k$, and let I^c, J^c be the complementary subsets. We denote by i_1, \dots, i_{n-k} (resp. j_1, \dots, j_{n-k}) the elements of I^c (resp. J^c) in increasing order. Let $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) be Grassmann variables as before. Then, for any $n \times n$ matrices A, B , we define

$$\left(\prod (A\bar{\eta})(B\eta) \right)_{I^c, J^c} = (A\bar{\eta})_{i_1} (B\eta)_{j_1} \cdots (A\bar{\eta})_{i_{n-k}} (B\eta)_{j_{n-k}}. \tag{5.16}$$

In particular, $((A\bar{\eta})(B\eta))_{\emptyset, \emptyset} = 1$.

Now suppose that we have N further sets of (real) Grassmann variables $\theta_i^{(\alpha)}$ ($1 \leq i \leq n, 1 \leq \alpha \leq N$) – the case $N = 0$ is also allowed – and suppose that $f(\eta, \bar{\eta}, \theta)$ is a polynomial in the scalar products $\bar{\eta}^T \eta, \bar{\eta}^T \theta^{(\alpha)}, \eta^T \theta^{(\alpha)}$ and $\theta^{(\alpha)T} \theta^{(\beta)}$. Then a Grassmann integral of the form

$$\int \mathcal{D}_n(\eta, \bar{\eta}, \theta) \left(\prod_{i \in L} \bar{\eta}_i \eta_i \right) f(\eta, \bar{\eta}, \theta) \tag{5.17}$$

obviously takes the same value for all sets $L \subseteq [n]$ of the same cardinality.

Let us now show how a Grassmann integral involving $(\prod (A\bar{\eta})(B\eta))_{I^c, J^c}$ and $f(\eta, \bar{\eta}, \theta)$ can be written as a determinant containing A and B multiplied by a purely combinatorial factor:

Lemma 5.1. *Let $I, J \subseteq [n]$ with $|I| = |J| = k$, let A, B be $n \times n$ matrices, and let $f(\eta, \bar{\eta}, \theta)$ be a polynomial in the scalar products as specified above. Suppose further that the number N of additional fermion species is even. Then*

$$\begin{aligned} & \int \mathcal{D}_n(\eta, \bar{\eta}, \theta) \left(\prod (A\bar{\eta})(B\eta) \right)_{I^c, J^c} f(\eta, \bar{\eta}, \theta) \\ &= \det[(AB^T)_{I^c, J^c}] \int \mathcal{D}_n(\eta, \bar{\eta}, \theta) \left(\prod_{i \in L} \bar{\eta}_i \eta_i \right) f(\eta, \bar{\eta}, \theta) \end{aligned} \tag{5.18}$$

where $L \subseteq [n]$ is any set of cardinality $n - k$.

Proof. The expansion of the product (5.16) produces

$$\left(\prod (A\bar{\eta})(B\eta) \right)_{I^c, J^c} = \sum_{\substack{r_1, \dots, r_{n-k} \in [n] \\ s_1, \dots, s_{n-k} \in [n]}} \prod_{p=1}^{n-k} A_{i_p, r_p} B_{j_p, s_p} \bar{\eta}_{r_p} \eta_{s_p}. \tag{5.19}$$

Note that, by nilpotency of Grassmann variables, the product on the right-hand side of (5.19) is nonvanishing only if the indices r_1, \dots, r_{n-k} are all distinct and the indices s_1, \dots, s_{n-k} are also all distinct. Let us now integrate (5.19) against one of the monomials arising in the expansion of $f(\eta, \bar{\eta}, \theta)$. For each “site” $i \in [n]$, this monomial contains an even number of factors $\eta_i, \bar{\eta}_i$ or $\theta_i^{(\alpha)}$; and it must contain each one of the factors $\theta_i^{(\alpha)}$ ($1 \leq \alpha \leq N$) precisely once if the Grassmann integral over θ is to be nonvanishing. Since N is even, this means that at each site i we must either have both of the factors η_i and $\bar{\eta}_i$, or neither. In order to have a nonvanishing Grassmann integral over η and $\bar{\eta}$, the former situation must occur at k sites and the latter at $n - k$ sites; moreover, the latter $n - k$ sites must correspond precisely to the factors $\bar{\eta}_{r_p} \eta_{s_p}$ in (5.19). We can therefore assume that $(r_1, \dots, r_{n-k}) = (s_{\sigma(1)}, \dots, s_{\sigma(n-k)})$ for some permutation $\sigma \in \mathcal{S}_{n-k}$; the contributing terms in (5.19) are then

$$\sum_{\substack{r_1, \dots, r_{n-k} \in [n] \\ \text{all distinct}}} \sum_{\sigma \in \mathcal{S}_{n-k}} \prod_{p=1}^{n-k} A_{i_p, r_p} B_{j_p, r_{\sigma^{-1}(p)}} \bar{\eta}_{r_p} \eta_{r_{\sigma^{-1}(p)}}. \tag{5.20}$$

The factors $B_{j_p, r_{\sigma^{-1}(p)}}$ can be reordered freely; reordering of the Grassmann factors $\eta_{r_{\sigma^{-1}(p)}}$ yields a prefactor $\text{sgn}(\sigma)$. We therefore obtain

$$\sum_{\substack{r_1, \dots, r_{n-k} \in [n] \\ \text{all distinct}}} \sum_{\sigma \in \mathcal{S}_{n-k}} \text{sgn}(\sigma) \prod_{p=1}^{n-k} A_{i_p, r_p} B_{j_{\sigma(p)}, r_p} \bar{\eta}_{r_p} \eta_{r_p}. \tag{5.21}$$

By the remarks around (5.17), the integral

$$\int \mathcal{D}_n(\eta, \bar{\eta}, \theta) \left(\prod_{p=1}^{n-k} \bar{\eta}_{r_p} \eta_{r_p} \right) f(\eta, \bar{\eta}, \theta) \tag{5.22}$$

is independent of the choice of r_1, \dots, r_{n-k} (provided that they are all distinct) and hence can be pulled out. We are left with the factor

$$\sum_{\substack{r_1, \dots, r_{n-k} \in [n] \\ \text{all distinct}}} \sum_{\sigma \in \mathcal{S}_{n-k}} \text{sgn}(\sigma) \prod_{p=1}^{n-k} A_{i_p, r_p} B_{j_{\sigma(p)}, r_p}. \tag{5.23}$$

We can now remove the restriction that the r_1, \dots, r_{n-k} be all distinct, because the terms with two or more r_i equal cancel out when we sum over permutations with the factor $\text{sgn}(\sigma)$. So (5.23) equals

$$\sum_{r_1, \dots, r_{n-k} \in [n]} \sum_{\sigma \in \mathcal{S}_{n-k}} \text{sgn}(\sigma) \prod_{p=1}^{n-k} A_{i_p, r_p} B_{j_{\sigma(p)}, r_p} = \sum_{\sigma \in \mathcal{S}_{n-k}} \text{sgn}(\sigma) \prod_{p=1}^{n-k} (AB^T)_{i_p, j_{\sigma(p)}} \tag{5.24a}$$

$$= \det[(AB^T)_{I^c, J^c}]. \quad \square \tag{5.24b}$$

Remark. If N is odd, the situation is different: at each site i the monomial coming from f must now provide *exactly one* of the factors η_i and $\bar{\eta}_i$. This means that we can get a nonzero contribution only

when n is even and $k = n/2$. For instance, suppose that $N = 1$, $n = 2$, $I^c = \{i\}$ and $J^c = \{j\}$, and that $f(\eta, \bar{\eta}, \theta) = (\eta^T \theta)(\bar{\eta}^T \theta)$. Then

$$\int \mathcal{D}_2(\eta, \bar{\eta}, \theta) \left(\prod_{I^c, J^c} (A\bar{\eta})(B\eta) \right) (\eta^T \theta)(\bar{\eta}^T \theta) = -A_{i1} B_{j2} + A_{i2} B_{j1}. \tag{5.25}$$

We shall not consider this situation further, as we shall not need it in the sequel.

The following identity is what was used in the proof of the all-minors ordinary Cayley identity (2.2) and will also be used in the proof of the all-minors symmetric Cayley identity (2.5):

Corollary 5.2. *Let $I, J \subseteq [n]$ with $|I| = |J| = k$, let A, B be $n \times n$ matrices, let M be an invertible $n \times n$ matrix, and let ℓ be a nonnegative integer. Then*

$$\int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} (A\bar{\eta})(B\eta) \right) (\bar{\eta}^T M \eta)^\ell = k! \delta_{\ell, k} (\det M) \det[(AM^{-T}B^T)_{I^c J^c}]. \tag{5.26}$$

First proof. The case $M = I$ is an easy consequence of Lemma 5.1. The case of a general invertible matrix M can be reduced to the case $M = I$ by the change of variables $(\eta', \bar{\eta}') \equiv (M\eta, \bar{\eta})$, which picks up a Jacobian $\det M$ and replaces B by BM^{-1} . \square

Here is a simple direct proof of Corollary 5.2 that avoids the combinatorial complexity of the full Lemma 5.1 and instead relies on standard facts from the theory of Grassmann–Berezin integration (Appendix A)²⁶:

Second proof. It is easy to see that the Grassmann integral is nonvanishing only when $\ell = k$. So in this case we can replace $(\bar{\eta}^T M \eta)^k$ in the integrand by $k! \exp(\bar{\eta}^T M \eta)$. The claim is then an immediate consequence of Wick’s theorem for “complex” fermions (Theorem A.16). \square

In order to handle the all-principal-minors versions of the antisymmetric identities, we shall need formulae analogous to Lemma 5.1 and Corollary 5.2, but for “real” rather than “complex” fermions. So let $I \subseteq [2m]$ with $|I| = 2k$, and let I^c be the complementary subset. We denote by i_1, \dots, i_{2m-2k} the elements of I^c in increasing order. Let $\theta_1, \dots, \theta_{2m}$ be “real” Grassmann variables. Then, for any $2m \times 2m$ matrix C , we define

$$\left(\prod_{I^c} (C\theta) \right) \equiv (C\theta)_{i_1} \cdots (C\theta)_{i_{2m-2k}} \tag{5.27}$$

(note that the order of factors θ is crucial here because they anticommute). In particular, $\left(\prod_{\emptyset} (C\theta) \right) = 1$.

Now suppose that we have N further sets of (real) Grassmann variables $\chi_i^{(\alpha)}$ ($1 \leq i \leq 2m$, $1 \leq \alpha \leq N$) – the case $N = 0$ is also allowed – and suppose that $f(\theta, \chi)$ is a polynomial in the symplectic products $\frac{1}{2}\theta J\theta$, $\theta J\chi^{(\alpha)}$, $\frac{1}{2}\chi^{(\alpha)} J\chi^{(\alpha)}$ and $\chi^{(\alpha)} J\chi^{(\beta)}$, where the $2m \times 2m$ matrix J is defined by

²⁶ Said another way: The proof of Wick’s theorem for “complex” fermions (Theorem A.16) requires combinatorial work similar in difficulty to that occurring in the proof of Lemma 5.1. But since Wick’s theorem is a “standard” result, we can employ it without reproving it *ab initio*.

$$J = \left(\begin{array}{cc|cc} 0 & 1 & & \\ -1 & 0 & & \\ \hline & & 0 & 1 \\ & & -1 & 0 \\ & & & \ddots \end{array} \right). \tag{5.28}$$

Now consider a Grassmann integral of the form

$$\int \mathcal{D}_{2m}(\theta, \chi) \left(\prod_{i \in L} \theta_i \right) f(\theta, \chi) \tag{5.29}$$

for some set $L \subseteq [2m]$, where the product $\prod_{i \in L} \theta_i$ is understood to be written from left to right in increasing order of the indices. Clearly $|L|$ must be even for the integral (5.29) to be nonvanishing. But more is true: L must in fact be a union of pairs $\{2j - 1, 2j\}$, otherwise (5.29) will again vanish.²⁷ To see this, it suffices to observe that all the variables θ or $\chi^{(\alpha)}$ with index $2j - 1$ or $2j$ appear in pairs in $f(\theta, \chi)$; therefore, in order to have an integrand that is even in these variables, the set L must contain either both of $2j - 1$ and $2j$ or neither. Let us call a set $L \subseteq [2m]$ well-paired if it is a union of pairs $\{2j - 1, 2j\}$. Obviously the integral (5.29) takes the same value for all well-paired sets $L \subseteq [2m]$ of the same cardinality.

Let us now show how a Grassmann integral involving $(\prod(C\theta))_{I_C}$ and $f(\theta, \chi)$ can be written as a pfaffian containing C multiplied by a purely combinatorial factor:

Lemma 5.3. *Let $I \subseteq [2m]$ with $|I| = 2k$, let C be a $2m \times 2m$ matrix, and let $f(\theta, \chi)$ be a polynomial in the symplectic products as specified above. Then*

$$\begin{aligned} & \int \mathcal{D}_{2m}(\theta, \chi) \left(\prod(C\theta) \right)_{I_C} f(\theta, \chi) \\ &= \text{pf}[(C J C^T)_{I_C}] \int \mathcal{D}_{2m}(\theta, \chi) \left(\prod_{i \in L} \theta_i \right) f(\theta, \chi) \end{aligned} \tag{5.30}$$

where $L \subseteq [2m]$ is any well-paired set of cardinality $2m - 2k$.

Proof. The expansion of the product (5.27) produces

$$\left(\prod(C\theta) \right)_{I_C} = \sum_{r_1, \dots, r_{2m-2k} \in [2m]} \left(\prod_{p=1}^{2m-2k} C_{i_p, r_p} \right) \theta_{r_1} \cdots \theta_{r_{2m-2k}}. \tag{5.31}$$

By nilpotency of Grassmann variables, the product on the right-hand side of (5.31) is nonvanishing only if the indices r_1, \dots, r_{2m-2k} are all distinct. Moreover, as noted above, the set $R = \{r_1, \dots, r_{2m-2k}\}$ must be well-paired if the integral is to be nonvanishing. Let us denote by r'_1, \dots, r'_{2m-2k} the elements

²⁷ This statement of course pertains to our specific choice of J , and would be modified in the obvious way if we had chosen a different convention for J .

of R in increasing order. Then we can write $(r_1, \dots, r_{2m-2k}) = (r'_{\sigma(1)}, \dots, r'_{\sigma(2m-2k)})$ for some permutation $\sigma \in \mathcal{S}_{2m-2k}$. The contributing terms in (5.31) are then

$$\sum_{\substack{R \subseteq [2m] \\ |R|=2m-2k \\ R \text{ well-paired}}} \sum_{\sigma \in \mathcal{S}_{2m-2k}} \left(\prod_{p=1}^{2m-2k} C_{i_p, r'_{\sigma(p)}} \right) \theta'_{r'_{\sigma(1)}} \cdots \theta'_{r'_{\sigma(2m-2k)}}. \tag{5.32}$$

We can now reorder the factors θ into increasing order, yielding a factor $\text{sgn}(\sigma)$, so that (5.32) becomes

$$\sum_{\substack{R \subseteq [2m] \\ |R|=2m-2k \\ R \text{ well-paired}}} \sum_{\sigma \in \mathcal{S}_{2m-2k}} \text{sgn}(\sigma) \left(\prod_{p=1}^{2m-2k} C_{i_p, r'_{\sigma(p)}} \right) \left(\prod_{i \in R} \theta_i \right). \tag{5.33}$$

We now multiply by $f(\theta, \chi)$ and integrate $\mathcal{D}_{2m}(\theta, \chi)$. As noted previously, the integral

$$\int \mathcal{D}_{2m}(\theta, \chi) \left(\prod_{i \in R} \theta_i \right) f(\theta, \chi) \tag{5.34}$$

is independent of the choice of R (provided that it is well-paired and of cardinality $2m - 2k$) and hence can be pulled out. In order to calculate the prefactor, let us substitute the defining structure of well-paired sets R by writing $r'_{2h-1} = 2j_h - 1$ and $r'_{2h} = 2j_h$ for suitable indices $j_1 < \dots < j_{m-k}$ in $[m]$. We are left with

$$\sum_{\substack{1 \leq j_1 < \dots < j_{m-k} \leq m \\ r'_{2h-1} \equiv 2j_h - 1, r'_{2h} \equiv 2j_h}} \sum_{\sigma \in \mathcal{S}_{2m-2k}} \text{sgn}(\sigma) \prod_{p=1}^{2m-2k} C_{i_p, r'_{\sigma(p)}}. \tag{5.35}$$

Alternatively, we can sum over all *distinct* $j_1, \dots, j_{m-k} \in [m]$ by inserting a factor $1/(m-k)!$.²⁸ Remarkably, we can now remove the restriction that the j_1, \dots, j_{m-k} be all distinct, because the terms with two or more j_α equal cancel out when we sum over permutations with the factor $\text{sgn}(\sigma)$.²⁹ We can also reorder the factors C freely. Therefore (5.35) equals

$$\begin{aligned} & \frac{1}{(m-k)!} \sum_{\substack{j_1, \dots, j_{m-k} \in [m] \\ r'_{2h-1} \equiv 2j_h - 1, r'_{2h} \equiv 2j_h}} \sum_{\sigma \in \mathcal{S}_{2m-2k}} \text{sgn}(\sigma) \prod_{p=1}^{2m-2k} C_{i_{\sigma^{-1}(p)}, r'_p} \\ &= \frac{1}{(m-k)!} \sum_{\substack{j_1, \dots, j_{m-k} \in [m] \\ r'_{2h-1} \equiv 2j_h - 1, r'_{2h} \equiv 2j_h}} \sum_{\sigma \in \mathcal{S}_{2m-2k}} \text{sgn}(\sigma) \prod_{q=1}^{m-k} C_{i_{\sigma^{-1}(2q-1)}, r'_{2q-1}} C_{i_{\sigma^{-1}(2q)}, r'_{2q}} \end{aligned}$$

²⁸ Such a factor might not exist in a general commutative ring R . But we can argue as follows: First we prove the identity when $R = \mathbb{R}$; then we observe that since both sides of the identity are polynomials with integer coefficients in the matrix elements of C , the identity must hold as *polynomials*; therefore the identity holds when the matrix elements of C are specialized to arbitrary values in an arbitrary commutative ring R .

²⁹ If $j_\alpha = j_\beta$ with $\alpha \neq \beta$, it suffices to use the involution exchanging $r'_{2\alpha-1}$ with $r'_{2\beta-1}$ (or alternatively the involution exchanging $r'_{2\alpha}$ with $r'_{2\beta}$).

$$\begin{aligned}
 &= \frac{1}{2^{m-k}(m-k)!} \sum_{\substack{s_1, \dots, s_{m-k} \in [2m] \\ t_1, \dots, t_{m-k} \in [2m]}} \sum_{\sigma \in S_{2m-2k}} \operatorname{sgn}(\sigma) \prod_{q=1}^{m-k} C_{i_{\sigma^{-1}(2q-1)}, s_q} C_{i_{\sigma^{-1}(2q)}, t_q} J_{s_q t_q} \\
 &= \operatorname{pf}[(CJC^T)_{I_c I_c}]. \quad \square
 \end{aligned} \tag{5.36}$$

The following identity is what will be used in the proof of the all-principal-minors antisymmetric Cayley identity (2.8):

Corollary 5.4. *Let $I \subseteq [2m]$ with $|I| = 2k$, let C be a $2m \times 2m$ matrix, let M be an invertible antisymmetric $2m \times 2m$ matrix, and let ℓ be a nonnegative integer. Then*

$$\int \mathcal{D}_{2m}(\theta) \left(\prod_{I^c} (C\theta) \right)_{I^c} \left(\frac{1}{2} \theta^T M \theta \right)^\ell = k! \delta_{\ell, k} (\operatorname{pf} M) \operatorname{pf}[(CM^{-T}C^T)_{I^c, I^c}]. \tag{5.37}$$

First proof. The case $M = J$ is an easy consequence of Lemma 5.3. For a general real antisymmetric matrix M , we can use the decomposition $M = AJA^T$ (Lemma B.17) and the change of variables $\theta' = A^T\theta$ (which picks up a Jacobian $\det A = \operatorname{pf} M$) to reduce to the case $M = J$.

Finally, if M is an antisymmetric matrix with coefficients in an arbitrary commutative ring R , we argue that both sides of (5.37) are polynomials with integer coefficients in the matrix elements of M ; since they agree for all real values of those matrix elements (or even for real values in a nonempty open set), they must agree as polynomials; and this implies that they agree for arbitrary values of the matrix elements in an arbitrary commutative ring R . \square

Here is a simple direct proof of Corollary 5.4 that avoids the combinatorial complexity of the full Lemma 5.3 and instead relies on standard facts from the theory of Grassmann–Berezin integration:

Second proof. It is easy to see that the Grassmann integral is nonvanishing only for $\ell = k$. So in this case we can replace $(\frac{1}{2}\theta^T M \theta)^k$ in the integrand by $k! \exp(\frac{1}{2}\theta^T M \theta)$. The claim is then an immediate consequence of Wick’s theorem for “real” fermions (Theorem A.15). \square

5.3. Symmetric Cayley identity

With Corollary 5.2 in hand, we can this time proceed directly to the proof of the all-minors identity.

Proof of Theorem 2.2. Recall that the matrix $\partial^{\operatorname{sym}}$ is given by

$$(\partial^{\operatorname{sym}})_{ij} = \begin{cases} \partial/\partial x_{ii} & \text{if } i = j, \\ \frac{1}{2} \partial/\partial x_{ij} & \text{if } i < j, \\ \frac{1}{2} \partial/\partial x_{ji} & \text{if } i > j. \end{cases} \tag{5.38}$$

As before, we introduce Grassmann variables $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) and use the representation

$$\det(\partial_{IJ}^{\operatorname{sym}}) = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} \bar{\eta} \eta \right) \exp \left[\sum_{i \leq j} \frac{1}{2} (\bar{\eta}_i \eta_j + \bar{\eta}_j \eta_i) \frac{\partial}{\partial x_{ij}} \right]. \tag{5.39}$$

By the translation formula (5.1), we have

$$\det(\partial_{IJ}^{\operatorname{sym}}) f(\{x_{ij}\}_{i \leq j}) = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} \bar{\eta} \eta \right) f \left(\left\{ x_{ij} + \frac{1}{2} (\bar{\eta}_i \eta_j - \eta_i \bar{\eta}_j) \right\}_{i \leq j} \right) \tag{5.40}$$

for an arbitrary polynomial f . We shall use this formula in the case $f(\{x_{ij}\}_{i \leq j}) = \det(X^{\text{sym}})^s$ where s is a positive integer, so that

$$\det(\partial_{IJ}^{\text{sym}}) \det(X^{\text{sym}})^s = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\eta} \eta \right)_{I^c, J^c} \det \left[X^{\text{sym}} + \frac{1}{2} (\bar{\eta} \eta^T - \eta \bar{\eta}^T) \right]^s. \tag{5.41}$$

It is convenient to introduce the shorthand

$$(X^{\text{sym}})^{\text{trans}} \equiv X^{\text{sym}} + \frac{1}{2} (\bar{\eta} \eta^T - \eta \bar{\eta}^T) \tag{5.42}$$

for the argument of \det .

Let us now change variables from $(\eta, \bar{\eta})$ to $(\eta', \bar{\eta}') \equiv (\eta, (X^{\text{sym}})^{-1} \bar{\eta})$, with Jacobian $(\det X^{\text{sym}})^{-1}$. Dropping primes from the new variables, we observe that the expression for the translated matrix can be written as

$$(X^{\text{sym}})^{\text{trans}} = X^{\text{sym}} \left[I + \frac{1}{2} (\bar{\eta} \eta^T - (X^{\text{sym}})^{-1} \eta \bar{\eta}^T X^{\text{sym}}) \right], \tag{5.43}$$

so that

$$\det(X^{\text{sym}})^{\text{trans}} = (\det X^{\text{sym}}) \det \left[I + \frac{1}{2} (\bar{\eta} \eta^T - (X^{\text{sym}})^{-1} \eta \bar{\eta}^T X^{\text{sym}}) \right]. \tag{5.44}$$

Applying Corollary B.12 to the rightmost determinant yields

$$\det \left[I + \frac{1}{2} (\bar{\eta} \eta^T - (X^{\text{sym}})^{-1} \eta \bar{\eta}^T X^{\text{sym}}) \right] = \left(1 - \frac{1}{2} \bar{\eta}^T \eta \right)^{-2} \tag{5.45}$$

so that we are left with the Grassmann-integral expression

$$\begin{aligned} & \det(\partial_{IJ}^{\text{sym}}) (\det X^{\text{sym}})^s \\ &= (\det X^{\text{sym}})^{s-1} \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod (X^{\text{sym}} \bar{\eta})(\eta) \right)_{I^c, J^c} \left(1 - \frac{1}{2} \bar{\eta}^T \eta \right)^{-2s}. \end{aligned} \tag{5.46}$$

Now insert the expansion

$$\left(1 - \frac{1}{2} \bar{\eta}^T \eta \right)^{-2s} = \sum_{\ell=0}^{\infty} \left(-\frac{1}{2} \right)^\ell \binom{-2s}{\ell} (\bar{\eta}^T \eta)^\ell \tag{5.47}$$

into (5.46) and use Corollary 5.2: we obtain

$$\det(\partial_{IJ}^{\text{sym}}) (\det X^{\text{sym}})^s = (\det X^{\text{sym}})^{s-1} \epsilon(I, J) (\det X_{I^c J^c}^{\text{sym}}) k! \left(-\frac{1}{2} \right)^k \binom{-2s}{k}. \tag{5.48}$$

This proves (2.5) when X^{sym} is an invertible real or complex symmetric $n \times n$ matrix and s is a positive integer; the general validity of the identity then follows from Proposition 2.18. \square

Remark. A slight variant of this proof employs the factorization $X^{\text{sym}} = AA^T$ when X^{sym} is a real symmetric positive-definite $n \times n$ matrix (Lemma B.16); we then use the change of variables $(\eta', \bar{\eta}') \equiv (A^{-1}\eta, A^{-1}\bar{\eta})$, with Jacobian $(\det A)^{-2} = (\det X^{\text{sym}})^{-1}$. This slightly shortens the calculations, but has the disadvantage that the proof is no longer purely algebraic/combinatorial, because it invokes a decomposition that is valid for *real symmetric positive-definite* matrices in order to define the needed change of variables (which involves A and not just X^{sym}). We therefore prefer to avoid matrix decompositions wherever we can (which is unfortunately not always). See also the Remark at the end of Section 5.4, and the discussion in Appendix B.3.

5.4. Antisymmetric Cayley identity

With Corollary 5.4 in hand, we can again proceed directly to the proof of the all-(principal-)minors identity.

Proof of Theorem 2.3. The $2m \times 2m$ matrix $\partial^{\text{antisym}}$ is given by

$$(\partial^{\text{antisym}})_{ij} = \begin{cases} 0 & \text{if } i = j, \\ \partial/\partial x_{ij} & \text{if } i < j, \\ -\partial/\partial x_{ji} & \text{if } i > j. \end{cases} \tag{5.49}$$

We introduce “real” Grassmann variables θ_i ($1 \leq i \leq 2m$) and use the representation

$$\text{pf}(\partial_{II}^{\text{antisym}}) = \epsilon(I) \int \mathcal{D}_{2m}(\theta) \left(\prod \theta \right)_{I^c} \exp \left[\sum_{i < j} \theta_i \theta_j \frac{\partial}{\partial x_{ij}} \right]. \tag{5.50}$$

By the translation formula (5.1),

$$\text{pf}(\partial_{II}^{\text{antisym}}) f(\{x_{ij}\}_{i < j}) = \epsilon(I) \int \mathcal{D}_{2m}(\theta) \left(\prod \theta \right)_{I^c} f(\{x_{ij} + \theta_i \theta_j\}_{i < j}) \tag{5.51}$$

for an arbitrary polynomial f . We shall use this formula in the case $f(\{x_{ij}\}_{i < j}) = (\text{pf } X^{\text{antisym}})^s$ where s is a positive even integer. It is convenient to introduce the shorthand

$$(X^{\text{antisym}})^{\text{trans}} = X^{\text{antisym}} + \theta \theta^T \tag{5.52}$$

for the argument of pf.

Suppose now that X^{antisym} is a $2m \times 2m$ real antisymmetric matrix of rank $2m$. Then Lemma B.17 guarantees that we can find a matrix $A \in GL(2m)$ such that $X^{\text{antisym}} = AJA^T$, where

$$J = \left(\begin{array}{cc|cc} 0 & 1 & & \\ -1 & 0 & & \\ \hline & & 0 & 1 \\ & & -1 & 0 \\ & & & \ddots \end{array} \right) \tag{5.53}$$

is the standard $2m \times 2m$ symplectic form (note that with this convention $\text{pf } J = +1$ for any m). We have

$$(X^{\text{antisym}})^{\text{trans}} = A(J + A^{-1}\theta\theta^T A^{-T})A^T \tag{5.54}$$

and

$$\text{pf}(X^{\text{antisym}})^{\text{trans}} = (\text{pf } X^{\text{antisym}}) \text{pf}(J + A^{-1}\theta\theta^T A^{-T}). \tag{5.55}$$

Now change variables from θ to $\theta' \equiv A^{-1}\theta$, with Jacobian $(\det A)^{-1} = (\text{pf } X)^{-1}$. Dropping primes, we are left with

$$\text{pf}(\theta_{II}^{\text{antisym}})(\text{pf } X^{\text{antisym}})^s = \epsilon(I)(\text{pf } X^{\text{antisym}})^{s-1} \int \mathcal{D}_{2m}(\theta) \left(\prod(A\theta)\right)_{Ic} \text{pf}(J + \theta\theta^T)^s. \tag{5.56}$$

We can now write

$$\text{pf}(J + \theta\theta^T)^s = \det(J + \theta\theta^T)^{s/2} = \det(I - J\theta\theta^T)^{s/2} \tag{5.57}$$

since $\text{pf } J = \det J = +1$ and $J^{-1} = -J$.³⁰ The matrix $I - J\theta\theta^T$ is a rank-1 perturbation of the identity matrix; applying Lemma B.11 with vectors $u = -J\theta$ and $v = \theta$, we obtain

$$\det(I - J\theta\theta^T) = (1 - \theta^T J\theta)^{-1}. \tag{5.58}$$

Using Corollary 5.4, we have

$$\begin{aligned} &\int \mathcal{D}_{2m}(\theta) \left(\prod(A\theta)\right)_{Ic} \text{pf}(J + \theta\theta^T)^s \\ &= \int \mathcal{D}_{2m}(\theta) \left(\prod(A\theta)\right)_{Ic} (1 - \theta^T J\theta)^{-s/2} \end{aligned} \tag{5.59a}$$

$$\begin{aligned} &= (-2)^k k! \binom{-s/2}{k} (\text{pf } J) \text{pf}[(AJ^{-T}A^T)_{IcIc}] \\ &= s(s+2) \cdots (s+2m-2) \text{pf}(X_{IcIc}^{\text{antisym}}) \end{aligned} \tag{5.59b}$$

since $\text{pf } J = +1$ and $J^{-T} = J$. This proves (2.8) when X^{antisym} is a real antisymmetric matrix of rank $2m$ and s is a positive even integer; the general validity of the identity then follows from Proposition 2.18. \square

Remark. Although the statement of Theorem 2.3 is purely algebraic/combinatorial, the foregoing proof is unfortunately not *purely* algebraic/combinatorial, as it invokes the decomposition $X^{\text{antisym}} = AJA^T$ that is valid for *real* full-rank antisymmetric matrices in order to define the needed change of variables (which involves A and not just X^{antisym}). A similar decomposition will be invoked in the proofs of the rectangular Cayley identities (Sections 5.5, 5.6, 5.7 and 5.9). This contrasts with the proofs of the ordinary and symmetric Cayley identities (Sections 5.1 and 5.3), where we were able to define the needed change of variables in terms of the original matrix X or X^{sym} . The matrix factorization lemmas needed for the former proofs are collected and proven in Appendix B.3.

³⁰ The equality $\text{pf}(J + \theta\theta^T) = \det(J + \theta\theta^T)^{1/2}$ follows from the general fact $(\text{pf } M)^2 = \det M$ together with the observation that both $\text{pf}(J + \theta\theta^T)$ and $\det(J + \theta\theta^T)$ are elements of the Grassmann algebra with constant term 1, and that an element of the Grassmann algebra with constant term 1 has a *unique* square root in the Grassmann algebra with this property.

5.5. Two-matrix rectangular Cayley identity

In the remaining subsections of this section we shall prove the various rectangular Cayley identities that were stated in Section 2.2. It is convenient to begin with the two-matrix rectangular Cayley identity (Theorem 2.6), whose proof is somewhat less intricate than that of the corresponding one-matrix identities (Theorems 2.7 and 2.8). Indeed, the one-matrix rectangular symmetric and antisymmetric Cayley identities are related to the two-matrix identity in roughly the same way as the symmetric and antisymmetric Cayley identities are related to the ordinary one.

Here we will need the full strength of Lemma 5.1 to handle the all-minors case.

Proof of Theorem 2.6. We begin once again by representing the differential operator as a Grassmann integral: exploiting Corollary B.6, we have

$$\det(\partial_X \partial_Y^T) = \det\left(\begin{array}{c|c} 0_m & \partial_X \\ \hline -\partial_Y^T & I_n \end{array}\right) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) e^{\bar{\eta}^T \eta + \bar{\psi}^T \partial_X \eta + \psi^T \partial_Y \bar{\eta}}. \tag{5.60}$$

Here $\psi_i, \bar{\psi}_i$ ($1 \leq i \leq m$) and $\eta_j, \bar{\eta}_j$ ($1 \leq j \leq n$) are Grassmann variables, and the subscripts on \mathcal{D} serve to remind us of the length of each vector; shorthand notations for index summations are understood, e.g. $\bar{\psi}^T \partial_X \eta \equiv \sum_{i=1}^m \sum_{j=1}^n \bar{\psi}_i \eta_j \partial / \partial x_{ij}$. For a general minor $I, J \subseteq [m]$ with $|I| = |J| = k$, we have (writing $L = \{m + 1, \dots, m + n\}$)

$$\det[(\partial_X \partial_Y^T)_{IJ}] = \det\left[\begin{array}{c|c} 0_m & \partial_X \\ \hline -\partial_Y^T & I_n \end{array}\right]_{I \cup L, J \cup L} \tag{5.61a}$$

$$= \epsilon(I, J) \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\psi} \psi\right)_{I^c, J^c} e^{\bar{\eta}^T \eta + \bar{\psi}^T \partial_X \eta + \psi^T \partial_Y \bar{\eta}}. \tag{5.61b}$$

Applying the translation formula (5.1) to the whole set of variables $\{x_{ij}, y_{ij}\}$ produces

$$\begin{aligned} &\det[(\partial_X \partial_Y^T)_{IJ}] f(X, Y) \\ &= \epsilon(I, J) \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\psi} \psi\right)_{I^c, J^c} e^{\bar{\eta}^T \eta} f(X + \bar{\psi} \eta^T, Y + \psi \bar{\eta}^T) \end{aligned} \tag{5.62}$$

for an arbitrary polynomial f . We shall use this formula in the case $f(X, Y) = \det(XY^T)^s$ where s is a positive integer. It is convenient to introduce the shorthands

$$X^{\text{trans}} \equiv X + \bar{\psi} \eta^T, \tag{5.63a}$$

$$Y^{\text{trans}} \equiv Y + \psi \bar{\eta}^T \tag{5.63b}$$

for the arguments of f .

Suppose now that X and Y are real $m \times n$ matrices of rank m that are sufficiently close to the matrix \hat{I}_{mn} defined by

$$(\hat{I}_{mn})_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{5.64}$$

[Note that $\hat{I}_{mn} \equiv (I_m, 0_{m \times (n-m)})$ when $m \leq n$, and $\hat{I}_{mn} = \hat{I}_{nm}^T$ otherwise. Henceforth we shall drop the subscripts mn on \hat{I}_{mn} to lighten the notation.] Then by Lemma B.18 we can find matrices $P, P' \in GL(m)$

and $Q \in GL(n)$ such that $X = P\hat{I}Q$ and $Y = P'\hat{I}Q^{-T}$. We have

$$\det(XY^T) = \det(P\hat{I}Q Q^{-1}\hat{I}^T P'^T) = \det(P P'^T) = \det(P) \det(P') \tag{5.65}$$

and

$$\begin{aligned} X^{\text{trans}}(Y^{\text{trans}})^T &\equiv (X + \bar{\psi}\eta^T)(Y^T - \bar{\eta}\psi^T) \\ &= P[\hat{I} + P^{-1}(\bar{\psi}\eta^T)Q^{-1}]Q Q^{-1}[\hat{I}^T - Q(\bar{\eta}\psi^T)P'^{-T}]P'^T. \end{aligned} \tag{5.66}$$

Let us now change variables from $(\psi, \bar{\psi}, \eta, \bar{\eta})$ to $(\psi', \bar{\psi}', \eta', \bar{\eta}') \equiv (P'^{-1}\psi, P^{-1}\bar{\psi}, Q^{-T}\eta, Q\bar{\eta})$, with Jacobian $(\det P)^{-1}(\det P')^{-1} = \det(XY^T)^{-1}$. In the new variables we have (dropping now the primes from the notation)

$$X^{\text{trans}}(Y^{\text{trans}})^T = P(\hat{I} + \bar{\psi}\eta^T)(\hat{I}^T - \bar{\eta}\psi^T)P'^T, \tag{5.67}$$

and the translated determinant is given by

$$\det[(X^{\text{trans}})(Y^{\text{trans}})^T] = \det(XY^T) \det[(\hat{I} + \bar{\psi}\eta^T)(\hat{I}^T - \bar{\eta}\psi^T)], \tag{5.68}$$

so that

$$\begin{aligned} &\det[(\partial_X \partial_Y^T)_{IJ}] \det(XY^T)^s \\ &= \epsilon(I, J) \det(XY^T)^{s-1} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I, J} (P\bar{\psi})(P'\psi) \right) \\ &\quad \times e^{\bar{\eta}^T \eta} \det[(\hat{I} + \bar{\psi}\eta^T)(\hat{I}^T - \bar{\eta}\psi^T)]^s. \end{aligned} \tag{5.69}$$

Let us now split the vectors η and $\bar{\eta}$ as

$$(\eta_1, \dots, \eta_n) = (\lambda_1, \dots, \lambda_m, \chi_1, \dots, \chi_{n-m}), \tag{5.70a}$$

$$(\bar{\eta}_1, \dots, \bar{\eta}_n) = (\bar{\lambda}_1, \dots, \bar{\lambda}_m, \bar{\chi}_1, \dots, \bar{\chi}_{n-m}) \tag{5.70b}$$

so that

$$(\hat{I} + \bar{\psi}\eta^T)(\hat{I}^T - \bar{\eta}\psi^T) = I_m + \bar{\psi}\lambda^T - \bar{\lambda}\psi^T + c\bar{\psi}\psi^T \tag{5.71}$$

with $c = \bar{\lambda}^T \lambda + \bar{\chi}^T \chi$. This matrix has the form of a low-rank perturbation $I_m + \sum_{\alpha=1}^2 u_\alpha v_\alpha^T$, with vectors $\{u_\alpha\}, \{v_\alpha\}$ given by:

$$\begin{array}{c|cc} \alpha & u_\alpha & v_\alpha \\ \hline 1 & \bar{\psi} & \lambda + c\psi \\ 2 & \bar{\lambda} & -\psi \end{array} \tag{5.72}$$

By Lemma B.11 we can write the needed determinant as the determinant of a 2×2 matrix:

$$\begin{aligned} & \det[(\hat{I} + \bar{\psi} \eta^T)(\hat{I}^T - \bar{\eta} \psi^T)] \\ &= \det^{-1} \begin{pmatrix} 1 + (\lambda + c\psi)^T \bar{\psi} & (\lambda + c\psi)^T \bar{\lambda} \\ -\psi^T \bar{\psi} & 1 - \psi^T \bar{\lambda} \end{pmatrix} \end{aligned} \tag{5.73a}$$

$$= \det^{-1} \begin{pmatrix} 1 + \lambda^T \bar{\psi} & c + \lambda^T \bar{\lambda} \\ -\psi^T \bar{\psi} & 1 - \psi^T \bar{\lambda} \end{pmatrix} \tag{5.73b}$$

$$= [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi) - (\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^{-1}, \tag{5.73c}$$

where the second equality is a row operation (row 1 \rightarrow row 1 + c row 2). We therefore have

$$\begin{aligned} \det[(\partial_X \partial_Y^T)_{IJ}] \det(XY^T)^s &= \epsilon(I, J) \det(XY^T)^{s-1} \int \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\chi}^T \chi} \\ &\quad \times \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_m(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} \left(\prod_{I \subset J^c} (P\bar{\psi})(P'\psi) \right)_{I \subset J^c} \\ &\quad \times [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi) - (\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^{-s}. \end{aligned} \tag{5.74}$$

Note in particular that the integrand depends on $\psi, \bar{\psi}, \lambda, \bar{\lambda}$ only via scalar products. This allows us to apply Lemma 5.1 to the integral over $\psi, \bar{\psi}, \lambda, \bar{\lambda}$; using also the fact that $P(P')^T = XY^T$, we obtain

$$\begin{aligned} \det[(\partial_X \partial_Y^T)_{IJ}] \det(XY^T)^s &= \epsilon(I, J) \det(XY^T)^{s-1} \det[(XY^T)_{I \subset J^c}] \\ &\quad \times \int \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\chi}^T \chi} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_m(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} \left(\prod_{a=k+1}^m \bar{\psi}_a \psi_a \right) \\ &\quad \times [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi) - (\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^{-s}. \end{aligned} \tag{5.75}$$

This formula expresses $\det[(\partial_X \partial_Y^T)_{IJ}] \det(XY^T)^s$ as the desired quantity $\det(XY^T)^{s-1} \epsilon(I, J) \det[(X \times Y^T)_{I \subset J^c}]$ multiplied by the purely combinatorial factor

$$\begin{aligned} p^{\text{tmrect}}(s, m, n, k) &\equiv \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_m(\lambda, \bar{\lambda}) \mathcal{D}_{n-m}(\chi, \bar{\chi}) \\ &\quad \times e^{\bar{\lambda}^T \lambda + \bar{\chi}^T \chi} \left(\prod_{a=k+1}^m \bar{\psi}_a \psi_a \right) \\ &\quad \times [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi) - (\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^{-s}, \end{aligned} \tag{5.76}$$

which we now proceed to calculate.

First note that the factor $\prod_{a=k+1}^m \bar{\psi}_a \psi_a$ forces the Taylor expansion of the square bracket to contain no variables $\psi_a, \bar{\psi}_a$ with $k + 1 \leq a \leq m$. We can therefore drop the factor $\prod_{a=k+1}^m \bar{\psi}_a \psi_a$, forget about the variables $\psi_a, \bar{\psi}_a$ with $k + 1 \leq a \leq m$, and consider $\psi, \bar{\psi}$ henceforth as vectors of length k . Let us also rename the vectors $\lambda, \bar{\lambda}$ as

$$(\lambda_1, \dots, \lambda_m) = (\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_{m-k}), \tag{5.77a}$$

$$(\bar{\lambda}_1, \dots, \bar{\lambda}_m) = (\bar{\lambda}_1, \dots, \bar{\lambda}_k, \bar{\mu}_1, \dots, \bar{\mu}_{m-k}). \tag{5.77b}$$

Then the quantity (5.76) is equivalent to

$$\begin{aligned}
 P^{\text{tmrect}}(s, m, n, k) &= \int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) \mathcal{D}_{m-k}(\mu, \bar{\mu}) \mathcal{D}_{n-m}(\chi, \bar{\chi}) \\
 &\quad \times e^{\bar{\lambda}^T \lambda + \bar{\mu}^T \mu + \bar{\chi}^T \chi} [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi) - (\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^{-s} \quad (5.78)
 \end{aligned}$$

where scalar products involving $\psi, \bar{\psi}, \lambda, \bar{\lambda}$ are understood as referring only to the first k variables. Integration over the variables $\mu, \bar{\mu}$ is trivial and produces just a factor 1. So we are left with

$$\begin{aligned}
 P^{\text{tmrect}}(s, m, n, k) &= \int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) \mathcal{D}_{n-m}(\chi, \bar{\chi}) \\
 &\quad \times e^{\bar{\lambda}^T \lambda + \bar{\chi}^T \chi} [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi) - (\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^{-s}. \quad (5.79)
 \end{aligned}$$

Note that $P^{\text{tmrect}}(s, m, n, k)$ depends on n and m only via the combination $n - m$.

The binomial expansion of the integrand in (5.79) yields

$$\begin{aligned}
 &[(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi) - (\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^{-s} \\
 &= \sum_{h=0}^{\infty} \binom{-s}{h} [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi)]^{-s-h} [-(\bar{\chi}^T \chi)(\bar{\psi}^T \psi)]^h. \quad (5.80)
 \end{aligned}$$

For each fixed value of h , integration over variables $\chi, \bar{\chi}$ gives

$$\int \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\chi}^T \chi} (-\bar{\chi}^T \chi)^h = (-1)^h \binom{n-m}{h} h!. \quad (5.81)$$

Now the factor $(\bar{\psi}^T \psi)^h$ will contain h pairs of variables, which can be chosen in $\binom{k}{h} h!$ ways (counting reorderings); then in the term $f(\psi, \bar{\psi}, \lambda, \bar{\lambda}) = [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi)]^{-s-h}$ we will have to use only the other $k - h$ pairs of variables for both $\psi, \bar{\psi}$ and $\lambda, \bar{\lambda}$, and with a reasoning as in Lemma 5.1 we can choose them to be the first $k - h$ indices, so that

$$\begin{aligned}
 &\int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} (\bar{\psi}^T \psi)^h f(\psi, \bar{\psi}, \lambda, \bar{\lambda}) \\
 &= \binom{k}{h} h! \int \mathcal{D}_{k-h}(\psi, \bar{\psi}) \mathcal{D}_{k-h}(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} f(\psi, \bar{\psi}, \lambda, \bar{\lambda}), \quad (5.82)
 \end{aligned}$$

and hence

$$\begin{aligned}
 P^{\text{tmrect}}(s, m, n, k) &= \sum_{h=0}^{\infty} (-1)^h \binom{n-m}{h} \binom{k}{h} (h!)^2 \binom{-s}{h} \\
 &\quad \times \int \mathcal{D}_{k-h}(\psi, \bar{\psi}) \mathcal{D}_{k-h}(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi)]^{-s-h}. \quad (5.83)
 \end{aligned}$$

Let us now make the change of variables $\psi_i \rightarrow \lambda_i, \lambda_i \rightarrow -\psi_i$ (whose Jacobian is cancelled by the reordering in the measure of integration); we have

$$\int \mathcal{D}_{k-h}(\psi, \bar{\psi}) \mathcal{D}_{k-h}(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} [(1 - \bar{\psi}^T \lambda)(1 + \bar{\lambda}^T \psi)]^{-s-h} \tag{5.84a}$$

$$= \int \mathcal{D}_{k-h}(\psi, \bar{\psi}) \mathcal{D}_{k-h}(\lambda, \bar{\lambda}) e^{-\bar{\lambda}^T \psi} [(1 + \bar{\psi}^T \psi)(1 + \bar{\lambda}^T \lambda)]^{-s-h} \tag{5.84b}$$

$$= \int \mathcal{D}_{k-h}(\psi, \bar{\psi}) \mathcal{D}_{k-h}(\lambda, \bar{\lambda}) [(1 + \bar{\psi}^T \psi)(1 + \bar{\lambda}^T \lambda)]^{-s-h} \tag{5.84c}$$

$$= \left[\int \mathcal{D}_{k-h}(\psi, \bar{\psi}) (1 + \bar{\psi}^T \psi)^{-s-h} \right]^2 \tag{5.84d}$$

$$= \left[\binom{-s-h}{k-h} (k-h)! \right]^2. \tag{5.84e}$$

Collecting all the factors, we have

$$P^{\text{tmrect}}(s, m, n, k) = \sum_{h=0}^{\infty} (-1)^h \binom{n-m}{h} \binom{k}{h} (h!)^2 \binom{-s}{h} \left[\binom{-s-h}{k-h} (k-h)! \right]^2 \tag{5.85a}$$

$$= \sum_{h=0}^{\infty} (-1)^h \binom{n-m}{h} (k!)^2 \binom{-s-h}{k-h} \binom{-s}{k} \tag{5.85b}$$

$$= \sum_{h=0}^{\infty} (-1)^h \binom{n-m}{h} (k!)^2 (-1)^{k-h} \binom{s+k-1}{k-h} (-1)^k \binom{s+k-1}{k} \tag{5.85c}$$

$$= (k!)^2 \binom{s+k-1}{k} \sum_{h=0}^{\infty} \binom{n-m}{h} \binom{s+k-1}{k-h} \tag{5.85d}$$

$$= (k!)^2 \binom{s+k-1}{k} \binom{s+k+n-m-1}{k} \tag{5.85e}$$

$$= \prod_{j=0}^{k-1} (s+j)(s+n-m+j), \tag{5.85f}$$

where the sum over h was performed using the Chu–Vandermonde convolution (Lemma B.2).

This proves Theorem 2.6 when X and Y are real $m \times n$ matrices of rank m lying in a sufficiently small neighborhood of \hat{I}_{mn} , and s is a positive integer. The general validity of the identity then follows from Proposition 2.18. \square

5.6. One-matrix rectangular symmetric Cayley identity

The proof of the one-matrix rectangular symmetric Cayley identity is extremely similar to that of the two-matrix identity, but is slightly more complicated because it involves a perturbation of rank 4 rather than rank 2 [compare (5.97) with (5.72)]. Luckily, the resulting 4×4 determinant turns out to be the square of a quantity involving only a 2×2 determinant [cf. (5.98b)/(5.99)]. Once again, we will need the full strength of Lemma 5.1 to handle the all-minors case.

Proof of Theorem 2.7. We begin once again by representing the differential operator as a Grassmann integral: exploiting Corollary B.6, we have

$$\det(\partial \partial^T) = \det \left(\begin{array}{c|c} 0_m & \partial \\ \hline -\partial^T & I_n \end{array} \right) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) e^{\bar{\eta}^T \eta + \psi^T \partial \bar{\eta} + \bar{\psi}^T \partial \eta}. \tag{5.86}$$

Here $\psi_i, \bar{\psi}_i$ ($1 \leq i \leq m$) and $\eta_j, \bar{\eta}_j$ ($1 \leq j \leq n$) are Grassmann variables, and we use the same conventions as in the preceding subsection. For a general minor $I, J \subseteq [m]$ with $|I| = |J| = k$, we have (writing $L = \{m + 1, \dots, m + n\}$)

$$\det[(\partial\partial^T)_{IJ}] = \det\left[\left(\begin{array}{c|c} 0_m & \partial \\ \hline -\partial^T & I_n \end{array}\right)_{I \cup L, J \cup L}\right] \tag{5.87a}$$

$$= \epsilon(I, J) \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\psi} \psi\right)_{I^c, J^c} e^{\bar{\eta}^T \eta + \psi^T \partial \bar{\eta} + \bar{\psi}^T \partial \eta}. \tag{5.87b}$$

Applying the translation formula (5.1) to the whole set of variables $\{x_{ij}\}$ produces

$$\begin{aligned} &\det[(\partial\partial^T)_{IJ}] f(X) \\ &= \epsilon(I, J) \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\psi} \psi\right)_{I^c, J^c} e^{\bar{\eta}^T \eta} f(X + \bar{\psi} \eta^T + \psi \bar{\eta}^T) \end{aligned} \tag{5.88}$$

for an arbitrary polynomial f . We shall use this formula in the case $f(X) = \det(XX^T)^s$ where s is a positive integer. It is convenient to introduce the shorthand

$$X^{\text{trans}} \equiv X + \bar{\psi} \eta^T + \psi \bar{\eta}^T \tag{5.89}$$

for the argument of f .

Suppose now that X is a real $m \times n$ matrix of rank m that is sufficiently close to the matrix \hat{I}_{mn} defined in (5.64). Then by Lemma B.19 we can find matrices $P \in GL(m)$ and $Q \in O(n)$ such that $X = P\hat{I}Q$ [we drop the subscripts mn on \hat{I}_{mn} to lighten the notation]. We have

$$\det(XX^T) = \det(P\hat{I}Q Q^T \hat{I}^T P^T) = \det(PP^T) = \det(P)^2 \tag{5.90}$$

and

$$X^{\text{trans}} \equiv X + \bar{\psi} \eta^T + \psi \bar{\eta}^T = P[\hat{I} + P^{-1}(\bar{\psi} \eta^T + \psi \bar{\eta}^T)Q^T]Q. \tag{5.91}$$

Let us now change variables from $(\psi, \bar{\psi}, \eta, \bar{\eta})$ to $(\psi', \bar{\psi}', \eta', \bar{\eta}') \equiv (P^{-1}\psi, P^{-1}\bar{\psi}, Q\eta, Q\bar{\eta})$, with Jacobian $(\det P)^{-2} = \det(XX^T)^{-1}$. In the new variables we have (dropping now the primes from the notation)

$$X^{\text{trans}} = P(\hat{I} + \bar{\psi} \eta^T + \psi \bar{\eta}^T)Q, \tag{5.92}$$

and the translated determinant is given by

$$\det[(X^{\text{trans}})(X^{\text{trans}})^T] = \det(XX^T) \det[(\hat{I} + \bar{\psi} \eta^T + \psi \bar{\eta}^T)(\hat{I}^T - \eta \bar{\psi}^T - \bar{\eta} \psi^T)], \tag{5.93}$$

so that

$$\begin{aligned} &\det[(\partial\partial^T)_{IJ}] \det(XX^T)^s \\ &= \epsilon(I, J) \det(XX^T)^{s-1} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod (P\bar{\psi})(P\psi)\right)_{I^c, J^c} \\ &\quad \times e^{\bar{\eta}^T \eta} \det[(\hat{I} + \bar{\psi} \eta^T + \psi \bar{\eta}^T)(\hat{I}^T - \eta \bar{\psi}^T - \bar{\eta} \psi^T)]^s. \end{aligned} \tag{5.94}$$

Let us now split the vectors η and $\bar{\eta}$ as

$$(\eta_1, \dots, \eta_n) = (\lambda_1, \dots, \lambda_m, \chi_1, \dots, \chi_{n-m}), \tag{5.95a}$$

$$(\bar{\eta}_1, \dots, \bar{\eta}_n) = (\bar{\lambda}_1, \dots, \bar{\lambda}_m, \bar{\chi}_1, \dots, \bar{\chi}_{n-m}) \tag{5.95b}$$

so that

$$\begin{aligned} &(\hat{I} + \bar{\psi}\eta^T + \psi\bar{\eta}^T)(\hat{I}^T - \eta\bar{\psi}^T - \bar{\eta}\psi^T) \\ &= I_m + \bar{\psi}\lambda^T + \psi\bar{\lambda}^T - \lambda\bar{\psi}^T - \bar{\lambda}\psi^T + c\bar{\psi}\psi^T - c\psi\bar{\psi}^T \end{aligned} \tag{5.96}$$

with $c = \bar{\lambda}^T\lambda + \bar{\chi}^T\chi$. This matrix has the form of a low-rank perturbation $I_m + \sum_{\alpha=1}^4 u_\alpha v_\alpha^T$, with vectors $\{u_\alpha\}, \{v_\alpha\}$ given by:

α	u_α	v_α	(5.97)
1	$\bar{\psi}$	$\lambda + c\psi$	
2	ψ	$\bar{\lambda} - c\bar{\psi}$	
3	$\bar{\lambda}$	$-\psi$	
4	λ	$-\bar{\psi}$	

By Lemma B.11 we can write the needed determinant as the determinant of a 4×4 matrix; after a few row and column manipulations we can write

$$\begin{aligned} &\det[(\hat{I} + \bar{\psi}\eta^T + \psi\bar{\eta}^T)(\hat{I}^T - \eta\bar{\psi}^T - \bar{\eta}\psi^T)] \\ &= \det^{-1} \left(\begin{array}{cc|cc} & A & 0 & \bar{\chi}^T\chi \\ & & -\bar{\chi}^T\chi & 0 \\ \hline 0 & -\bar{\psi}^T\psi & & \\ \bar{\psi}^T\psi & 0 & & A^T \end{array} \right) \end{aligned} \tag{5.98a}$$

$$= [\det A - (\bar{\psi}^T\psi)(\bar{\chi}^T\chi)]^{-2} \tag{5.98b}$$

where

$$A = \begin{pmatrix} 1 + \lambda^T\bar{\psi} & \lambda^T\psi \\ \bar{\lambda}^T\bar{\psi} & 1 + \bar{\lambda}^T\psi \end{pmatrix} \tag{5.99}$$

and hence

$$\det A = 1 + \lambda^T\bar{\psi} + \bar{\lambda}^T\psi + (\lambda^T\bar{\psi})(\bar{\lambda}^T\psi) + (\bar{\lambda}^T\bar{\psi})(\psi^T\lambda). \tag{5.100}$$

We therefore have

$$\begin{aligned} \det[(\partial\partial^T)_{IJ}] \det(XX^T)^s &= \epsilon(I, J) \det(XX^T)^{s-1} \int \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\chi}^T\chi} \\ &\times \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_m(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T\lambda} \left(\prod (P\bar{\psi})(P\psi) \right)_{I^c, J^c} \\ &\times [\det A - (\bar{\psi}^T\psi)(\bar{\chi}^T\chi)]^{-2s}. \end{aligned} \tag{5.101}$$

Note in particular that $e^{\bar{\lambda}^T \lambda}$ and $\det A - (\bar{\psi}^T \psi)(\bar{\chi}^T \chi)$ depend on $\psi, \bar{\psi}, \lambda, \bar{\lambda}$ only via scalar products. This allows us to apply Lemma 5.1 to the integral over $\psi, \bar{\psi}, \lambda, \bar{\lambda}$; using also the fact that $PP^T = XX^T$, we obtain

$$\begin{aligned} \det[(\partial\bar{\partial}^T)_{IJ}] \det(XX^T)^s &= \epsilon(I, J) \det(XX^T)^{s-1} \det[(XX^T)_{I^c J^c}] \\ &\times \int \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\chi}^T \chi} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_m(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} \\ &\times \left(\prod_{a=k+1}^m \bar{\psi}_a \psi_a \right) [\det A - (\bar{\psi}^T \psi)(\bar{\chi}^T \chi)]^{-2s}. \end{aligned} \tag{5.102}$$

This formula expresses $\det[(\partial\bar{\partial}^T)_{IJ}] \det(XX^T)^s$ as the desired quantity $\det(XX^T)^{s-1} \epsilon(I, J) \det[(X \times X^T)_{I^c J^c}]$ multiplied by the purely combinatorial factor

$$\begin{aligned} P^{\text{symrect}}(s, m, n, k) &\equiv \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_m(\lambda, \bar{\lambda}) \mathcal{D}_{n-m}(\chi, \bar{\chi}) \\ &\times e^{\bar{\lambda}^T \lambda + \bar{\chi}^T \chi} \left(\prod_{a=k+1}^m \bar{\psi}_a \psi_a \right) [\det A - (\bar{\psi}^T \psi)(\bar{\chi}^T \chi)]^{-2s}, \end{aligned} \tag{5.103}$$

which we now proceed to calculate.

First note that the factor $\prod_{a=k+1}^m \bar{\psi}_a \psi_a$ forces the Taylor expansion of $[\det A - (\bar{\psi}^T \psi)(\bar{\chi}^T \chi)]^{-2s}$ to contain no variables $\psi_a, \bar{\psi}_a$ with $k+1 \leq a \leq m$. We can therefore drop the factor $\prod_{a=k+1}^m \bar{\psi}_a \psi_a$, forget about the variables $\psi_a, \bar{\psi}_a$ with $k+1 \leq a \leq m$, and consider $\psi, \bar{\psi}$ henceforth as vectors of length k . Let us also rename the vectors $\lambda, \bar{\lambda}$ as

$$(\lambda_1, \dots, \lambda_m) = (\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_{m-k}), \tag{5.104a}$$

$$(\bar{\lambda}_1, \dots, \bar{\lambda}_m) = (\bar{\lambda}_1, \dots, \bar{\lambda}_k, \bar{\mu}_1, \dots, \bar{\mu}_{m-k}). \tag{5.104b}$$

Then the quantity (5.103) is equivalent to

$$\begin{aligned} P^{\text{symrect}}(s, m, n, k) &= \int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) \mathcal{D}_{m-k}(\mu, \bar{\mu}) \mathcal{D}_{n-m}(\chi, \bar{\chi}) \\ &\times e^{\bar{\lambda}^T \lambda + \bar{\mu}^T \mu + \bar{\chi}^T \chi} [\det A - (\bar{\psi}^T \psi)(\bar{\chi}^T \chi)]^{-2s} \end{aligned} \tag{5.105}$$

where $\det A$ has the same expression as in (5.100) but scalar products involving $\psi, \bar{\psi}, \lambda, \bar{\lambda}$ are understood as referring only to the first k variables. Integration over the variables $\mu, \bar{\mu}$ is trivial and produces just a factor 1. So we are left with

$$\begin{aligned} P^{\text{symrect}}(s, m, n, k) &= \int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\lambda}^T \lambda + \bar{\chi}^T \chi} [\det A - (\bar{\psi}^T \psi)(\bar{\chi}^T \chi)]^{-2s}. \end{aligned} \tag{5.106}$$

Note that $P^{\text{symrect}}(s, m, n, k)$ depends on n and m only via the combination $n - m$.

The multinomial expansion of the integrand in (5.106) (except for $e^{\bar{\chi}^T \chi}$) is

$$\sum_{\substack{r \geq 0 \\ t_1, \dots, t_5 \geq 0}} \binom{-2s}{t_1 + \dots + t_5} \binom{t_1 + \dots + t_5}{t_1, \dots, t_5} e^{\bar{\chi}^T \chi} \frac{(-1)^{t_5}}{r!} \times (\bar{\lambda}^T \lambda)^r (\bar{\chi}^T \chi)^{t_5} (\lambda^T \bar{\psi})^{t_1+t_3} (\bar{\lambda}^T \psi)^{t_2+t_3} [(\bar{\lambda}^T \bar{\psi})(\psi^T \lambda)]^{t_4} (\bar{\psi}^T \psi)^{t_5}. \tag{5.107}$$

The integration

$$\int \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\chi}^T \chi} (\bar{\chi}^T \chi)^\ell = \frac{(n-m)!}{(n-m-\ell)!} \tag{5.108}$$

is trivial. So we are left with two integrations over complex Grassmann vectors of length k . This is performed through a lemma that we shall prove at the end of this subsection:

Lemma 5.5. For all integers $a, a', b, b', c, c' \geq 0$, we have

$$\int \mathcal{D}_n(\psi, \bar{\psi}) \mathcal{D}_n(\lambda, \bar{\lambda}) (\bar{\lambda}^T \lambda)^a (\bar{\psi}^T \psi)^{a'} (\bar{\lambda}^T \psi)^b (\lambda^T \bar{\psi})^{b'} (\bar{\lambda}^T \bar{\psi})^c (\psi^T \lambda)^{c'} = \delta_{aa'} \delta_{bb'} \delta_{cc'} \delta_{a+b+c, n} \binom{n}{a, b, c} (a! b! c!)^2. \tag{5.109}$$

Combining formula (5.107) (after integration of $\bar{\chi}, \chi$) and the statement of the lemma, we are left with three independent summations; we choose the summation indices to be t_1, t_3 and t_5 , which we relabel as h, l and j , respectively. The remaining indices are given by $r = j, t_2 = h, t_4 = m - j - h - l$. The resulting expression is

$$p^{\text{symrect}}(s, m, n, k) = (k!)^2 \binom{-2s}{k} \sum_{\substack{j, h, l \geq 0 \\ j+h+l \leq k}} (-1)^j \binom{-2s-k}{h} \binom{n-m}{j} \binom{h+l}{h}. \tag{5.110}$$

Renaming $-2s - k \equiv a$ and $n - m \equiv b$ highlights the fact that there is no direct dependence of the summands on k . From Lemma B.4 we have

$$\sum_{\substack{j, h, l \geq 0 \\ j+h+l \leq k}} (-1)^j \binom{a}{h} \binom{b}{j} \binom{h+l}{h} = \binom{a-b+k+1}{k}, \tag{5.111}$$

which finally gives

$$p^{\text{symrect}}(s, m, n, k) = (k!)^2 \binom{-2s}{k} \binom{-(2s+n-m-1)}{k} \tag{5.112a}$$

$$= \prod_{j=0}^{k-1} (2s+j)(2s+n-m-1+j). \tag{5.112b}$$

This proves (2.19) when X is a real $m \times n$ matrix of rank m lying in a sufficiently small neighborhood of \hat{I}_{mn} , and s is a positive integer; the general validity of the identity then follows from Proposition 2.18. \square

Proof of Lemma 5.5. Let us rewrite (5.109) by forming an exponential generating function: that is, we multiply both sides by $\alpha^a (\alpha')^a \beta^b (\beta')^b \gamma^c (\gamma')^c / [a! a! b! b! c! c!]$ and sum over $a, a', b, b', c, c' \geq 0$. So (5.109) is equivalent to

$$\int \mathcal{D}_n(\psi, \bar{\psi}) \mathcal{D}_n(\lambda, \bar{\lambda}) e^{\alpha \bar{\lambda}^T \lambda + \alpha' \bar{\psi}^T \psi + \beta \bar{\lambda}^T \psi + \beta' \lambda^T \bar{\psi} + \gamma \bar{\lambda}^T \bar{\psi} + \gamma' \psi^T \lambda} = (\alpha \alpha' + \beta \beta' + \gamma \gamma')^n, \tag{5.113}$$

and this is the formula that we shall prove. Note first that the measure can be rewritten as

$$\mathcal{D}_n(\psi, \bar{\psi}) \mathcal{D}_n(\lambda, \bar{\lambda}) = d\psi_n d\bar{\psi}_n d\lambda_n d\bar{\lambda}_n \cdots d\psi_1 d\bar{\psi}_1 d\lambda_1 d\bar{\lambda}_1 \tag{5.114}$$

with no minus signs. So let us assemble $\psi, \bar{\psi}, \lambda, \bar{\lambda}$ into a single Grassmann vector θ of length $4n$, with $\psi_k = \theta_{4k}, \bar{\psi}_k = \theta_{4k-1}, \lambda_k = \theta_{4k-2}$ and $\bar{\lambda}_k = \theta_{4k-3}$; then the measure becomes $\mathcal{D}_{4n}(\theta)$. In the exponential we have an antisymmetric bilinear form $\frac{1}{2} \theta Q \theta$, in which Q is block-diagonal with n identical 4×4 blocks that we call Q' :

$$Q' = \begin{pmatrix} 0 & \alpha & \gamma & \beta \\ -\alpha & 0 & \beta' & -\gamma' \\ -\gamma & -\beta' & 0 & \alpha' \\ -\beta & \gamma' & -\alpha' & 0 \end{pmatrix}. \tag{5.115}$$

So the result of the integration is $\text{pf } Q = (\text{pf } Q')^n$. And indeed, $\text{pf } Q' = \alpha \alpha' + \beta \beta' + \gamma \gamma'$, as was to be proven. \square

5.7. One-matrix rectangular antisymmetric Cayley identity

Let us now prove the one-matrix rectangular antisymmetric Cayley identity (Theorem 2.8). Again, this proof is extremely similar to that of the two-matrix and the symmetric one-matrix rectangular identities. As always with pfaffians, we will deal only with principal minors, for which we will need the full strength of Lemma 5.3.

Proof of Theorem 2.8. We begin once again by representing the differential operator as a Grassmann integral: exploiting Corollary B.9, we have

$$\text{pf}(\partial J \partial^T) = (-1)^m \text{pf} \left(\begin{array}{c|c} 0_{2m} & \partial \\ \hline -\partial^T & J_{2n} \end{array} \right) \tag{5.116a}$$

$$= (-1)^m \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2n}(\eta) e^{\frac{1}{2} \eta^T J \eta + \psi^T \partial \eta}, \tag{5.116b}$$

where ψ_i ($1 \leq i \leq 2m$) and η_j ($1 \leq j \leq 2n$) are “real” Grassmann variables, and we use the same conventions as in the preceding subsection. For a general even-dimensional principal minor $I \subseteq [2m]$ with $|I| = 2k$, we similarly have (writing $L = \{m+1, \dots, m+n\}$)

$$\text{pf}[(\partial J \partial^T)_{II}] = (-1)^k \text{pf} \left[\left(\begin{array}{c|c} 0_{2m} & \partial \\ \hline -\partial^T & J_{2n} \end{array} \right)_{I \cup L, I \cup L} \right] \tag{5.117a}$$

$$= (-1)^k \epsilon(I) \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2n}(\eta) \left(\prod \psi \right)_{I^c} e^{\frac{1}{2} \eta^T J \eta + \psi^T \partial \eta}. \tag{5.117b}$$

Applying the translation formula (5.1) to the whole set of variables $\{x_{ij}\}$ produces

$$\text{pf}[(\partial J \partial^T)_{II}] f(X) = (-1)^k \epsilon(I) \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2n}(\eta) \left(\prod \psi \right)_{I^c} e^{\frac{1}{2} \eta^T J \eta} f(X + \psi \eta^T) \tag{5.118}$$

for an arbitrary polynomial f . We shall use this formula in the case $f(X) = \text{pf}(XJX^T)^s$ where s is a positive integer. It is convenient to introduce the shorthand

$$X^{\text{trans}} \equiv X + \psi \eta^T \tag{5.119}$$

for the argument of f .

Suppose now that X is a real $2m \times 2n$ matrix of rank $2m$ that is sufficiently close to the matrix $\hat{I}_{2m,2n}$ defined in (5.64). Then by Lemma B.20 we can find matrices $P \in GL(2m)$ and $Q \in Sp(2n)$ such that $X = P\hat{I}_{2m,2n}Q$. We recall the defining property of $Sp(2n)$, namely that $QJ_{2n}Q^T = J_{2n}$. We have

$$XJ_{2n}X^T = P\hat{I}_{2m,2n}QJ_{2n}Q^T\hat{I}_{2m,2n}^T P^T = PJ_{2m}P^T, \tag{5.120}$$

so that

$$\text{pf}(XJ_{2n}X^T) = \text{pf}(PJ_{2m}P^T) = \det P \tag{5.121}$$

and

$$X^{\text{trans}} \equiv X + \psi \eta^T = P[\hat{I} + P^{-1}\psi \eta^T Q^{-1}]Q \tag{5.122}$$

(we drop the subscripts on $\hat{I}_{2m,2n}$ to lighten the notation). Let us now change variables from (ψ, η) to $(\psi', \eta') \equiv (P^{-1}\psi, Q^{-T}\eta)$, with Jacobian $(\det P)^{-1} = \text{pf}(XJX^T)^{-1}$. In the new variables we have (dropping now the primes from the notation)

$$X^{\text{trans}} = P(\hat{I} + \psi \eta^T)Q, \tag{5.123}$$

and the translated pfaffian is given by

$$\begin{aligned} \text{pf}[(X^{\text{trans}})J(X^{\text{trans}})^T] &= \text{pf}[P(\hat{I} + \psi \eta^T)J(\hat{I}^T - \eta \psi^T)P^T] \\ &= \det(P) \text{pf}[(\hat{I} + \psi \eta^T)J(\hat{I}^T - \eta \psi^T)] \\ &= \text{pf}(XJX^T) \text{pf}[(\hat{I} + \psi \eta^T)J(\hat{I}^T - \eta \psi^T)], \end{aligned} \tag{5.124}$$

so that

$$\begin{aligned} &\text{pf}[(\partial J \partial^T)_{II}] \text{pf}(XJX^T)^s \\ &= \text{pf}(XJX^T)^{s-1} (-1)^k \epsilon(I) \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2n}(\eta) \left(\prod (P\psi) \right)_{Ic} \\ &\quad \times e^{\frac{1}{2}\eta^T J \eta} \text{pf}[(\hat{I} + \psi \eta^T)J(\hat{I}^T - \eta \psi^T)]^s. \end{aligned} \tag{5.125}$$

Let us now split the vector η as

$$(\eta_1, \dots, \eta_{2n}) = (\lambda_1, \dots, \lambda_{2m}, \chi_1, \dots, \chi_{2(n-m)}), \tag{5.126}$$

so that

$$(\hat{I} + \psi \eta^T)J_{2n}(\hat{I}^T - \eta \psi^T) = J_{2m} - \psi \lambda^T J_{2m}^T - J_{2m} \lambda \psi^T - c \psi \psi^T \tag{5.127}$$

with $c = \eta^T J_{2n} \eta = \lambda^T J_{2m} \lambda + \chi^T J_{2(n-m)} \chi$. The matrix (5.127) is manifestly antisymmetric, and has the form of a low-rank perturbation of J_{2m} . Since $\text{pf } J_{2m} = 1$ and the perturbation is purely Grassmannian, we can write

$$\text{pf}(J - \psi \lambda^T J^T - J \lambda \psi^T - c \psi \psi^T) = \det^{1/2}(I + J \psi \lambda^T J^T - \lambda \psi^T + c J \psi \psi^T). \tag{5.128}$$

Here the argument of \det is a matrix of the form $I_{2m} + \sum_{\alpha=1}^2 u_\alpha v_\alpha^T$, with vectors $\{u_\alpha\}, \{v_\alpha\}$ given by:

$$\begin{array}{c|cc} \alpha & u_\alpha & v_\alpha \\ \hline 1 & J\psi & J\lambda + c\psi \\ 2 & -\lambda & \psi \end{array} \tag{5.129}$$

By Lemma B.11 we can write the needed determinant as the determinant of a 2×2 matrix, which, after a slight row manipulation, can be written as

$$\det(I + J \psi \lambda^T J^T - \lambda \psi^T + c J \psi \psi^T) = \det^{-1} \begin{pmatrix} 1 - \psi^T \lambda & -\chi^T J \chi \\ \psi^T J \psi & 1 - \psi^T \lambda \end{pmatrix}. \tag{5.130}$$

We therefore have

$$\begin{aligned} \text{pf}[(\partial J \partial^T)_{II}] \text{pf}(X J X^T)^s &= \text{pf}(X J X^T)^{s-1} (-1)^k \epsilon(I) \int \mathcal{D}_{2(n-m)}(\chi) e^{\frac{1}{2} \chi^T J \chi} \\ &\quad \times \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2m}(\lambda) e^{\frac{1}{2} \lambda^T J \lambda} \left(\prod (P \psi) \right)_{I^c} \\ &\quad \times [(1 - \psi^T \lambda)^2 + (\psi^T J \psi)(\chi^T J \chi)]^{-s/2}. \end{aligned} \tag{5.131}$$

In this expression we have both an ordinary scalar product ($\psi^T \lambda$) and symplectic scalar products ($\lambda^T J \lambda$, $\psi^T J \psi$ and $\chi^T J \chi$). By a further change of variables $\lambda \rightarrow \lambda' = -J\lambda$, we can reduce to symplectic products only (the Jacobian is 1). Dropping the primes, we have (also using $J^T J J = J$)

$$\begin{aligned} \text{pf}[(\partial J \partial^T)_{II}] \text{pf}(X J X^T)^s &= \text{pf}(X J X^T)^{s-1} (-1)^k \epsilon(I) \int \mathcal{D}_{2(n-m)}(\chi) e^{\frac{1}{2} \chi^T J \chi} \\ &\quad \times \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2m}(\lambda) e^{\frac{1}{2} \lambda^T J \lambda} \left(\prod (P \psi) \right)_{I^c} \\ &\quad \times [(1 - \psi^T J \lambda)^2 + (\psi^T J \psi)(\chi^T J \chi)]^{-s/2}. \end{aligned} \tag{5.132}$$

Now the integral on the ψ and λ fields is of the form described in Lemma 5.3. Applying this lemma, and making use of (5.120), we have

$$\begin{aligned} \text{pf}[(\partial J \partial^T)_{II}] \text{pf}(X J X^T)^s &= \epsilon(I) \text{pf}(X J X^T)^{s-1} \text{pf}[(X J X^T)_{I^c I^c}] \\ &\quad \times (-1)^k \int \mathcal{D}_{2(n-m)}(\chi) e^{\frac{1}{2} \chi^T J \chi} \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2m}(\lambda) e^{\frac{1}{2} \lambda^T J \lambda} \\ &\quad \times \left(\prod \psi \right)_{\{2k+1, \dots, 2m\}} [(1 - \psi^T J \lambda)^2 + (\psi^T J \psi)(\chi^T J \chi)]^{-s/2}. \end{aligned} \tag{5.133}$$

This formula expresses $\text{pf}[(\partial J \partial^T)_{II}] \text{pf}(X J X^T)^s$ as the desired quantity $\text{pf}(X J X^T)^{s-1} \epsilon(I) \text{pf}[(X J X^T)_{I^c I^c}]$ multiplied by the purely combinatorial factor

$$p^{\text{asrect}}(s, m, n, k) \equiv (-1)^k \int \mathcal{D}_{2(n-m)}(\chi) e^{\frac{1}{2}\chi^T J \chi} \int \mathcal{D}_{2m}(\psi) \mathcal{D}_{2m}(\lambda) e^{\frac{1}{2}\lambda^T J \lambda} \times \left(\prod_{\{2k+1, \dots, 2m\}} \psi \right) [(1 - \psi^T J \lambda)^2 + (\psi^T J \psi)(\chi^T J \chi)]^{-s/2}, \quad (5.134)$$

which we now proceed to calculate.

Note, first of all, that the factor $\prod_{a=2k+1}^{2m} \psi_a$ forces the Taylor expansion of $[(1 - \psi^T J \lambda)^2 + (\psi^T J \psi)(\chi^T J \chi)]^{-s/2}$ to contain no variables ψ_a with $2k + 1 \leq a \leq 2m$, and thus also no variables λ_a with $2k + 1 \leq a \leq 2m$. The latter must all come from the trivial diagonal exponential $e^{\frac{1}{2}\lambda^T J \lambda}$, and their integration can then be performed easily (it gives 1). We can therefore drop the factor $\prod_{a=2k+1}^{2m} \psi_a$, forget about the variables ψ_a and λ_a with $2k + 1 \leq a \leq 2m$, and consider ψ and λ henceforth as vectors of length $2k$. We have

$$p^{\text{asrect}}(s, m, n, k) \equiv (-1)^k \int \mathcal{D}_{2(n-m)}(\chi) e^{\frac{1}{2}\chi^T J \chi} \int \mathcal{D}_{2k}(\psi) \mathcal{D}_{2k}(\lambda) \times e^{\frac{1}{2}\lambda^T J \lambda} [(1 - \psi^T J \lambda)^2 + (\psi^T J \psi)(\chi^T J \chi)]^{-s/2}. \quad (5.135)$$

We see that $p^{\text{asrect}}(s, m, n, k)$ depends on n and m only via the combination $n - m$.

In the expression (5.135) we have symplectic scalar products of “real” Grassmann fields. However, our summation lemmas (such as Lemma 5.5) have heretofore been developed only for ordinary scalar products of “complex” Grassmann fields. However, through a relabeling of the fields, we can rewrite (5.135) in terms of ordinary scalar products of “complex” fields of half the dimensions. More precisely, we relabel $(\psi_1, \dots, \psi_{2k}) \rightarrow (\bar{\psi}_1, \psi_1, \dots, \bar{\psi}_k, \psi_k)$, $(\lambda_1, \dots, \lambda_{2k}) \rightarrow (\bar{\lambda}_1, \lambda_1, \dots, \bar{\lambda}_k, \lambda_k)$ and $(\chi_1, \dots, \chi_{2(n-m)}) \rightarrow (\bar{\chi}_1, \chi_1, \dots, \bar{\chi}_{n-m}, \chi_{n-m})$. No signs arise in the measure of integration, and we have the correspondences

$$\frac{1}{2} \psi^T J \psi \rightarrow \bar{\psi}^T \psi, \quad (5.136a)$$

$$\frac{1}{2} \lambda^T J \lambda \rightarrow \bar{\lambda}^T \lambda, \quad (5.136b)$$

$$\psi^T J \lambda \rightarrow \bar{\psi}^T \lambda + \bar{\lambda}^T \psi, \quad (5.136c)$$

$$\frac{1}{2} \chi^T J \chi \rightarrow \bar{\chi}^T \chi, \quad (5.136d)$$

which rewrites (5.135) as

$$p^{\text{asrect}}(s, m, n, k) = (-1)^k \int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) \mathcal{D}_{n-m}(\chi, \bar{\chi}) \times e^{\bar{\lambda}^T \lambda + \bar{\chi}^T \chi} [(1 - \bar{\psi}^T \lambda - \bar{\lambda}^T \psi)^2 + 4(\bar{\psi}^T \psi)(\bar{\chi}^T \chi)]^{-s/2}. \quad (5.137)$$

We begin by getting rid of variables χ and $\bar{\chi}$, writing

$$\int \mathcal{D}_{n-m}(\chi, \bar{\chi}) e^{\bar{\chi}^T \chi} [(1 - \bar{\psi}^T \lambda - \bar{\lambda}^T \psi)^2 + 4(\bar{\psi}^T \psi)(\bar{\chi}^T \chi)]^{-s/2} = \sum_{a \geq 0} \binom{-\frac{s}{2}}{a} (4\bar{\psi}^T \psi)^a (1 - \bar{\psi}^T \lambda - \bar{\lambda}^T \psi)^{-s-2a} \binom{n-m}{a} a!. \quad (5.138)$$

So we have

$$\begin{aligned}
 P^{\text{asrect}}(s, m, n, k) &= (-1)^k \int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) e^{\bar{\lambda}^T \lambda} \\
 &\quad \times \sum_{a \geq 0} \binom{-\frac{s}{2}}{a} (4 \bar{\psi}^T \psi)^a (1 - \bar{\psi}^T \lambda - \bar{\lambda}^T \psi)^{-s-2a} \binom{n-m}{a} a!. \tag{5.139}
 \end{aligned}$$

Next we expand fully the integrand in (5.139), yielding

$$\begin{aligned}
 P^{\text{asrect}}(s, m, n, k) &= \sum_{\substack{a, a' \geq 0 \\ b, b' \geq 0}} 4^a (-1)^{k-b} \binom{-\frac{s}{2}}{a} \binom{n-m}{a} a! \binom{-s-2a}{b+b'} \binom{b+b'}{b} \frac{1}{a!} \\
 &\quad \times \int \mathcal{D}_k(\psi, \bar{\psi}) \mathcal{D}_k(\lambda, \bar{\lambda}) (\bar{\psi}^T \psi)^a (\bar{\lambda}^T \lambda)^{a'} (\bar{\psi}^T \lambda)^b (\psi^T \bar{\lambda})^{b'}. \tag{5.140}
 \end{aligned}$$

The fermionic integration here is a special case of Lemma 5.5, which thus gives

$$\begin{aligned}
 P^{\text{asrect}}(s, m, n, k) &= \sum_{\substack{a, a' \geq 0 \\ b, b' \geq 0}} 4^a (-1)^{k-b} \binom{-\frac{s}{2}}{a} \binom{n-m}{a} a! \binom{-s-2a}{b+b'} \binom{b+b'}{b} \frac{1}{a!} \\
 &\quad \times \delta_{aa'} \delta_{bb'} \delta_{a+b, k} \binom{k}{a} (a! b!)^2 \\
 &= \sum_{a=0}^k (-4)^a \binom{-\frac{s}{2}}{a}^a (-s-2a)^{\frac{2k-2a}{2}} \binom{n-m}{a} \binom{k}{a} a! \tag{5.141}
 \end{aligned}$$

where $x^k = x(x-1) \cdots (x-k+1)$. Notice now that

$$\begin{aligned}
 &\binom{-\frac{s}{2}}{a}^a (-s-2a)^{\frac{2k-2a}{2}} \\
 &= \left(-\frac{1}{2}\right)^a s(s+2) \cdots (s+2k-2)(s+2a+1)(s+2a+3) \cdots (s+2k-1). \tag{5.142}
 \end{aligned}$$

The factors of the form $s+2j$, namely $s(s+2) \cdots (s+2k-2)$, are independent of the summation variable a . What remains is

$$\begin{aligned}
 &\sum_{a=0}^k 2^a (s+2a+1)(s+2a+3) \cdots (s+2k-1) \binom{n-m}{a} \binom{k}{a} a! \\
 &= \sum_{a=0}^k 2^k \binom{\frac{s-1}{2} + k}{k-a} (k-a)! \binom{n-m}{a} \binom{k}{a} a! \\
 &= 2^k k! \sum_{a=0}^k \binom{\frac{s-1}{2} + k}{k-a} \binom{n-m}{a} \\
 &= 2^k k! \binom{\frac{s-1}{2} + n - m + k}{k}
 \end{aligned}$$

$$= \prod_{j=0}^{k-1} (s + 1 + 2n - 2m + 2j) \tag{5.143}$$

where the sum over a was a Chu–Vandermonde convolution (Lemma B.2). Restoring the factors of the form $s + 2j$, we conclude that

$$p^{\text{asrect}}(s, m, n, k) = \prod_{j=0}^{k-1} (s + 2j)(s + 1 + 2n - 2m + 2j). \tag{5.144}$$

This proves (2.22) when X is a real $2m \times 2n$ matrix of rank $2m$ lying in a sufficiently small neighborhood of $\hat{I}_{2m, 2n}$, and s is a positive integer; the general validity of the identity then follows from Proposition 2.18. \square

5.8. A lemma on Grassmann integrals of scalar products

In the preceding proofs we have frequently had to evaluate Grassmann integrals over one or more sets of Grassmann variables, in which the integrand depends only on scalar products among those sets. A simple version of such a formula occurred already in (5.9), while more complicated versions arose in (5.79) ff. and (5.109). [Also, a version for real fermions arose in (5.59).] So far we have simply treated such integrals “by hand”. But since a much more complicated such integral will arise in the proof of the multi-matrix rectangular Cayley identity (Theorem 2.9) in the next subsection, it is worth stating now a general lemma that allows us to systematize such calculations. Indeed, we think that this lemma (Proposition 5.6 below) is of some independent interest.

Let us start by rephrasing the simplest case (5.9) in a suggestive way. Let n be a positive integer, and introduce Grassmann variables $\psi_i, \bar{\psi}_i$ ($1 \leq i \leq n$). Let $f(x) = \sum_{k=0}^{\infty} a_k x^k$ be a formal power series in one indeterminate. We then have by (5.9)

$$\int \mathcal{D}_n(\psi, \bar{\psi}) f(\bar{\psi}^T \psi) = n! a_n. \tag{5.145}$$

On the other hand, we also have trivially

$$\left. \frac{d^n}{dx^n} f(x) \right|_{x=0} = n! a_n \tag{5.146}$$

and

$$\left. f\left(\frac{d}{dx}\right) x^n \right|_{x=0} = n! a_n. \tag{5.147}$$

Hence

$$\int \mathcal{D}_n(\psi, \bar{\psi}) f(\bar{\psi}^T \psi) = \left. \frac{d^n}{dx^n} f(x) \right|_{x=0} = \left. f\left(\frac{d}{dx}\right) x^n \right|_{x=0}. \tag{5.148}$$

Surprisingly enough, a similar formula exists for Grassmann integrals involving multiple sets of fermionic variables:

Proposition 5.6. Let ℓ and n be positive integers, and introduce Grassmann variables $\psi_i^\alpha, \bar{\psi}_i^\alpha$ ($1 \leq \alpha \leq \ell, 1 \leq i \leq n$). Let $f(X)$ be a formal power series in commuting indeterminates $X = (x_{\alpha\beta})_{\alpha,\beta=1}^\ell$, and write $\partial = (\partial/\partial x_{\alpha\beta})$. We then have

$$\int \mathcal{D}_n(\psi^1, \bar{\psi}^1) \cdots \mathcal{D}_n(\psi^\ell, \bar{\psi}^\ell) f(\{\bar{\psi}^{\alpha T} \psi^\beta\}) = \det(\partial)^n f(X)|_{X=0} \tag{5.149a}$$

$$= f(\partial)(\det X)^n|_{X=0}. \tag{5.149b}$$

Proof. We write

$$f(X) = \sum_{N \in \mathbb{N}^{\ell \times \ell}} f_N \prod_{\alpha,\beta=1}^\ell x_{\alpha\beta}^{n_{\alpha\beta}} \tag{5.150}$$

where the sum runs over matrices $N = (n_{\alpha\beta})_{\alpha,\beta=1}^\ell$ of nonnegative integers. Note first that it suffices to prove (5.149) for *polynomials* f , since all three expressions in (5.149) have the property that only finitely many coefficients N contribute (namely, those matrices N with all row and column sums equal to n). So it suffices to prove (5.149) for all monomials $X^N := \prod_{\alpha,\beta=1}^\ell x_{\alpha\beta}^{n_{\alpha\beta}}$. But then it suffices to prove (5.149) for the exponential generating function

$$\phi_\Omega(X) = \exp \operatorname{tr}(\Omega^T X) = \sum_{N \in \mathbb{N}^{\ell \times \ell}} \prod_{\alpha,\beta=1}^\ell \frac{(\omega_{\alpha\beta} x_{\alpha\beta})^{n_{\alpha\beta}}}{n_{\alpha\beta}!} \tag{5.151}$$

where $\Omega = (\omega_{\alpha\beta})_{\alpha,\beta=1}^\ell$ are indeterminates, since the value at the monomial X^N can be obtained by extracting the coefficient $[\Omega^N]$. Our goal is to prove that all the three expressions in (5.149), specialized to $f(X) = \phi_\Omega(X)$, equal $(\det \Omega)^n$.

For the Grassmann-integral expression we have

$$\begin{aligned} & \int \mathcal{D}_n(\psi^1, \bar{\psi}^1) \cdots \mathcal{D}_n(\psi^\ell, \bar{\psi}^\ell) \phi_\Omega(\{\bar{\psi}^{\alpha T} \psi^\beta\}) \\ &= \int \mathcal{D}_n(\psi^1, \bar{\psi}^1) \cdots \mathcal{D}_n(\psi^\ell, \bar{\psi}^\ell) \exp\left(\sum_{i=1}^n \sum_{\alpha,\beta=1}^\ell \bar{\psi}_i^\alpha \omega_{\alpha\beta} \psi_i^\beta\right) \end{aligned} \tag{5.152a}$$

$$= (\det \Omega)^n. \tag{5.152b}$$

For the power series in derivative operators applied to a power of a determinant, a simple application of the translation formula (5.1) yields

$$\phi_\Omega(\partial)(\det X)^n|_{X=0} = \exp\left(\sum_{\alpha,\beta=1}^\ell \omega_{\alpha\beta} \frac{\partial}{\partial x_{\alpha\beta}}\right)(\det X)^n|_{X=0} \tag{5.153a}$$

$$= (\det(X + \Omega))^n|_{X=0} \tag{5.153b}$$

$$= (\det \Omega)^n. \tag{5.153c}$$

For the power of a determinant in derivative operators applied to a power series, it suffices to cite the proof (5.153) and invoke transposition duality $X \leftrightarrow \partial$ in the Weyl algebra; but for completeness let us give a direct proof. Use a fermionic representation of the differential operator

$$\det(\partial) = \int \mathcal{D}_\ell(\eta, \bar{\eta}) \exp\left(\sum_{\alpha, \beta=1}^{\ell} \bar{\eta}_\alpha \frac{\partial}{\partial X_{\alpha\beta}} \eta_\beta\right) \tag{5.154}$$

and apply again the translation formula (5.1):

$$\det(\partial)[\exp \operatorname{tr}(\Omega^T X)] = \int \mathcal{D}_\ell(\eta, \bar{\eta}) \exp\left(\sum_{\alpha, \beta=1}^{\ell} \bar{\eta}_\alpha \frac{\partial}{\partial X_{\alpha\beta}} \eta_\beta\right) [\exp \operatorname{tr}(\Omega^T X)] \tag{5.155a}$$

$$= \int \mathcal{D}_\ell(\eta, \bar{\eta}) \exp \operatorname{tr}(\Omega^T (X + \bar{\eta} \eta^T)) \tag{5.155b}$$

$$= [\exp \operatorname{tr}(\Omega^T X)] \int \mathcal{D}_\ell(\eta, \bar{\eta}) \exp\left(\sum_{\alpha, \beta=1}^{\ell} \bar{\eta}_\alpha \omega_{\alpha\beta} \eta_\beta\right) \tag{5.155c}$$

$$= (\det \Omega) \exp \operatorname{tr}(\Omega^T X). \tag{5.155d}$$

Iterating n times, we get

$$\det(\partial)^n \phi_\Omega(X)|_{X=0} = (\det \Omega)^n \exp \operatorname{tr}(\Omega^T X)|_{X=0} = (\det \Omega)^n. \quad \square \tag{5.156}$$

5.9. Multi-matrix rectangular Cayley identity

In this subsection we shall prove the multi-matrix rectangular Cayley identity (Theorem 2.9), which is the most difficult of the identities proven in this paper. The difficulty arises principally from the fact that the number of matrices appearing in the identity (which we call ℓ) can be arbitrarily large, and we are required to provide a proof valid for all ℓ . The proof nevertheless follows the basic pattern of the proofs of the other identities (notably the two-matrix rectangular identity), and divides naturally into two parts:

- (i) We represent the differential operator as a Grassmann integral, and after several manipulations we are able to express $\det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}] \det(X^{(1)} \dots X^{(\ell)})^s$ as the desired quantity $\det(X^{(1)} \dots X^{(\ell)})^{s-1} \epsilon(I, J) \det[(X^{(1)} \dots X^{(\ell)})_{I^c J^c}]$ multiplied by a purely combinatorial factor $b(s)$ that is given as a Grassmann integral.
- (ii) We evaluate this Grassmann integral and prove that $b(s) = \prod_{\alpha=1}^{\ell} \prod_{j=1}^k (s + n_\alpha - j)$.

The major new complications, as compared to the preceding proofs, come from the fact that our Grassmann integral involves ℓ sets of fermionic fields, while in the preceding proofs we only had one or two, and the matrix arising from the application of the “low-rank perturbation lemma” (Lemma B.11) is of size $\ell \times \ell$, while in the preceding proofs we never had a matrix larger than 4×4 . Despite these complications, step (i) follows closely the model established in the preceding proofs and is not much more difficult than them: the main novelty is that we need to use a variant of the low-rank perturbation lemma (Corollary B.13) that is specially adapted to perturbations of the product form (5.170) encountered here. The big trouble arises in step (ii): to evaluate the Grassmann integral involving scalar products among ℓ sets of fermionic fields, we shall first rewrite it as a differential operator acting on a determinant (using Proposition 5.6) and then exploit the fact that the matrix arising in this differential operator is upper Hessenberg (i.e. has zero entries below the first subdiagonal); the latter leads to some special combinatorial/algebraic computations (Theorem 5.7 and Corollary 5.8).

It is convenient to formally divide the statement of Theorem 2.9 into two parts, corresponding to steps (i) and (ii) above:

Theorem 2.9, part (i). Fix integers $\ell \geq 1$ and $n_1, \dots, n_\ell \geq 0$ and write $n_{\ell+1} = n_1$. For $1 \leq \alpha \leq \ell$, let $X^{(\alpha)}$ be an $n_\alpha \times n_{\alpha+1}$ matrix of indeterminates, and let $\partial^{(\alpha)}$ be the corresponding matrix of partial derivatives. If $I, J \subseteq [n_1]$ with $|I| = |J| = k$, then

$$\det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}] \det(X^{(1)} \dots X^{(\ell)})^s = b_{n_1, \dots, n_\ell; k}(s) \det(X^{(1)} \dots X^{(\ell)})^{s-1} \epsilon(I, J) \det[(X^{(1)} \dots X^{(\ell)})_{I^c J^c}] \tag{5.157}$$

where

$$b_{n_1, \dots, n_\ell; k}(s) = (-1)^{k\ell} \int \mathcal{D}_k(\psi^1, \bar{\psi}^1) \dots \mathcal{D}_k(\psi^\ell, \bar{\psi}^\ell) \mathcal{D}_{n_2-n_1}(\eta^2, \bar{\eta}^2) \dots \mathcal{D}_{n_\ell-n_1}(\eta^\ell, \bar{\eta}^\ell) \times \exp \left[\sum_{\alpha=2}^{\ell} (\bar{\psi}^{\alpha T} \psi^{\alpha-1} + \bar{\eta}^{\alpha T} \eta^\alpha) \right] \det(I_\ell + M)^{-s} \tag{5.158}$$

and

$$M_{\alpha\beta} = \begin{cases} \bar{\psi}^{\alpha T} \psi^\beta & \text{if } \alpha \leq \beta, \\ -\bar{\eta}^{\alpha T} \eta^\alpha & \text{if } \alpha = \beta + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{5.159}$$

Theorem 2.9, part (ii). We have

$$b_{n_1, \dots, n_\ell; k}(s) = \prod_{\alpha=1}^{\ell} \prod_{j=1}^k (s + n_\alpha - j). \tag{5.160}$$

Let us now begin the proof of part (i) of Theorem 2.9. As before, we will need the full strength of Lemma 5.1 to handle the all-minors case.

Proof of Theorem 2.9, part (i). We can assume that $n_\alpha \geq n_1$ for $2 \leq \alpha \leq \ell$, since otherwise $\det(X^{(1)} \dots X^{(\ell)})$ is the zero polynomial.

We begin, as usual, by representing the differential operator as a Grassmann integral, this time exploiting Lemma B.14 (rather than just Corollary B.6). We have

$$\det(\partial^{(1)} \dots \partial^{(\ell)}) = \int \mathcal{D}_{n_1}(\psi^1, \bar{\psi}^1) \dots \mathcal{D}_{n_\ell}(\psi^\ell, \bar{\psi}^\ell) \times \exp \left[\sum_{\alpha=2}^{\ell} (\bar{\psi}^{\alpha T} \psi^\alpha - \bar{\psi}^{\alpha-1 T} \partial^{(\alpha-1)} \psi^\alpha) + \bar{\psi}^{\ell T} \partial^{(\ell)} \psi^1 \right]. \tag{5.161}$$

Here $\psi_i^\alpha, \bar{\psi}_i^\alpha$ ($1 \leq \alpha \leq \ell, 1 \leq i \leq n_\alpha$) are Grassmann variables, and the subscripts on \mathcal{D} serve to remind us of the length of each vector; shorthand notations for index summations are understood, e.g. $\bar{\psi}^{\alpha T} \partial^{(a)} \psi^{a+1} \equiv \sum_{i=1}^{n_\alpha} \sum_{j=1}^{n_{a+1}} \bar{\psi}_i^\alpha \psi_j^{a+1} \partial / \partial x_{ij}^{(a)}$. For a general minor $I, J \subseteq [n_1]$ with $|I| = |J| = k$, the quantity $\det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}]$ has a representation like (5.161) but with an extra factor $\epsilon(I, J)(\prod \bar{\psi}^1 \psi^1)_{I^c, J^c}$ in the integrand.

It is convenient to make the change of variables $\psi^1 \rightarrow -\psi^1$, as this makes the summand $\bar{\psi}^\ell \partial^{(\ell)} \psi^1$ analogous to the $\bar{\psi}^{\alpha-1} \partial^{(\alpha-1)} \psi^\alpha$ arising for $2 \leq \alpha \leq \ell$. We shall exploit this structure by writing $\psi^{\ell+1}$ as a synonym for ψ^1 . This change of variables introduces an overall factor $(-1)^{n_1} (-1)^{n_1-k} = (-1)^k$.

Applying the translation formula (5.1) to the whole set of variables $\{X_{ij}^{(\alpha)}\}$ produces

$$\begin{aligned} & \det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}] f(X^{(1)}, \dots, X^{(\ell)}) \\ &= \epsilon(I, J) \int \mathcal{D}_{n_1}(\psi^1, \bar{\psi}^1) \dots \mathcal{D}_{n_\ell}(\psi^\ell, \bar{\psi}^\ell) \\ & \quad \times (-1)^k \exp\left[\sum_{\alpha=2}^{\ell} \bar{\psi}^\alpha \psi^\alpha\right] \left(\prod \bar{\psi}^1 \psi^1\right)_{I^c, J^c} f(X_{\text{trans}}^{(1)}, \dots, X_{\text{trans}}^{(\ell)}) \end{aligned} \tag{5.162}$$

for an arbitrary polynomial f , where we have introduced the shorthand

$$X_{\text{trans}}^{(\alpha)} = X^{(\alpha)} - \bar{\psi}^\alpha (\psi^{\alpha+1})^T \tag{5.163}$$

for the arguments of f . We shall use the formula (5.162) in the case $f(X^{(1)}, \dots, X^{(\ell)}) = \det(X^{(1)} \dots X^{(\ell)})^s$ where s is a positive integer.

Suppose now that the $X^{(\alpha)}$ are real matrices of rank $\min(n_\alpha, n_{\alpha+1})$ that are sufficiently close to the matrix $\hat{I}_{n_\alpha n_{\alpha+1}}$ defined in (5.64). Then by Lemma B.21 we can find matrices $P_\alpha \in GL(n_\alpha)$ for $1 \leq \alpha \leq \ell + 1$ such that $X^{(\alpha)} = P_\alpha \hat{I}_{n_\alpha n_{\alpha+1}} P_{\alpha+1}^{-1}$. We have $X^{(1)} \dots X^{(\ell)} = P_1 P_{\ell+1}^{-1}$ since $\hat{I}_{n_1 n_2} \hat{I}_{n_2 n_3} \dots \hat{I}_{n_\ell n_1} = I_{n_1}$ as a consequence of the fact that $n_\alpha \geq n_1$ for $2 \leq \alpha \leq \ell$. Therefore

$$\det(X^{(1)} \dots X^{(\ell)}) = \det(P_1) \det(P_{\ell+1})^{-1}. \tag{5.164}$$

We also have

$$X_{\text{trans}}^{(1)} \dots X_{\text{trans}}^{(\ell)} = \prod_{\alpha=1}^{\ell} P_\alpha [\hat{I}_{n_\alpha n_{\alpha+1}} - P_\alpha^{-1} \bar{\psi}^\alpha (\psi^{\alpha+1})^T P_{\alpha+1}] P_{\alpha+1}^{-1} \tag{5.165}$$

where the product is taken from left ($\alpha = 1$) to right ($\alpha = \ell$).

Let us now change variables from $(\psi^\alpha, \bar{\psi}^\alpha)$ to $(\psi'^\alpha, \bar{\psi}'^\alpha)$ defined by

$$\psi'^\alpha = \begin{cases} P_\alpha^T \psi^\alpha & \text{for } 2 \leq \alpha \leq \ell, \\ P_{\ell+1}^T \psi^1 & \text{for } \alpha = 1, \end{cases} \tag{5.166}$$

$$\bar{\psi}'^\alpha = P_\alpha^{-1} \bar{\psi}^\alpha. \tag{5.167}$$

The Jacobian is $(\det P_1^{-1})(\det P_{\ell+1}) = \det(X^{(1)} \dots X^{(\ell)})^{-1}$ [using (5.164)]. In the new variables we have (dropping now the primes from the notation)

$$X_{\text{trans}}^{(1)} \dots X_{\text{trans}}^{(\ell)} = P_1 \left(\prod_{\alpha=1}^{\ell} [\hat{I}_{n_\alpha n_{\alpha+1}} - \bar{\psi}^\alpha (\psi^{\alpha+1})^T] \right) P_{\ell+1}^{-1}, \tag{5.168}$$

so again using (5.164) we see that the translated determinant is given by

$$\det(X_{\text{trans}}^{(1)} \dots X_{\text{trans}}^{(\ell)}) = \det(X^{(1)} \dots X^{(\ell)}) (\det M_{n_1, \dots, n_\ell}) \tag{5.169}$$

where the matrix M_{n_1, \dots, n_ℓ} depends only on the Grassmann variables:

$$M_{n_1, \dots, n_\ell} = \prod_{\alpha=1}^{\ell} [\hat{I}_{n_\alpha n_{\alpha+1}} - \bar{\psi}^\alpha (\psi^{\alpha+1})^T]. \tag{5.170}$$

Therefore

$$\begin{aligned} & \det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}] \det(X^{(1)} \dots X^{(\ell)})^s \\ &= \epsilon(I, J) \det(X^{(1)} \dots X^{(\ell)})^{s-1} (-1)^k \int \mathcal{D}_{n_1}(\psi^1, \bar{\psi}^1) \dots \mathcal{D}_{n_\ell}(\psi^\ell, \bar{\psi}^\ell) \\ & \quad \times \left(\prod (P_1 \bar{\psi}^1) (P_{\ell+1}^{-T} \psi^1) \right)_{J^c, J^c} \exp \left[\sum_{\alpha=2}^{\ell} \bar{\psi}^\alpha \psi^\alpha \right] \det(M_{n_1, \dots, n_\ell})^s. \end{aligned} \tag{5.171}$$

To evaluate $\det(M_{n_1, \dots, n_\ell})$, we shall use Corollary B.13, which is a variant of the low-rank perturbation lemma that is specially adapted to matrices of the form (5.170). Before doing so, it is convenient to split the vectors ψ^α and $\bar{\psi}^\alpha$ as

$$(\psi_1^\alpha, \dots, \psi_{n_\alpha}^\alpha) = (\lambda_1^\alpha, \dots, \lambda_{n_1}^\alpha, \zeta_1^\alpha, \dots, \zeta_{m_\alpha}^\alpha), \tag{5.172a}$$

$$(\bar{\psi}_1^\alpha, \dots, \bar{\psi}_{n_\alpha}^\alpha) = (\bar{\lambda}_1^\alpha, \dots, \bar{\lambda}_{n_1}^\alpha, \bar{\zeta}_1^\alpha, \dots, \bar{\zeta}_{m_\alpha}^\alpha), \tag{5.172b}$$

where $m_\alpha := n_\alpha - n_1$ (note in particular that $\psi^1 = \lambda^1$ and $\bar{\psi}^1 = \bar{\lambda}^1$). We now apply Corollary B.13 with $x_\alpha = \bar{\psi}^\alpha$, $y_\alpha = \psi^{\alpha+1}$ and $\epsilon = -1$ to obtain

$$\det(M_{n_1, \dots, n_\ell}) = (\det N)^{-1} \tag{5.173}$$

where the $\ell \times \ell$ matrix N is defined by

$$N_{\alpha\beta} = \begin{cases} \sum_{i=1}^{m_{\alpha+1, \beta}} \zeta_i^{\alpha+1} \bar{\zeta}_i^\beta & \text{if } \alpha < \beta, \\ \delta_{\alpha\beta} - \lambda^{\alpha+1T} \bar{\lambda}^\beta & \text{if } \alpha \geq \beta \end{cases} \tag{5.174}$$

and $m_{\alpha, \beta} := \min_{\alpha \leq \gamma \leq \beta} m_\gamma$. We now fall into the conditions for the application of Lemma 5.1 with $\eta = \psi^1 = \lambda^1$, $\bar{\eta} = \bar{\psi}^1 = \bar{\lambda}^1$ and $\theta = \{\lambda^2, \dots, \lambda^\ell, \bar{\lambda}^2, \dots, \bar{\lambda}^\ell\}$ (here the variables ζ and $\bar{\zeta}$ just go for the ride), yielding

$$\begin{aligned} & \det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}] \det[X^{(1)} \dots X^{(\ell)}]^s \\ &= \epsilon(I, J) \det[X^{(1)} \dots X^{(\ell)}]^{s-1} \det[(X^{(1)} \dots X^{(\ell)})_{J^c, J^c}] (-1)^k \\ & \quad \times \int \mathcal{D}_{n_1}(\lambda^1, \bar{\lambda}^1) \dots \mathcal{D}_{n_1}(\lambda^\ell, \bar{\lambda}^\ell) \mathcal{D}_{m_2}(\zeta^2, \bar{\zeta}^2) \dots \mathcal{D}_{m_\ell}(\zeta^\ell, \bar{\zeta}^\ell) \\ & \quad \times \left(\prod_{j=k+1}^{n_1} \bar{\lambda}_j^1 \lambda_j^1 \right) \exp \left[\sum_{\alpha=2}^{\ell} (\bar{\lambda}^{\alpha T} \lambda^\alpha + \bar{\zeta}^{\alpha T} \zeta^\alpha) \right] \det(N)^{-s}. \end{aligned} \tag{5.175}$$

We have thus represented $\det[(\partial^{(1)} \dots \partial^{(\ell)})_{IJ}] \det(X^{(1)} \dots X^{(\ell)})^s$ as the desired quantity $\det(X^{(1)} \dots X^{(\ell)})^{s-1} \epsilon(I, J) \det[(X^{(1)} \dots X^{(\ell)})_{J^c, J^c}]$ multiplied by a purely combinatorial factor $b_{n_1, \dots, n_\ell; k}(s)$ that is given as a Grassmann integral:

$$\begin{aligned}
 b_{n_1, \dots, n_\ell; k}(s) &:= (-1)^k \int \mathcal{D}_{n_1}(\lambda^1, \bar{\lambda}^1) \cdots \mathcal{D}_{n_1}(\lambda^\ell, \bar{\lambda}^\ell) \mathcal{D}_{m_2}(\zeta^2, \bar{\zeta}^2) \cdots \mathcal{D}_{m_\ell}(\zeta^\ell, \bar{\zeta}^\ell) \\
 &\quad \times \left(\prod_{j=k+1}^{n_1} \bar{\lambda}_j^1 \lambda_j^1 \right) \exp \left[\sum_{\alpha=2}^{\ell} (\bar{\lambda}^{\alpha T} \lambda^\alpha + \bar{\zeta}^{\alpha T} \zeta^\alpha) \right] \det(N)^{-s}. \tag{5.176}
 \end{aligned}$$

In order to handle the factor $\prod_{j=k+1}^{n_1} \bar{\lambda}_j^1 \lambda_j^1$, it is convenient to further split the fields λ and $\bar{\lambda}$ as

$$(\lambda_1^\alpha, \dots, \lambda_{n_1}^\alpha) = (\lambda_1^\alpha, \dots, \lambda_k^\alpha, \chi_1^\alpha, \dots, \chi_{n_1-k}^\alpha), \tag{5.177a}$$

$$(\bar{\lambda}_1^\alpha, \dots, \bar{\lambda}_{n_1}^\alpha) = (\bar{\lambda}_1^\alpha, \dots, \bar{\lambda}_k^\alpha, \bar{\chi}_1^\alpha, \dots, \bar{\chi}_{n_1-k}^\alpha). \tag{5.177b}$$

Notice now that the overall factor $\prod_{j=k+1}^{n_1} \bar{\lambda}_j^1 \lambda_j^1 = \prod_{i=1}^{n_1-k} \bar{\chi}_i^1 \chi_i^1$ in the integrand kills all monomials in the expansion of the rest of the integrand that contain any field χ^1 or $\bar{\chi}^1$. Now, the factor $(\det N)^{-s}$ in (5.176) depends on the fields $\lambda, \bar{\lambda}, \chi, \bar{\chi}, \zeta, \bar{\zeta}$ only through products of the forms $\{\lambda_i^\alpha \bar{\lambda}_i^\beta\}_{1 \leq \beta < \alpha \leq \ell}, \{\lambda_i^1 \bar{\lambda}_i^\beta\}_{1 \leq \beta \leq \ell}, \{\chi_i^\alpha \bar{\chi}_i^\beta\}_{1 \leq \beta < \alpha \leq \ell}$ and $\{\zeta_i^\alpha \bar{\zeta}_i^\beta\}_{2 \leq \alpha \leq \beta \leq \ell}$, while the exponential depends only on combinations $\bar{\lambda}^\alpha \lambda^\alpha, \bar{\chi}^\alpha \chi^\alpha$ and $\bar{\zeta}^\alpha \zeta^\alpha$ that are “charge-neutral” in each field separately. Therefore, the only monomials in the expansion of $(\det N)^{-s}$ that can contribute to the integral must also be charge-neutral in each field separately. But since no monomial containing any χ^1 can arise, it is impossible to make such a charge-neutral combination using any other χ or $\bar{\chi}$ since all such terms are of the form $\chi^\alpha \bar{\chi}^\beta$ with $\beta < \alpha$. In a similar way, the terms $\zeta^\alpha \bar{\zeta}^\beta$ with $2 \leq \alpha < \beta \leq \ell$ cannot contribute. (The combinations $\zeta^\alpha \bar{\zeta}^\alpha$ do survive.) The integral (5.176) will therefore be unchanged if we replace N by a new matrix N' in which all these “forbidden combinations” are set to zero:

$$N'_{\alpha\beta} = \begin{cases} \zeta^{\beta T} \bar{\zeta}^\beta & \text{if } \alpha = \beta - 1, \\ 0 & \text{if } \alpha < \beta - 1, \\ \delta_{\alpha\beta} - \lambda^{\alpha+1 T} \bar{\lambda}^\beta & \text{if } \alpha \geq \beta. \end{cases} \tag{5.178}$$

Note that the matrix N' is lower Hessenberg (i.e. has zero entries above the first superdiagonal). Summarizing, we have

$$\begin{aligned}
 b_{n_1, \dots, n_\ell; k}(s) &= (-1)^k \int \mathcal{D}_k(\lambda^1, \bar{\lambda}^1) \cdots \mathcal{D}_k(\lambda^\ell, \bar{\lambda}^\ell) \\
 &\quad \times \mathcal{D}_{n_1-k}(\chi^1, \bar{\chi}^1) \cdots \mathcal{D}_{n_1-k}(\chi^\ell, \bar{\chi}^\ell) \mathcal{D}_{m_2}(\zeta^2, \bar{\zeta}^2) \cdots \mathcal{D}_{m_\ell}(\zeta^\ell, \bar{\zeta}^\ell) \\
 &\quad \times \left(\prod_{i=1}^{n_1-k} \bar{\chi}_i^1 \chi_i^1 \right) \exp \left[\sum_{\alpha=2}^{\ell} (\bar{\lambda}^{\alpha T} \lambda^\alpha + \bar{\chi}^{\alpha T} \chi^\alpha + \bar{\zeta}^{\alpha T} \zeta^\alpha) \right] \det(N')^{-s}. \tag{5.179}
 \end{aligned}$$

Since N' does not contain χ or $\bar{\chi}$, we can immediately perform the integrations over these fields, yielding 1.

To make the indices in the matrix N' look nicer, we perform the change of variables from λ to λ' defined by $(\lambda')^\alpha = \lambda^{\alpha+1}$ for $1 \leq \alpha \leq \ell$ (and recalling that $\lambda^{\ell+1}$ is a shorthand for λ^1); the variables $\bar{\lambda}$ are left as is. The Jacobian is $(-1)^{k(\ell-1)}$. So, dropping primes, we have

$$\begin{aligned}
 b_{n_1, \dots, n_\ell; k}(s) &= (-1)^k \int \mathcal{D}_k(\lambda^1, \bar{\lambda}^1) \cdots \mathcal{D}_k(\lambda^\ell, \bar{\lambda}^\ell) \mathcal{D}_{m_2}(\zeta^2, \bar{\zeta}^2) \cdots \mathcal{D}_{m_\ell}(\zeta^\ell, \bar{\zeta}^\ell) \\
 &\quad \times \exp \left[\sum_{\alpha=2}^{\ell} (\bar{\lambda}^{\alpha T} \lambda^{\alpha-1} + \bar{\zeta}^{\alpha T} \zeta^\alpha) \right] \det(N'')^{-s} \tag{5.180}
 \end{aligned}$$

where

$$N''_{\alpha\beta} = \begin{cases} \zeta^{\beta T} \bar{\zeta}^\beta & \text{if } \alpha = \beta - 1, \\ 0 & \text{if } \alpha < \beta - 1, \\ \delta_{\alpha\beta} - \lambda^{\alpha T} \bar{\lambda}^\beta & \text{if } \alpha \geq \beta. \end{cases} \tag{5.181}$$

Finally, in order to put our scalar products of complex fermions in the standard forms $\bar{\zeta}^T \zeta$ and $\bar{\lambda}^T \lambda$, we anticommute all the bilinears (obtaining a minus sign); and in order to keep the indices in a natural notation, we replace N'' by its transpose. After a renaming $\lambda \rightarrow \psi$ and $\zeta \rightarrow \eta$, the result is (5.158)/(5.159), where $(N'')^T = I_\ell + M$.

This proves part (i) of Theorem 2.9 when the $X^{(\alpha)}$ are real matrices of rank $\min(n_\alpha, n_{\alpha+1})$ lying in a sufficiently small neighborhood of $\hat{I}_{n_\alpha n_{\alpha+1}}$, and s is a positive integer; the general validity of the identity then follows from Proposition 2.18. \square

We now turn to the evaluation of the Grassmann integral (5.158)/(5.159) for $b_{n_1, \dots, n_\ell; k}(s)$. The integrand depends on the fields only through the scalar products $\{\bar{\psi}^{\alpha T} \psi^\beta\}_{\alpha, \beta=1}^\ell$ and $\{\bar{\eta}^{\alpha T} \eta^\alpha\}_{\alpha=2}^\ell$. We can therefore apply Proposition 5.6 once to the entire set of variables $\psi, \bar{\psi}$ and separately to the variables $\eta^\alpha, \bar{\eta}^\alpha$ for each α ($2 \leq \alpha \leq \ell$). We therefore introduce indeterminates $X = (x_{\alpha\beta})_{\alpha, \beta=1}^\ell$ and $y = (y_\alpha)_{\alpha=2}^\ell$, along with the corresponding differential operators $\partial/\partial x_{\alpha\beta}$ and $\partial/\partial y_\alpha$. Using Proposition 5.6 in the form (5.149b), we obtain (after renaming y_2, \dots, y_ℓ as $y_1, \dots, y_{\ell-1}$)

$$b_{n_1, \dots, n_\ell; k}(s) = (-1)^{k\ell} (\det \hat{M})^{-s} \times \exp \left[\sum_{\alpha=1}^{\ell-1} \left(\frac{\partial}{\partial x_{\alpha+1, \alpha}} + \frac{\partial}{\partial y_\alpha} \right) \right] \det(X)^k \prod_{\alpha=1}^{\ell-1} y_\alpha^{m_\alpha} \Big|_{X=y=0}, \tag{5.182}$$

where now \hat{M} reads

$$\hat{M} = \begin{pmatrix} 1 + \frac{\partial}{\partial x_{11}} & \frac{\partial}{\partial x_{12}} & \frac{\partial}{\partial x_{13}} & \cdots & \frac{\partial}{\partial x_{1\ell}} \\ -\frac{\partial}{\partial y_1} & 1 + \frac{\partial}{\partial x_{22}} & \frac{\partial}{\partial x_{23}} & \cdots & \frac{\partial}{\partial x_{2\ell}} \\ 0 & -\frac{\partial}{\partial y_2} & 1 + \frac{\partial}{\partial x_{33}} & \cdots & \frac{\partial}{\partial x_{3\ell}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -\frac{\partial}{\partial y_{\ell-1}} & 1 + \frac{\partial}{\partial x_{\ell\ell}} \end{pmatrix}. \tag{5.183}$$

We can now apply the translation formula (5.1) to the exponential of the differential operator: this transforms $\prod_{\alpha=1}^{\ell-1} y_\alpha^{m_\alpha}$ into $\prod_{\alpha=1}^{\ell-1} (1 + y_\alpha)^{m_\alpha}$ and $x_{\alpha\beta}$ into $x'_{\alpha\beta} = x_{\alpha\beta} + \delta_{\alpha, \beta+1}$. Thus

$$b_{n_1, \dots, n_\ell; k}(s) = (-1)^{k\ell} (\det \hat{M})^{-s} \det(X')^k \prod_{\alpha=1}^{\ell-1} (1 + y_\alpha)^{m_\alpha} \Big|_{X=y=0}. \tag{5.184}$$

Finally, we observe that \hat{M} does not contain the differential operators $\partial/\partial x_{\alpha\beta}$ with $\alpha > \beta$, so in X' we can set those variables $x_{\alpha\beta}$ to zero immediately. We thus have (5.184) where now

$$X' = \begin{pmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1\ell} \\ 1 & x_{22} & x_{23} & \cdots & x_{2\ell} \\ 0 & 1 & x_{33} & \cdots & x_{3\ell} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & x_{\ell\ell} \end{pmatrix}. \tag{5.185}$$

Since the evaluation of (5.184) will involve a recursion in ℓ , it is convenient to introduce an infinite set of indeterminates $\{x_{\alpha\beta}\}_{1 \leq \alpha \leq \beta < \infty}$, along with the corresponding set of differential operators $\partial_{\alpha\beta} = \partial/\partial x_{\alpha\beta}$, as well as another infinite set of indeterminates $\{a_\alpha\}_{1 \leq \alpha \leq n-1}$ (only finitely many of these will play any role at any given stage). Then define, for each $\ell \geq 1$, the quantities

$$D_\ell(a) = \det \begin{pmatrix} 1 + \partial_{11} & \partial_{12} & \partial_{13} & \cdots & \partial_{1\ell} \\ a_1 & 1 + \partial_{22} & \partial_{23} & \cdots & \partial_{2\ell} \\ 0 & a_2 & 1 + \partial_{33} & \cdots & \partial_{3\ell} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{\ell-1} & 1 + \partial_{\ell\ell} \end{pmatrix} \tag{5.186}$$

and

$$X_\ell = \det \begin{pmatrix} x_{11} & x_{12} & x_{13} & \cdots & x_{1\ell} \\ 1 & x_{22} & x_{23} & \cdots & x_{2\ell} \\ 0 & 1 & x_{33} & \cdots & x_{3\ell} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & x_{\ell\ell} \end{pmatrix} \tag{5.187}$$

where we also set $D_0(a) = 1$ and $X_0 = 1$. Note also that $D_\ell(a)$ [resp. X_ℓ] involves only those $\partial_{\alpha\beta}$ [resp. $x_{\alpha\beta}$] with $\alpha \leq \beta \leq \ell$.

Given a formal indeterminate s and a nonnegative integer k , our goal in the remainder of this subsection is to compute the expression

$$D_\ell(a)^{-s} X_\ell^k \Big|_{X=0}, \tag{5.188}$$

which abstracts the relevant features of (5.184) [as s is a formal variable, the choice between s and $-s$ has no special role and is made here just for convenience]. Since $D_\ell(a)$ is a polynomial in the quantities $\{\partial_{\alpha\beta}\}$ and $\{a_\alpha\}$ with constant term 1, $D_\ell(a)^{-s}$ is here to be understood as the series

$$D_\ell(a)^{-s} = \sum_{h=0}^{\infty} \binom{-s}{h} [D_\ell(a) - 1]^h \tag{5.189}$$

which, when applied to X_ℓ^k as in (5.188), can be truncated to $h \leq k\ell$. We shall prove the following:

Theorem 5.7. *With the definitions above, we have*

$$D_\ell(a)^{-s} X_\ell^k \Big|_{X=0} = k! \binom{-s}{k} \prod_{\alpha=1}^{\ell-1} \sum_{b=0}^k k! \binom{-s-b}{k-b} \frac{a_\alpha^b}{b!}. \tag{5.190}$$

We remark that

$$\sum_{b=0}^k \binom{-s-b}{k-b} \frac{z^b}{b!} = \binom{-s}{k} {}_1F_1(-k; s; z) \tag{5.191}$$

although we will not use this.

What we shall actually need is a specific corollary of Theorem 5.7. Let us introduce the variables $\{y_\alpha\}_{1 \leq \alpha < \infty}$ and the associated derivatives $\hat{\partial}_\alpha = \partial/\partial y_\alpha$, as well as the further indeterminates (or non-negative integers) $\{m_\alpha\}_{1 \leq \alpha < \infty}$.

Corollary 5.8. *With the definitions above, we have*

$$D_\ell(-\hat{\partial})^{-s} X_\ell^k \prod_{\alpha=1}^{\ell-1} (1 + y_\alpha)^{m_\alpha} \Big|_{X=y=0} = (-1)^{k\ell} \prod_{\alpha=0}^{\ell-1} \prod_{i=0}^{k-1} (s + m_\alpha + i) \tag{5.192}$$

with the convention $m_0 = 0$.

Given Corollary 5.8, the proof of Theorem 2.9, part (ii) is a triviality:

Proof of Theorem 2.9, part (ii), given Corollary 5.8. It suffices to recognize $\det[\hat{M}]$ and $\det(X')$ as the operators $D_\ell(-\hat{\partial})$ and X_ℓ in Corollary 5.8. The sign $(-1)^{k\ell}$ in (5.184) combines with the one in (5.192), leaving exactly the prefactor claimed in Theorem 2.9. \square

Let us next show how to deduce Corollary 5.8 from Theorem 5.7:

Proof of Corollary 5.8, given Theorem 5.7. We evaluate the left-hand side of (5.192) by using (5.190) with a_α replaced by $-\hat{\partial}_\alpha$. Since

$$(-\hat{\partial}_\alpha)^b (1 + y_\alpha)^{m_\alpha} \Big|_{y_\alpha=0} = (-1)^b b! \binom{m_\alpha}{b} \tag{5.193}$$

and everything is factorized over α , we obtain

$$D_\ell(-\hat{\partial})^{-s} X_\ell^k \prod_{\alpha=1}^{\ell-1} (1 + y_\alpha)^{m_\alpha} \Big|_{X=y=0} = k! \binom{-s}{k} \prod_{\alpha=1}^{\ell-1} \sum_{b=0}^k k! \binom{-s-b}{k-b} (-1)^b \binom{m_\alpha}{b}. \tag{5.194}$$

Then for each α we have

$$\begin{aligned} k! \sum_{b=0}^k \binom{-s-b}{k-b} (-1)^b \binom{m_\alpha}{b} &= k! \sum_{b=0}^k (-1)^{k-b} \binom{s+k-1}{k-b} (-1)^b \binom{m_\alpha}{b} \\ &= (-1)^k k! \sum_{b=0}^k \binom{s+k-1}{k-b} \binom{m_\alpha}{b} \\ &= (-1)^k k! \binom{s+k+m_\alpha-1}{k} \\ &= (-1)^k \prod_{i=0}^{k-1} (s + m_\alpha + i) \end{aligned} \tag{5.195}$$

where we used the Chu–Vandermonde convolution (Lemma B.2) in going from the second to the third line. The prefactor $k! \binom{-s}{k}$ gives an analogous contribution with $m_0 = 0$. \square

Finally, we turn to the proof of Theorem 5.7. We shall need two main lemmas: one that essentially provides an inductive step, and another dealing with a special sum of multinomial coefficients. Henceforth we shall write D_ℓ as a shorthand for $D_\ell(a)$.

We start with a pair of easy recursive formulae, obtained by expansion of the determinant on the last column:

Lemma 5.9.

$$D_\ell = D_{\ell-1} + \sum_{\alpha=1}^{\ell} (-1)^{\ell-\alpha} a_\alpha \cdots a_{\ell-1} D_{\alpha-1} \partial_{\alpha\ell} \tag{5.196}$$

(the empty product $a_\alpha \cdots a_{\ell-1}$ for $\alpha = \ell$ should be understood as 1) and

$$X_\ell = \sum_{\alpha=1}^{\ell} (-1)^{\ell-\alpha} X_{\alpha-1} X_{\alpha\ell}. \tag{5.197}$$

Related formulae for the determinant of a Hessenberg matrix can be found in [99,100]. The induction lemma is the following:

Lemma 5.10. *Let $m \geq 1$ and $c_1, \dots, c_m \geq 0$ be integers, and let t be an indeterminate. Then*

$$\begin{aligned} D_m^t \prod_{\alpha=1}^m D_\alpha^{c_\alpha} \prod_{\alpha=1}^m X_\alpha^{c_\alpha} \Big|_{x_{\alpha m}=0} &= \sum_{\substack{b_1, \dots, b_m \geq 0 \\ b_1 + \dots + b_m = c_m}} \binom{t + c_m}{c_m} c_m! \binom{c_m}{b_1, \dots, b_m} \prod_{\alpha=1}^{m-1} a_\alpha^{\sum_{\beta=1}^{\alpha} b_\beta} \\ &\times D_{m-1}^t \prod_{\alpha=1}^{m-1} D_\alpha^{c_\alpha + b_{\alpha+1}} \prod_{\alpha=1}^{m-1} X_\alpha^{c_\alpha + b_{\alpha+1}}. \end{aligned} \tag{5.198}$$

Proof. Notice, first of all, that the factors D_α and X_α for $\alpha < m$ do not play any role, i.e. we can rewrite the left-hand side of (5.198) as

$$\left(\prod_{\alpha=1}^{m-1} D_\alpha^{c_\alpha} \right) (D_m^{c_m+t} X_m^{c_m} |_{x_{\alpha m}=0}) \left(\prod_{\alpha=1}^{m-1} X_\alpha^{c_\alpha} \right) \tag{5.199}$$

and concentrate on the central factor alone. To compute $D_m^{c_m+t} X_m^{c_m} |_{x_{\alpha m}=0}$, we expand $D_m^{c_m+t}$ and $X_m^{c_m}$ using Lemma 5.9:

$$\begin{aligned} D_m^{t+c_m} &= \sum_{b_1, \dots, b_m \geq 0} \binom{t + c_m}{b_1 + \dots + b_m} \binom{b_1 + \dots + b_m}{b_1, \dots, b_m} D_{m-1}^{t+c_m-(b_1+\dots+b_m)} \\ &\times \prod_{\alpha=1}^m [(-1)^{m-\alpha} (a_\alpha \cdots a_{m-1}) D_{\alpha-1} \partial_{\alpha m}]^{b_\alpha}, \end{aligned} \tag{5.200}$$

$$X_m^{c_m} = \sum_{\substack{b_1, \dots, b_m \geq 0 \\ b_1 + \dots + b_m = c_m}} \binom{c_m}{b_1, \dots, b_m} \prod_{\alpha=1}^m [(-1)^{m-\alpha} X_{\alpha m} X_{\alpha-1}]^{b_\alpha}. \tag{5.201}$$

Since $\partial_{\alpha m}^b X_{\alpha m}^{b'} |_{x_{\alpha m}=0} = \delta_{b,b'} b!$, the two sets of summation indices, when combined inside $D_m^{c_m+t} X_m^{c_m} |_{x_{\alpha m}=0}$, must coincide, and we get

$$\begin{aligned}
 D_m^{c_m+t} X_m^{c_m} \Big|_{X_{\alpha m}=0} &= \sum_{\substack{b_1, \dots, b_m \geq 0 \\ b_1 + \dots + b_m = c_m}} \binom{t + c_m}{c_m} c_m! \binom{c_m}{b_1, \dots, b_m} D_{m-1}^t \\
 &\times \prod_{\alpha=1}^m [(a_\alpha \cdots a_{m-1}) D_{\alpha-1}]^{b_\alpha} \prod_{\alpha=1}^m (X_{\alpha-1})^{b_\alpha}. \tag{5.202}
 \end{aligned}$$

In the two final products, we can drop the factors $D_0^{b_1}$ and $X_0^{b_1}$ since $D_0 = X_0 = 1$. Reintroducing the missing factors from (5.199), we obtain (5.198). \square

We will now apply Lemma 5.10 for ℓ “rounds”, starting with the initial conditions $m = \ell, t = -s - k, c_1 = \dots = c_{\ell-1} = 0$ and $c_\ell = k$. At round i we have $m = \ell + 1 - i$. Let us denote by $c_1^i, \dots, c_{\ell+1-i}^i$ the parameters $\{c_j\}$ immediately before entering round i , and let us denote by $b_1^i, \dots, b_{\ell+1-i}^i$ the summation indices in round i . We therefore have the initial conditions

$$c_1^1 = \dots = c_{\ell-1}^1 = 0, \quad c_\ell^1 = k \tag{5.203}$$

and the recursion

$$c_j^{i+1} = c_j^i + b_{j+1}^i \quad \text{for } 1 \leq j \leq \ell - i \tag{5.204}$$

[cf. (5.198)]. The summation indices $b_1^i, \dots, b_{\ell+1-i}^i$ obey the constraint

$$\sum_{j=1}^{\ell+1-i} b_j^i = c_{\ell+1-i}^i. \tag{5.205}$$

Using (5.203) and (5.204) we prove by induction that

$$c_{\ell+1-i}^i = \sum_{h=1}^{i-1} b_{\ell+2-i}^h \quad \text{for } 2 \leq i \leq \ell. \tag{5.206}$$

It is convenient to arrange the summation indices $\{b_j^i\}_{i+j \leq \ell+1}$ into a matrix

$$B = \begin{pmatrix}
 b_1^1 & b_2^1 & b_3^1 & \dots & b_\ell^1 \\
 b_1^2 & b_2^2 & \dots & b_{\ell-1}^2 & \\
 b_1^3 & \vdots & \ddots & & \\
 \vdots & b_2^{\ell-1} & & & \\
 b_1^\ell & & & &
 \end{pmatrix} \tag{5.207}$$

in which the first row sums to k and, for $2 \leq i \leq \ell$, row i and column $\ell + 2 - i$ have the same sum. Such matrices can be characterized as follows:

Lemma 5.11. Fix integers $\ell \geq 1$ and $k \geq 0$. For a matrix $B = (b_j^i)_{i+j \leq \ell+1}$ of nonnegative integers as in (5.207), the following conditions are equivalent:

- (a) The first row sums to k and, for $2 \leq i \leq \ell$, row i and column $\ell + 2 - i$ have the same sum.
- (b) For $1 \leq h \leq \ell$ we have

$$\sum_{\substack{1 \leq i \leq h \\ 1 \leq j \leq \ell + 1 - h}} b_j^i = k.$$

- (c) There exist nonnegative integers $b_\ell^2, b_{\ell-1}^3, \dots, b_2^\ell$ completing the matrix B to one in which all the row and column sums are equal to k . (Such numbers are obviously unique if they exist, and must lie in the interval $[0, k]$.)

Note in particular that statement (b) with $h = \ell$ tells us that the first column sums to k .

Proof of Lemma 5.11. (a) \implies (b): By induction on h . By hypothesis the equality holds for $h = 1$. Then for $h \geq 2$ we have

$$\sum_{\substack{1 \leq i \leq h \\ 1 \leq j \leq \ell + 1 - h}} b_j^i - \sum_{\substack{1 \leq i \leq h-1 \\ 1 \leq j \leq \ell + 2 - h}} b_j^i = \sum_{1 \leq j \leq \ell - h} b_j^h - \sum_{1 \leq i \leq h-1} b_{\ell + 2 - h}^i = 0 \tag{5.208}$$

by hypothesis.

(b) \implies (c): It is easily checked that the definition

$$b_{\ell + 2 - h}^h = \sum_{\substack{1 \leq i \leq h-1 \\ 1 \leq j \leq \ell + 1 - h}} b_j^i \quad \text{for } 2 \leq h \leq \ell \tag{5.209}$$

does what is required.

(c) \implies (a) is obvious. \square

Remark. An analogous equivalence holds, with the same proof (*mutatis mutandis*), for matrices B of nonnegative real numbers where k is a fixed nonnegative real number.

The ℓ -fold application of Lemma 5.10 with the initial conditions (5.203) gives rise to a sum over matrices B satisfying the equivalent conditions (a)–(c) of Lemma 5.11. To each such matrix there corresponds an ℓ -tuple $\bar{c} = (\bar{c}^1, \dots, \bar{c}^\ell)$ of integers in the range $[0, k]$, where $\bar{c}^i = \sum_{j=1}^{\ell+1-i} b_j^i$ is the sum of row i (it also equals $c_{\ell+1-i}^i$) and of course $\bar{c}^1 = k$. It will be convenient to partition the sum over matrices B according to the vector \bar{c} , so let us denote by $\mathcal{B}(\bar{c})$ be the set of matrices satisfying the given conditions with the row sums \bar{c} :

$$\mathcal{B}(\bar{c}) = \left\{ B = (b_j^i)_{i+j \leq \ell+1} \in \mathbb{N}^{\ell(\ell+1)/2} : \sum_{j=1}^{\ell+1-i} b_j^i = \bar{c}^i \text{ for } 1 \leq i \leq \ell \right. \\ \left. \text{and } \sum_{h=1}^{i-1} b_{\ell+2-i}^h = \bar{c}^i \text{ for } 2 \leq i \leq \ell \right\}. \tag{5.210}$$

The summand is then

$$\left(\prod_{i=1}^{\ell} \binom{t + \bar{c}^i}{\bar{c}^i} \bar{c}^i! \binom{\bar{c}^i}{b_1^i, \dots, b_{\ell+1-i}^i} \right) \left(\prod_{\alpha=1}^{\ell-1} a_\alpha^{\sum_{i=1}^{\ell-\alpha} \sum_{j=1}^{\alpha} b_j^i} \right). \tag{5.211}$$

Furthermore, it follows from Lemma 5.11(b) that

$$\sum_{i=1}^{\ell-\alpha} \sum_{j=1}^{\alpha} b_j^i = k - \bar{c}^{\ell+1-\alpha}. \tag{5.212}$$

Let us now show how to perform the sum over matrices B with a given vector \bar{c} :

Lemma 5.12. *Let $\ell \geq 1, k \geq 0$ and $\bar{c}^1, \dots, \bar{c}^\ell \geq 0$ be integers, with $\bar{c}^1 = k$. Then*

$$\sum_{B=(b_j^i)_{i+j \leq \ell+1} \in \mathcal{B}(\bar{c})} \prod_{i=1}^{\ell} \binom{\bar{c}^i}{b_1^i, \dots, b_{\ell+1-i}^i} = \prod_{i=1}^{\ell} \binom{k}{\bar{c}^i}. \tag{5.213}$$

Before proving Lemma 5.12, let us show how it can be used to complete the proof of Theorem 5.7.

Proof of Theorem 5.7, given Lemma 5.12. Summing (5.211) over $B \in \mathcal{B}(\bar{c})$ and using (5.212) and (5.213) along with $t = -s - k$ and $\bar{c}^1 = k$, we obtain

$$\left(\prod_{i=1}^{\ell} \binom{-s - k + \bar{c}^i}{\bar{c}^i} \bar{c}^{i!} \binom{k}{\bar{c}^i} \right) \left(\prod_{\alpha=1}^{\ell-1} a_{\alpha}^{k - \bar{c}^{\ell+1-\alpha}} \right) = k! \binom{-s}{k} \prod_{\alpha=1}^{\ell-1} \binom{-s - \hat{c}^{\alpha}}{k - \hat{c}^{\alpha}} k! \frac{a_{\alpha}^{\hat{c}^{\alpha}}}{\hat{c}^{\alpha}!} \tag{5.214}$$

where $\hat{c}^{\alpha} = k - \bar{c}^{\ell+1-\alpha}$. The sum over $\bar{c}^2, \dots, \bar{c}^{\ell}$ – or equivalently over $\hat{c}^1, \dots, \hat{c}^{\ell-1}$ – now factorizes and gives precisely (5.190). \square

Proof of Lemma 5.12. The sum on the left-hand side of (5.213) is nontrivial because the “row” and “column” constraints (5.210) are entangled. We shall replace one of the two sets of constraints (say, the “column” one) by a generating function: that is, we introduce indeterminates ξ_j ($1 \leq j \leq \ell$) and consider

$$\Phi_{\ell,k,\bar{c}}(\xi) = \sum_{\substack{\{b_j^i\}_{i+j \leq \ell+1} \\ \sum_{j=1}^{\ell+1-i} b_j^i = \bar{c}^i}} \prod_{j=1}^{\ell} \xi_j^{\sum_{i=1}^{\ell+1-j} b_j^i} \prod_{i=1}^{\ell} \binom{\bar{c}^i}{b_1^i, \dots, b_{\ell+1-i}^i}. \tag{5.215}$$

The sum now factorizes over rows: we have

$$\Phi_{\ell,k,\bar{c}}(\xi) = \prod_{i=1}^{\ell} \sum_{\substack{b_1^i, \dots, b_{\ell+1-i}^i \geq 0 \\ b_1^i + \dots + b_{\ell+1-i}^i = \bar{c}^i}} \left(\prod_{j=1}^{\ell+1-i} \xi_j^{b_j^i} \right) \binom{\bar{c}^i}{b_1^i, \dots, b_{\ell+1-i}^i} \tag{5.216a}$$

$$= \prod_{i=1}^{\ell} \left(\sum_{j=1}^{\ell+1-i} \xi_j \right)^{\bar{c}^i} \tag{5.216b}$$

$$= (\xi_1 + \dots + \xi_{\ell})^k (\xi_1 + \dots + \xi_{\ell-1})^{\bar{c}^2} (\xi_1 + \dots + \xi_{\ell-2})^{\bar{c}^3} \dots \xi_1^{\bar{c}^{\ell}}. \tag{5.216c}$$

We must now extract the coefficient of the monomial

$$\xi_1^k \prod_{j=2}^{\ell} \xi_j^{\bar{c}^{\ell+2-j}} = \xi_1^k \xi_2^{\bar{c}^{\ell}} \xi_3^{\bar{c}^{\ell-1}} \dots \xi_{\ell}^{\bar{c}^2} \tag{5.217}$$

in $\Phi_{\ell,k,\bar{c}}(\xi)$. We first extract $[\xi_{\ell}^{\bar{c}^2}]$ from (5.216c): here ξ_{ℓ} occurs only in the first factor, so we get $\binom{k}{\bar{c}^2} (\xi_1 + \dots + \xi_{\ell-1})^{k-\bar{c}^2}$ times the remaining factors, i.e.

$$\binom{k}{\bar{c}^2} (\xi_1 + \dots + \xi_{\ell-1})^k (\xi_1 + \dots + \xi_{\ell-2})^{\bar{c}^3} \dots \xi_1^{\bar{c}^{\ell}}. \tag{5.218}$$

We can then extract $[\xi_{\ell-1}^{\bar{c}^3}]$ in the same way, and so forth until the end, yielding right-hand side of (5.213). □

Let us conclude by apologizing for the combinatorial complexity involved in the proof of Theorem 5.7. The simplicity of the final formula (5.190), together with the simplicity of (5.213) and the miraculous simplifications observed in its proof, suggest to us that there ought to exist a much simpler and shorter proof of Theorem 5.7. But we have thus far been unable to find one.

6. Proofs of diagonal-parametrized Cayley identities

In this section we prove the diagonal-parametrized Cayley identities (Theorems 2.10 and 2.11). We give two proofs of each result: the first deduces the diagonal-parametrized Cayley identity from the corresponding Cayley identity by a change of variables; the second is a direct proof using a Grassmann representation of the differential operator.

6.1. Diagonal-parametrized ordinary Cayley identity

Proof of Theorem 2.10. We change variables from $(x_{ij})_{i,j=1}^n$ to new variables $(t_i)_{i=1}^n$ and $(y_{ij})_{1 \leq i \neq j \leq n}$ defined by

$$t_i = x_{ii}, \tag{6.1a}$$

$$y_{ij} = x_{ii}^{-\alpha_i} x_{jj}^{-(1-\alpha_j)} x_{ij}. \tag{6.1b}$$

We also set $y_{ii} = 1$ for all i and define the matrices $T_{\alpha} = \text{diag}(t_i^{\alpha_i})$, $T_{1-\alpha} = \text{diag}(t_i^{1-\alpha_i})$ and $Y = (y_{ij})$, so that

$$X = T_{\alpha} Y T_{1-\alpha} \tag{6.2}$$

and hence

$$(\det X)^s = \left(\prod_{i=1}^n t_i^s \right) (\det Y)^s. \tag{6.3}$$

A straightforward computation shows that the differential operators (vector fields) $\partial/\partial x_{ij}$ can be rewritten in the new variables as

$$\frac{\partial}{\partial x_{ij}} = \begin{cases} t_i^{-1} [t_i \frac{\partial}{\partial t_i} - \alpha_i \sum_{k \neq i} y_{ik} \frac{\partial}{\partial y_{ik}} - (1 - \alpha_i) \sum_{l \neq i} y_{li} \frac{\partial}{\partial y_{li}}] & \text{if } i = j, \\ t_i^{-\alpha_i} t_j^{-(1-\alpha_j)} \frac{\partial}{\partial y_{ij}} & \text{if } i \neq j. \end{cases} \tag{6.4}$$

Let us denote by Δ_α the matrix of differential operators whose elements are given by the right-hand side of (6.4); please note that each element commutes with each other element not in the same row or column [that is, $(\Delta_\alpha)_{ij}$ commutes with $(\Delta_\alpha)_{i'j'}$ whenever $i \neq i'$ and $j \neq j'$].

Now (6.3), considered as a function of $t = (t_i)$, is of the form $\text{const} \times \prod_i t_i^{s_i}$; therefore, acting on (6.3), each operator $t_i \partial/\partial t_i$ is equivalent to multiplication by s_i . It follows that the action of Δ_α on (6.3) is identical to that of $\Delta_{\alpha,s}$ defined by

$$(\Delta_{\alpha,s})_{ij} = \begin{cases} t_i^{-1} [s - \alpha_i \sum_{k \neq i} y_{ik} \frac{\partial}{\partial y_{ik}} - (1 - \alpha_i) \sum_{l \neq i} y_{li} \frac{\partial}{\partial y_{li}}] & \text{if } i = j, \\ t_i^{-\alpha_i} t_j^{-(1-\alpha_j)} \frac{\partial}{\partial y_{ij}} & \text{if } i \neq j. \end{cases} \tag{6.5}$$

Furthermore, $(\Delta_\alpha)_{ij}$ and $(\Delta_{\alpha,s})_{ij}$ both commute with $(\Delta_{\alpha,s})_{i'j'}$ whenever $i \neq i'$ and $j \neq j'$.

Now suppose we have a product $(\Delta_\alpha)_{i_1 j_1} \cdots (\Delta_\alpha)_{i_\ell j_\ell}$ acting on (6.3), in which i_1, \dots, i_ℓ are all distinct and also j_1, \dots, j_ℓ are all distinct (this is the case that will arise when we take the determinant of any submatrix). Then the rightmost factor $(\Delta_\alpha)_{i_\ell j_\ell}$ can be replaced by $(\Delta_{\alpha,s})_{i_\ell j_\ell}$; and the same can be done for the other factors by commuting them to the right, changing them from Δ_α to $\Delta_{\alpha,s}$, and commuting them back. It follows that, under the given condition on the indices, $(\Delta_\alpha)_{i_1 j_1} \cdots (\Delta_\alpha)_{i_\ell j_\ell} (\det X)^s = (\Delta_{\alpha,s})_{i_1 j_1} \cdots (\Delta_{\alpha,s})_{i_\ell j_\ell} (\det X)^s$.

Translating back to the original variables $(x_{ij})_{i,j=1}^n$, $\Delta_{\alpha,s}$ equals $D_{\alpha,s}$ defined by

$$(D_{\alpha,s})_{ij} = \begin{cases} x_{ii}^{-1} [s - \alpha_i \sum_{k \neq i} x_{ik} \frac{\partial}{\partial x_{ik}} - (1 - \alpha_i) \sum_{l \neq i} x_{li} \frac{\partial}{\partial x_{li}}] & \text{if } i = j, \\ \frac{\partial}{\partial x_{ij}} & \text{if } i \neq j, \end{cases} \tag{6.6}$$

and the Cayley formula (2.2) tells us that

$$\det((D_{\alpha,s})_{IJ}) (\det X)^s = s(s+1) \cdots (s+k-1) (\det X)^{s-1} \epsilon(I, J) (\det X_{I^c J^c}). \tag{6.7}$$

On the other hand, $D_{\alpha,\beta,s} = \hat{X}_\beta D_{\alpha,s} \hat{X}_{1-\beta}$ where $\hat{X}_\beta = \text{diag}(x_{ii}^{\beta_i})$ and $\hat{X}_{1-\beta} = \text{diag}(x_{ii}^{1-\beta_i})$, so that

$$\det((D_{\alpha,\beta,s})_{IJ}) = \left(\prod_{i \in I} x_{ii}^{\beta_i} \right) \left(\prod_{j \in J} x_{jj}^{1-\beta_j} \right) \det((D_{\alpha,s})_{IJ}). \tag{6.8}$$

Combining (6.7) and (6.8), we obtain (2.27). \square

Now let us show an *ab initio* proof based on Grassmann representation of the differential operator $\det((D_{\alpha,s})_{IJ})$. Since $D_{\alpha,s}$ contains not only terms $\partial/\partial x_{ij}$ but also terms $x_{ij} \partial/\partial x_{ij}$, the role played in Section 5 by the translation formula (5.1) will here be played by the dilation–translation formula (B.60): see Appendix B.4 for discussion.

Alternate proof of Theorem 2.10. For notational simplicity let us assume that the diagonal elements x_{ii} are all equal to 1; the general case can be recovered by a simple scaling.

Consider the matrix of differential operators $D_{\alpha,s}$ defined by (6.6). In terms of the matrices E^{ij} defined by (3.4), we can write

$$D_{\alpha,s} = sI + \sum_{i \neq j} (E^{ij} - x_{ij} [\alpha_i E^{ii} + (1 - \alpha_j) E^{jj}]) \frac{\partial}{\partial x_{ij}}. \tag{6.9}$$

As before, we introduce Grassmann variables $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) and use the representation

$$\det((D_{\alpha,s})_{IJ}) = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\eta}\eta \right)_{I^c, J^c} \times \exp \left[s \sum_i \bar{\eta}_i \eta_i + \sum_{i \neq j} [\bar{\eta}_i \eta_j - x_{ij}(\alpha_i \bar{\eta}_i \eta_i + (1 - \alpha_j) \bar{\eta}_j \eta_j)] \frac{\partial}{\partial x_{ij}} \right]. \tag{6.10}$$

Let us now apply this operator to a generic polynomial $f(X)$: using the dilation–translation formula (B.60) on all the variables x_{ij} ($i \neq j$), we obtain

$$\det((D_{\alpha,s})_{IJ}) f(X) = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\eta}\eta \right)_{I^c, J^c} \exp \left[s \sum_i \bar{\eta}_i \eta_i \right] \times f(\{(1 - \alpha_i \bar{\eta}_i \eta_i) x_{ij} (1 - (1 - \alpha_j) \bar{\eta}_j \eta_j) + \bar{\eta}_i \eta_j\}_{i \neq j}). \tag{6.11}$$

Note that the diagonal terms remain unchanged at their original value $x_{ii} = 1$. Defining the diagonal matrices $M_\alpha = \text{diag}(1 - \alpha_i \bar{\eta}_i \eta_i)$ and $M_{1-\alpha} = \text{diag}[1 - (1 - \alpha_i) \bar{\eta}_i \eta_i]$, we see that the argument of f is

$$X' = M_\alpha X M_{1-\alpha} + \bar{\eta}\eta^T \tag{6.12}$$

(note that also the diagonal elements $i = j$ come out right). We are interested in $f(X) = \det(X)^s$, and we have

$$\det X' = (\det M_\alpha)(\det M_{1-\alpha}) \det(X + M_\alpha^{-1} \bar{\eta}\eta^T M_{1-\alpha}^{-1}) \tag{6.13a}$$

$$= (\det M_\alpha)(\det M_{1-\alpha}) \det(X + \bar{\eta}\eta^T) \tag{6.13b}$$

where $M_\alpha^{-1} \bar{\eta} = \bar{\eta}$ and $\eta^T M_{1-\alpha}^{-1} = \eta^T$ by nilpotency. The factor $(\det M_\alpha)^s (\det M_{1-\alpha})^s$ exactly cancels the factor $\exp[s \sum_i \bar{\eta}_i \eta_i]$ in the integrand, and we are left with

$$\det((D_{\alpha,s})_{IJ}) \det(X)^s = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\eta}\eta \right)_{I^c, J^c} \det(X + \bar{\eta}\eta^T)^s, \tag{6.14}$$

which coincides with (5.13a). Arguing exactly as in (5.13)–(5.15), we therefore obtain

$$\det((D_{\alpha,s})_{IJ}) (\det X)^s = s(s + 1) \cdots (s + k - 1) (\det X)^{s-1} \epsilon(I, J) (\det X)_{I^c, J^c}. \quad \square \tag{6.15}$$

6.2. Diagonal-parametrized symmetric Cayley identity

An analogous proof gives the symmetric analogue:

Proof of Theorem 2.11. We change variables from $(x_{ij})_{1 \leq i \leq j \leq n}$ to new variables $(t_i)_{i=1}^n$ and $(y_{ij})_{1 \leq i < j \leq n}$ defined by

$$t_i = x_{ii}, \tag{6.16a}$$

$$y_{ij} = x_{ii}^{-\frac{1}{2}} x_{jj}^{-\frac{1}{2}} x_{ij}. \tag{6.16b}$$

We also set $y_{ii} = 1$ for all i , use the synonymous $y_{ji} = y_{ij}$ for $j > i$, and define the matrices $T = \text{diag}(t_i^{\frac{1}{2}})$ and $Y = (y_{ij})$, so that

$$X = TYT \tag{6.17}$$

and hence

$$(\det X)^s = \left(\prod_{i=1}^n t_i^s \right) (\det Y)^s. \tag{6.18}$$

A straightforward computation shows that the differential operators (vector fields) $\partial/\partial x_{ij}$ can be rewritten in the new variables as

$$\frac{\partial}{\partial x_{ij}} = \begin{cases} t_i^{-1} [t_i \frac{\partial}{\partial t_i} - \frac{1}{2} \sum_{k \neq i} y_{ik} \frac{\partial}{\partial y_{ik}} - \frac{1}{2} \sum_{l \neq i} y_{li} \frac{\partial}{\partial y_{li}}] & \text{if } i = j, \\ t_i^{-\frac{1}{2}} t_j^{-\frac{1}{2}} \frac{\partial}{\partial y_{ij}} & \text{if } i \neq j. \end{cases} \tag{6.19}$$

Let us denote by Δ the matrix of differential operators whose elements are given by the right-hand side of (6.19); please note that each element commutes with each other element not in the same row or column [that is, Δ_{ij} commutes with $\Delta_{i'j'}$ whenever $i \neq i'$ and $j \neq j'$].

Now (6.18), considered as a function of $t = (t_i)$, is of the form $\text{const} \times \prod_i t_i^s$; therefore, acting on (6.18), each operator $t_i \partial/\partial t_i$ is equivalent to multiplication by s . It follows that the action of Δ is identical to that of Δ_s defined by

$$(\Delta_s)_{ij} = \begin{cases} t_i^{-1} [s - \frac{1}{2} \sum_{k \neq i} y_{ik} \frac{\partial}{\partial y_{ik}} - \frac{1}{2} \sum_{l \neq i} y_{li} \frac{\partial}{\partial y_{li}}] & \text{if } i = j, \\ t_i^{-\frac{1}{2}} t_j^{-\frac{1}{2}} \frac{\partial}{\partial y_{ij}} & \text{if } i \neq j. \end{cases} \tag{6.20}$$

Furthermore, Δ_{ij} and $(\Delta_s)_{ij}$ both commute with $(\Delta_s)_{i'j'}$ whenever $i \neq i'$ and $j \neq j'$.

Now suppose we have a product $\Delta_{i_1 j_1} \cdots \Delta_{i_\ell j_\ell}$ acting on (6.18), in which i_1, \dots, i_ℓ are all distinct and also j_1, \dots, j_ℓ are all distinct. Then the rightmost factor $\Delta_{i_\ell j_\ell}$ can be replaced by $(\Delta_s)_{i_\ell j_\ell}$; and the same can be done for the other factors by commuting them to the right, changing them from Δ to Δ_s , and commuting them back. It follows that, under the given condition on the indices, $\Delta_{i_1 j_1} \cdots \Delta_{i_\ell j_\ell} (\det X)^s = (\Delta_s)_{i_1 j_1} \cdots (\Delta_s)_{i_\ell j_\ell} (\det X)^s$.

Translating back to the original variables $(x_{ij})_{1 \leq i \leq j \leq n}$, Δ_s equals D_s^{sym} defined by

$$(D_s^{\text{sym}})_{ij} = \begin{cases} x_{ii}^{-1} [s - \frac{1}{2} \sum_{k \neq i} x_{ik} \frac{\partial}{\partial x_{ik}} - \frac{1}{2} \sum_{l \neq i} x_{li} \frac{\partial}{\partial x_{li}}] & \text{if } i = j, \\ \frac{\partial}{\partial x_{ij}} & \text{if } i \neq j, \end{cases} \tag{6.21}$$

and the Cayley formula (2.2) tells us that

$$\det((D_s^{\text{sym}})_{IJ}) (\det X^{\text{sym}})^s = s \left(s + \frac{1}{2} \right) \cdots \left(s + \frac{k-1}{2} \right) (\det X^{\text{sym}})^{s-1} \epsilon(I, J) (\det X_{IcJc}^{\text{sym}}). \tag{6.22}$$

On the other hand, $D_{\beta,s}^{\text{sym}} = \hat{X}_\beta D_s^{\text{sym}} \hat{X}_{1-\beta}$ where $\hat{X}_\beta = \text{diag}(x_{ii}^{\beta_i})$ and $\hat{X}_{1-\beta} = \text{diag}(x_{ii}^{1-\beta_i})$, so that

$$\det((D_{\beta,s}^{\text{sym}})_{IJ}) = \left(\prod_{i \in I} x_{ii}^{\beta_i} \right) \left(\prod_{j \in J} x_{jj}^{1-\beta_j} \right) \det((D_s^{\text{sym}})_{IJ}). \tag{6.23}$$

Combining (6.22) and (6.23), we obtain (2.32). \square

Now let us show the proof based on Grassmann representation of the differential operator:

Alternate proof of Theorem 2.11. For notational simplicity let us assume that the diagonal elements x_{ii} are all equal to 1; the general case can be recovered by a simple scaling. We are therefore using the matrix X^{sym} defined by

$$(X^{\text{sym}})_{ij} = \begin{cases} 1 & \text{for } i = j, \\ x_{ij} & \text{for } i < j, \\ x_{ji} & \text{for } i > j. \end{cases} \tag{6.24}$$

Consider the matrix of differential operators $D_{\beta,s}^{\text{sym}}$ defined by (2.30). In terms of the matrices E^{ij} defined by (3.4), we can write

$$D_{\beta,s}^{\text{sym}} = sI + \frac{1}{2} \sum_{i < j} [E^{ij} + E^{ji} - x_{ij}(E^{ii} + E^{jj})] \frac{\partial}{\partial x_{ij}}. \tag{6.25}$$

We introduce Grassmann variables $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) and use the representation

$$\det((D_{\beta,s}^{\text{sym}})_{IJ}) = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\eta} \eta \right)_{I^c, J^c} \times \exp \left[s \sum_i \bar{\eta}_i \eta_i + \frac{1}{2} \sum_{i < j} [\bar{\eta}_i \eta_j - \eta_i \bar{\eta}_j - x_{ij}(\bar{\eta}_i \eta_i + \bar{\eta}_j \eta_j)] \frac{\partial}{\partial x_{ij}} \right]. \tag{6.26}$$

Applying the dilation–translation formula (B.60) on all the variables x_{ij} ($i < j$), we obtain

$$\begin{aligned} & \det((D_{\beta,s}^{\text{sym}})_{IJ}) f(\{x_{ij}\}_{i < j}) \\ &= \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod \bar{\eta} \eta \right)_{I^c, J^c} \exp \left[s \sum_i \bar{\eta}_i \eta_i \right] \\ & \times f \left(\left\{ \left(1 - \frac{1}{2} \bar{\eta}_i \eta_i \right) x_{ij} \left(1 - \frac{1}{2} \bar{\eta}_j \eta_j \right) + \frac{1}{2} (\bar{\eta}_i \eta_j - \eta_i \bar{\eta}_j) \right\}_{i < j} \right). \end{aligned} \tag{6.27}$$

Note that the diagonal terms remain unchanged at their original value $x_{ii} = 1$. Defining the diagonal matrix $M = \text{diag}(1 - \frac{1}{2} \bar{\eta}_i \eta_i)$, we see that the argument of f is

$$X^{\text{sym}'} = M X^{\text{sym}} M + \frac{1}{2} (\bar{\eta} \eta^T - \eta \bar{\eta}^T) \tag{6.28}$$

(note that also the diagonal elements $i = j$ come out right). We are interested in $f(X^{\text{sym}}) = \det(X^{\text{sym}})^s$, and we have

$$\det X^{\text{sym}'} = (\det M)^2 \det \left[X + \frac{1}{2} M^{-1} (\bar{\eta} \eta^T - \eta \bar{\eta}^T) M^{-1} \right] \tag{6.29a}$$

$$= (\det M)^2 \det \left[X + \frac{1}{2} (\bar{\eta} \eta^T - \eta \bar{\eta}^T) \right] \tag{6.29b}$$

where the last equality again follows by nilpotency. The factor $(\det M)^{2s}$ exactly cancels the factor $\exp[s \sum_i \bar{\eta}_i \eta_i]$ in the integrand, and we are left with

$$\det((D_{\beta,s}^{\text{sym}})_{IJ}) \det(X^{\text{sym}})^s = \epsilon(I, J) \int \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod_{I^c, J^c} \bar{\eta} \eta \right) \det \left[X^{\text{sym}} + \frac{1}{2} (\bar{\eta} \eta^T - \eta \bar{\eta}^T) \right]^s, \tag{6.30}$$

which coincides with (5.41). The remainder of the proof is as in (5.41)–(5.48). □

7. Proofs of Laplacian-parametrized Cayley identities

In this section we prove Theorems 2.12 and 2.14. The proofs use a Grassmann representation of the differential operator, and are closely patterned after the proofs of the ordinary and symmetric Cayley identities in Sections 5.1 and 5.3, respectively.

7.1. Laplacian-parametrized ordinary Cayley identity

Let us begin by recalling the definitions of the matrices arising in Theorem 2.12:

$$(X^{\text{row-Lap}})_{ij} = \begin{cases} x_{ij} & \text{if } i \neq j, \\ -\sum_{k \neq i} x_{ik} & \text{if } i = j, \end{cases} \tag{7.1a}$$

$$(\partial^{\text{row-Lap}})_{ij} = \begin{cases} \partial / \partial x_{ij} & \text{if } i \neq j, \\ 0 & \text{if } i = j, \end{cases} \tag{7.1b}$$

and of course $T = \text{diag}(t_i)$. In order to maximize the correspondences with the proof in Section 5.1, it is convenient to prove instead the dual (and of course equivalent) result for column-Laplacian matrices:

$$(X^{\text{col-Lap}})_{ij} = \begin{cases} x_{ij} & \text{if } i \neq j, \\ -\sum_{k \neq i} x_{ki} & \text{if } i = j, \end{cases} \tag{7.2a}$$

$$(\partial^{\text{col-Lap}})_{ij} = \begin{cases} \partial / \partial x_{ij} & \text{if } i \neq j, \\ 0 & \text{if } i = j \end{cases} \tag{7.2b}$$

(note that $\partial^{\text{col-Lap}} = \partial^{\text{row-Lap}}$). In what follows we shall drop the superscripts “col-Lap” in order to lighten the notation, but it is important to remember the definitions (7.2).

Proof of Theorem 2.12. We introduce Grassmann variables $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) and use the representation

$$\det(U + \partial) - \det(\partial) = \int \mathcal{D}_n(\eta, \bar{\eta}) (e^{\bar{\eta}^T U \eta} - 1) e^{\bar{\eta}^T \partial \eta}. \tag{7.3}$$

It is convenient to introduce the Grassmann quantities $\Theta = \sum_{\ell} \eta_{\ell}$ and $\bar{\Theta} = \sum_{\ell} \bar{\eta}_{\ell}$, so that $\bar{\eta}^T U \eta = \bar{\Theta}^T \Theta$. Since $\Theta^2 = \bar{\Theta}^2 = 0$, the first exponential in (7.3) is easily expanded, and we have

$$\det(U + \partial) - \det(\partial) = \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\Theta}^T \Theta) e^{\bar{\eta}^T \partial \eta}. \tag{7.4}$$

Let us now apply (7.4) to $\det(T + X)^s$ where s is a positive integer, using the translation formula (5.1): by (7.2b) we get $x_{ij} \rightarrow x_{ij} + \bar{\eta}_i \eta_j$ for $i \neq j$, and by (7.2a) this induces $X \rightarrow X + \bar{\eta} \eta^T - \text{diag}(\bar{\Theta} \eta_i)$. We therefore obtain

$$\begin{aligned}
 & [\det(U + \vartheta) - \det(\vartheta)] \det(T + X)^s \\
 &= \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\vartheta}^T \vartheta) \det(T + X + \bar{\eta} \eta^T - \text{diag}(\bar{\vartheta} \eta_i))^s.
 \end{aligned} \tag{7.5}$$

But now comes an amazing simplification: because of the prefactor $\bar{\vartheta}^T \vartheta$ and the nilpotency $\vartheta^2 = 0$, all terms in the expansion of the determinant arising from the term $\text{diag}(\bar{\vartheta} \eta_i)$ simply vanish, so we can drop $\text{diag}(\bar{\vartheta} \eta_i)$:

$$[\det(U + \vartheta) - \det(\vartheta)] \det(T + X)^s = \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\vartheta}^T \vartheta) \det(T + X + \bar{\eta} \eta^T). \tag{7.6}$$

Assuming that $T + X$ is an invertible real or complex matrix, we can write this as

$$\begin{aligned}
 & [\det(U + \vartheta) - \det(\vartheta)] \det(T + X)^s \\
 &= \det(T + X)^s \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T U \eta) \det[I + (T + X)^{-1} \bar{\eta} \eta^T]^s.
 \end{aligned} \tag{7.7}$$

Let us now change variables from $(\eta, \bar{\eta})$ to $(\eta', \bar{\eta}') \equiv (\eta, (T + X)^{-1} \bar{\eta})$: we pick up a Jacobian $\det(T + X)^{-1}$, and dropping primes we have

$$\begin{aligned}
 & [\det(U + \vartheta) - \det(\vartheta)] \det(T + X)^s \\
 &= \det(T + X)^{s-1} \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T (T + X^T) U \eta) \det(I + \bar{\eta} \eta^T)^s.
 \end{aligned} \tag{7.8}$$

But since X is column-Laplacian, we have $UX = 0$ and hence $X^T U = 0$, so that $\bar{\eta}^T (T + X^T) U \eta$ reduces to $\bar{\eta}^T T U \eta$.³¹ This expresses the left-hand side of the identity as the desired quantity $\det(T + X)^{s-1}$ times a factor

$$P(s, n, T) \equiv \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T T U \eta) \det(I + \bar{\eta} \eta^T)^s, \tag{7.9}$$

which we now proceed to calculate. The matrix $I + \bar{\eta} \eta^T$ is a rank-1 perturbation of the identity matrix; by Lemma B.11 we have

$$\det(I + \bar{\eta} \eta^T)^s = (1 - \bar{\eta}^T \eta)^{-s} \tag{7.10a}$$

$$= \sum_{\ell=0}^{\infty} (-1)^\ell \binom{-s}{\ell} (\bar{\eta}^T \eta)^\ell. \tag{7.10b}$$

Now

$$\bar{\eta}^T T U \eta = \sum_i t_i \bar{\eta}_i \eta_i + \sum_i \sum_{j \neq i} t_i \bar{\eta}_i \eta_j, \tag{7.11}$$

³¹ For a row-Laplacian matrix we would have instead chosen to multiply $\bar{\eta} \eta^T$ by $(T + X)^{-1}$ on the right, and then made the change of variables from $(\eta, \bar{\eta})$ to $(\eta', \bar{\eta}') \equiv ((T + X^T)^{-1} \eta, \bar{\eta})$, picking up a Jacobian $(\det(T + X^T))^{-1} = (\det(T + X))^{-1}$ and obtaining a prefactor $\bar{\eta}^T U (T + X^T) \eta$; then $UX^T = 0$ because X is row-Laplacian.

but the terms $\bar{\eta}_i \eta_j$ with $i \neq j$ cannot contribute to an integral in which the rest of the integrand depends on $\bar{\eta}$ and η only through products $\bar{\eta}_i \eta_i$; so in the integrand we can replace $\bar{\eta}^T T U \eta$ with $\bar{\eta}^T T \eta$. Since

$$\int \mathcal{D}_n(\eta, \bar{\eta}) \bar{\eta}_i \eta_i (\bar{\eta}^T \eta)^\ell = (n-1)! \delta_{\ell, n-1} \tag{7.12}$$

for each index i , it follows that

$$P(s, n, T) = \left(\sum_i t_i \right) (-1)^{n-1} \binom{-s}{n-1} (n-1)! \tag{7.13a}$$

$$= \left(\sum_i t_i \right) s(s+1) \cdots (s+n-2). \tag{7.13b}$$

This proves (2.36) when X is a real or complex matrix, $\mathbf{t} = (t_i)$ are real or complex values such that $T + X$ is invertible, and s is a positive integer; the general validity of the identity then follows from Proposition 2.18. \square

7.2. Laplacian-parametrized symmetric Cayley identity

Now we prove the corresponding result for symmetric Laplacian matrices. Let us recall the definitions:

$$(X^{\text{sym-Lap}})_{ij} = \begin{cases} x_{ij} & \text{if } i < j, \\ x_{ji} & \text{if } i > j, \\ -\sum_{k \neq i} x_{ik} & \text{if } i = j, \end{cases} \tag{7.14a}$$

$$(\partial^{\text{sym-Lap}})_{ij} = \begin{cases} \partial/\partial x_{ij} & \text{if } i < j, \\ \partial/\partial x_{ji} & \text{if } i > j, \\ 0 & \text{if } i = j. \end{cases} \tag{7.14b}$$

Once again we drop the superscripts “sym-Lap” to lighten the notation.

Proof of Theorem 2.14. Again we introduce Grassmann variables $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) and apply the representation (7.4) to $\det(T + X)^s$ where s is a positive integer. Using the translation formula (5.1) and defining Θ and $\bar{\Theta}$ as before, we obtain

$$\begin{aligned} & [\det(U + \partial) - \det(\partial)] \det(T + X)^s \\ &= \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\Theta} \Theta) \det(T + X + \bar{\eta} \eta^T - \eta \bar{\eta}^T - \text{diag}(\bar{\eta}_i \Theta) - \text{diag}(\bar{\Theta} \eta_i))^s. \end{aligned} \tag{7.15}$$

We again argue that because of the prefactor $\bar{\Theta}^T \Theta$ and the nilpotencies $\Theta^2 = \bar{\Theta}^2 = 0$, all terms in the expansion of the determinant arising from the terms $\text{diag}(\bar{\eta}_i \Theta)$ and $\text{diag}(\bar{\Theta} \eta_i)$ simply vanish, so we can drop these two terms:

$$\begin{aligned} & [\det(U + \partial) - \det(\partial)] \det(T + X)^s \\ &= \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\Theta} \Theta) \det(T + X + \bar{\eta} \eta^T - \eta \bar{\eta}^T)^s. \end{aligned} \tag{7.16}$$

Assuming that $T + X$ is an invertible real or complex matrix, we can write this as

$$\begin{aligned}
 & [\det(U + \partial) - \det(\partial)] \det(T + X)^s \\
 &= \det(T + X)^s \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T U \eta) \det[I + (T + X)^{-1} (\bar{\eta} \eta^T - \eta \bar{\eta}^T)]^s. \tag{7.17}
 \end{aligned}$$

Let us now change variables from $(\eta, \bar{\eta})$ to $(\eta', \bar{\eta}') \equiv (\eta, (T + X)^{-1} \bar{\eta})$; we pick up a Jacobian $(\det(T + X))^{-1}$, and dropping primes we have

$$\begin{aligned}
 & [\det(U + \partial) - \det(\partial)] \det(T + X)^s \\
 &= \det(T + X)^{s-1} \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T (T + X^T) U \eta) \\
 &\quad \times \det[I + \bar{\eta} \eta^T - (T + X)^{-1} \eta \bar{\eta}^T (T + X)^T]^s. \tag{7.18}
 \end{aligned}$$

But since X is symmetric Laplacian (hence column-Laplacian), we have $X^T U = 0$, and the prefactor reduces to $\bar{\eta}^T T U \eta$. Now we apply Corollary B.12 to the determinant expression, and obtain

$$\begin{aligned}
 & [\det(U + \partial) - \det(\partial)] \det(T + X)^s \\
 &= \det(T + X)^{s-1} \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T T U \eta) (1 - \bar{\eta}^T \eta)^{-2s}. \tag{7.19}
 \end{aligned}$$

This expresses the left-hand side of the identity as the desired quantity $\det(T + X)^{s-1}$ times a factor

$$P(s, n, T) \equiv \int \mathcal{D}_n(\eta, \bar{\eta}) (\bar{\eta}^T T U \eta) (1 - \bar{\eta}^T \eta)^{-2s}, \tag{7.20}$$

which we now proceed to calculate. As in the row-Laplacian case, the terms in $\bar{\eta}^T (T U) \eta$ of the form $\bar{\eta}_i \eta_j$ with $i \neq j$ cannot contribute to the integral, as the rest of the integrand depends on η and $\bar{\eta}$ only through products $\bar{\eta}_i \eta_i$, so we can replace $\bar{\eta}^T (T U) \eta$ with $\bar{\eta}^T T \eta$. Since

$$\int \mathcal{D}_n(\eta, \bar{\eta}) \bar{\eta}_i \eta_i (\bar{\eta}^T \eta)^\ell = (n - 1)! \delta_{\ell, n-1} \tag{7.21}$$

for each i , it follows that

$$P(s, n, T) = \left(\sum_i t_i \right) (-1)^{n-1} \binom{-2s}{n-1} (n-1)! \tag{7.22a}$$

$$= \left(\sum_i t_i \right) 2s(2s+1) \cdots (2s+n-2). \tag{7.22b}$$

This proves (2.40) when X is a symmetric real or complex matrix, $\mathbf{t} = (t_i)$ are real or complex values such that $T + X$ is invertible, and s is a positive integer; the general validity of the identity then follows from Proposition 2.18. \square

8. Proofs of product-parametrized and border-parametrized rectangular Cayley identities

In this section we prove the product-parametrized and border-parametrized rectangular Cayley identities (Theorems 2.16 and 2.17) and then discuss the close relationship between them.

8.1. Product-parametrized rectangular Cayley identity

Before beginning the proof of Theorem 2.16, let us observe that the quantity $\det(MA)$ appearing in the statement of the theorem [cf. (2.44) for the definition of the matrix M in terms of X and B] has an alternate expression as follows:

Lemma 8.1. *Let A, B, X and M be as in Theorem 2.16. Then*

$$\det(MA) = \sum_{\substack{L \subseteq [m] \\ |L|=k}} \epsilon(I, L) (\det(B^T A)_{JL}) (\det(XA)_{I^c L^c}). \tag{8.1}$$

Proof. The definition (2.44) of M can be rewritten as

$$M_{I^c \star} = X_{I^c \star}, \tag{8.2a}$$

$$M_{I \star} = (B^T)_{J \star}. \tag{8.2b}$$

Therefore

$$(MA)_{I^c \star} = (XA)_{I^c \star}, \tag{8.3a}$$

$$(MA)_{I \star} = (B^T A)_{J \star} \tag{8.3b}$$

for any matrix A . We now apply multi-row Laplace expansion (A.8) with row set I and summation variable L ; this yields (8.1). \square

We are now ready to prove Theorem 2.16. In order to bring out the ideas behind the proof as clearly as possible, we will first fully develop the reasoning proving the “basic” identity (2.42) – which is actually quite simple – and then describe the modifications needed to handle the all-minors case.

Proof of Theorem 2.16. Let us apply Corollary B.6 to the determinant $\det(\partial B)$ and then introduce a Grassmann representation for the resulting block determinant: we obtain

$$\det(\partial B) = \det \left(\begin{array}{c|c} 0_m & \partial \\ \hline -B & I_n \end{array} \right) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \exp[\bar{\psi}^T \partial \eta - \bar{\eta}^T B \psi + \bar{\eta}^T \eta] \tag{8.4}$$

where $\psi_i, \bar{\psi}_i$ ($1 \leq i \leq m$) and $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n$) are Grassmann variables. By the translation formula (5.1), we have

$$\det(\partial B) f(X) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \exp[\bar{\eta}^T (-B\psi + \eta)] f(X + \bar{\psi} \eta^T) \tag{8.5}$$

for an arbitrary polynomial f . We shall use this formula in the case $f(X) = \det(XA)^s$ where s is a positive integer, so that

$$\det(\partial B) \det(XA)^s = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \exp[\bar{\eta}^T (-B\psi + \eta)] \det[(X + \bar{\psi} \eta^T) A]^s. \tag{8.6}$$

It is convenient to introduce the shorthand

$$X^{\text{trans}} \equiv X + \bar{\psi} \eta^T. \tag{8.7}$$

Suppose now that XA is an invertible real or complex matrix. Then we have

$$X^{\text{trans}} A \equiv (X + \bar{\psi} \eta^T) A = (XA) [I_m + ((XA)^{-1} \bar{\psi})(\eta^T A)]. \tag{8.8}$$

Let us now change variables from $(\psi, \bar{\psi}, \eta, \bar{\eta})$ to $(\psi', \bar{\psi}', \eta', \bar{\eta}') \equiv (\psi, (XA)^{-1} \bar{\psi}, \eta, \bar{\eta})$, with Jacobian $\det(XA)^{-1} = (\det XA)^{-1}$. In the new variables we have (dropping now the primes from the notation)

$$X^{\text{trans}} A = (XA)(I_m + \bar{\psi} \eta^T A), \tag{8.9}$$

and the translated determinant is given by

$$\det(X^{\text{trans}} A) = \det(XA) \det(I_m + \bar{\psi} \eta^T A), \tag{8.10}$$

so that

$$\det(\partial B) \det(XA)^s = \det(XA)^{s-1} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) e^{-\bar{\eta}^T B \psi + \bar{\eta}^T \eta} \det(I + \bar{\psi} \eta^T A)^s. \tag{8.11}$$

Applying Lemma B.11 to the rightmost determinant yields

$$\det(I + \bar{\psi} \eta^T A) = (1 - \bar{\psi}^T A^T \eta)^{-1}, \tag{8.12}$$

so that we are left with the Grassmann-integral expression

$$\det(\partial B) \det(XA)^s = \det(XA)^{s-1} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) e^{-\bar{\eta}^T B \psi + \bar{\eta}^T \eta} (1 - \bar{\psi}^T A^T \eta)^{-s}. \tag{8.13}$$

We have therefore proven that $\det(\partial B) \det(XA)^s$ equals the desired quantity $\det(XA)^{s-1}$ multiplied by a factor

$$b(s, A, B) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) e^{-\bar{\eta}^T B \psi + \bar{\eta}^T \eta} (1 - \bar{\psi}^T A^T \eta)^{-s} \tag{8.14}$$

that does not involve the variables X (but still involves the parameters A and B). Now, in the expansion of

$$(1 - \bar{\psi}^T A^T \eta)^{-s} = \sum_{k=0}^{\infty} (-1)^k \binom{-s}{k} (\bar{\psi}^T A^T \eta)^k, \tag{8.15}$$

only the term $k = m$ survives the integration over the variables $\bar{\psi}$, so we can replace $(1 - \bar{\psi}^T A^T \eta)^{-s}$ in the integrand of (8.14) by $(-1)^m \binom{-s}{m} (\bar{\psi}^T A^T \eta)^m$. Moreover, the same reasoning shows that we can replace $(\bar{\psi}^T A^T \eta)^m$ by $m! \exp(\bar{\psi}^T A^T \eta)$. We are therefore left with a combinatorial prefactor

$$(-1)^m \binom{-s}{m} m! = s(s+1) \cdots (s+m-1) \tag{8.16}$$

multiplying the Grassmann integral

$$\int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \exp \left[\begin{pmatrix} \bar{\psi} \\ \bar{\eta} \end{pmatrix}^T \begin{pmatrix} 0_m & A^T \\ -B & I_n \end{pmatrix} \begin{pmatrix} \psi \\ \eta \end{pmatrix} \right] = \det \begin{pmatrix} 0_m & A^T \\ -B & I_n \end{pmatrix}, \tag{8.17}$$

which equals $\det(A^T B)$ by Corollary B.6.

This proves the “basic” identity (2.42) whenever XA is an invertible real or complex matrix and s is a positive integer. Now, if A has rank $< m$, then both sides of (2.42) are identically zero; while if A has rank m , then XA is invertible for a nonempty open set of matrices X . The general validity of the identity (2.42) therefore follows from Proposition 2.18.

Now let us consider the modifications needed to prove the all-minors identity (2.43). For a while these modifications will run along the same lines as those used in the proof of the two-matrix rectangular Cayley identity (Section 5.5). Thus, a factor $\epsilon(I, J)(\prod \bar{\psi} \psi)_{I^c, J^c}$ gets inserted into the Grassmann integral (8.4)–(8.6); after the change of variables (and dropping of primes) it becomes $\epsilon(I, J)(\prod (XA\bar{\psi})\psi)_{I^c, J^c}$. So we have, in place of Eq. (8.13), the modified expression

$$\det[(\partial B)_{IJ}] \det(XA)^s = \epsilon(I, J) \det(XA)^{s-1} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod (XA\bar{\psi})\psi \right)_{I^c, J^c} \times \exp[\bar{\eta}^T \eta - \bar{\eta}^T B \psi] (1 - \bar{\psi}^T A^T \eta)^{-s}. \tag{8.18}$$

Once again we argue that the integration over variables $\bar{\psi}$ allows to replace

$$(1 - \bar{\psi}^T A^T \eta)^{-s} \rightarrow (-1)^k \binom{-s}{k} k! \exp[\bar{\psi}^T A^T \eta] \tag{8.19}$$

since, in both cases, only the k -th term of the expansion survives. We therefore have

$$\det[(\partial B)_{IJ}] \det(XA)^s = \epsilon(I, J) \det(XA)^{s-1} s(s+1) \cdots (s+k-1) \times \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \left(\prod (\bar{\psi} A^T X^T) \psi \right)_{I^c, J^c} \exp[\eta^T \eta - \eta^T B \psi + \bar{\psi}^T A^T \eta]. \tag{8.20}$$

Now we perform the integration over η and $\bar{\eta}$ using Wick’s theorem for “complex” fermions in the “source” form [cf. (A.94)], yielding

$$\det[(\partial B)_{IJ}] \det(XA)^s = \epsilon(I, J) \det(XA)^{s-1} s(s+1) \cdots (s+k-1) \times \int \mathcal{D}_m(\psi, \bar{\psi}) \left(\prod (\bar{\psi} A^T X^T) \psi \right)_{I^c, J^c} \exp[\bar{\psi}^T A^T B \psi]. \tag{8.21}$$

Next we perform the integration over ψ and $\bar{\psi}$ using Wick’s theorem for “complex” fermions in the “correlation function” form (A.97), yielding³²

³² Here we have made in (A.97) the substitutions $A \rightarrow A^T B$, $B \rightarrow (I)_{J^c, *}$, $C \rightarrow (A^T X^T)_{*I^c}$, $I \rightarrow K$, $J \rightarrow L$.

$$\begin{aligned}
 & \int \mathcal{D}_m(\psi, \bar{\psi}) \left(\prod (\bar{\psi} A^T X^T) \psi \right)_{J^c, J^c} \exp[\bar{\psi}^T A^T B \psi] \\
 &= \sum_{|K|=|L|=n-k} \epsilon(K, L) (\det I_{J^c L}) (\det(A^T B)_{K^c L^c}) (\det(A^T X^T)_{K L^c}) \\
 &= \sum_{|K|=n-k} \epsilon(K, J^c) (\det(A^T B)_{K^c J}) (\det(A^T X^T)_{K L^c}) \\
 &= \sum_{|K|=k} \epsilon(K^c, J^c) (\det(A^T B)_{K J}) (\det(A^T X^T)_{K^c J^c}). \tag{8.22}
 \end{aligned}$$

We now use $\epsilon(K^c, J^c) = \epsilon(J, K)$ and $\epsilon(I, J)\epsilon(J, K) = \epsilon(I, K)$; it follows that $\epsilon(I, J)$ times (8.22) equals (8.1). \square

Remark. The $s = 1$ special case of the all-minors identity (2.43) has an easy elementary proof, which actually proves a stronger result. Note first that by the multilinearity of the determinant $\det(XA)$ in the variables $\{x_{ij}\}$, we have

$$\det(MA) = \left(\prod_p (\partial B)_{i_p j_p} \right) \det(XA). \tag{8.23}$$

Moreover, because permuting the indices j_1, \dots, j_k amounts to permuting the rows of B^T and hence permuting a subset of the rows of M , we have

$$\det(MA) = \text{sgn}(\sigma) \left(\prod_p (\partial B)_{i_p j_{\sigma(p)}} \right) \det(XA) \tag{8.24}$$

for any permutation $\sigma \in \mathcal{S}_k$. Summing this over $\sigma \in \mathcal{S}_k$, we obtain

$$k! \det(MA) = [\det(\partial B)_{I J}] \det(XA), \tag{8.25}$$

which is nothing other than the $s = 1$ case of (2.43). But it is amusing to note that (8.24) holds for each $\sigma \in \mathcal{S}_k$, not just when summed over $\sigma \in \mathcal{S}_k$.

8.2. Border-parametrized rectangular Cayley identity

Proof of Theorem 2.17. We introduce Grassmann variables $\psi_i, \bar{\psi}_i$ ($1 \leq i \leq m$) and $\eta_i, \bar{\eta}_i$ ($1 \leq i \leq n - m$). We use the representation

$$\det(\hat{\partial}) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_{n-m}(\eta, \bar{\eta}) \exp \left[\bar{\psi}^T \partial \begin{pmatrix} \psi \\ \eta \end{pmatrix} + \bar{\eta}^T B \begin{pmatrix} \psi \\ \eta \end{pmatrix} \right]. \tag{8.26}$$

By the translation formula (5.1), we have

$$\det(\hat{\partial}) f(X) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_{n-m}(\eta, \bar{\eta}) \exp \left[\bar{\eta}^T B \begin{pmatrix} \psi \\ \eta \end{pmatrix} \right] f \left(X + \bar{\psi} \begin{pmatrix} \psi \\ \eta \end{pmatrix}^T \right) \tag{8.27}$$

for an arbitrary polynomial f . We shall use this formula in the case $f(X) = \det(\hat{X})^s$ where s is a positive integer, so that

$$\det(\hat{\partial}) \det(\hat{X})^s = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_{n-m}(\eta, \bar{\eta}) \exp\left[\bar{\eta}^T B \begin{pmatrix} \psi \\ \eta \end{pmatrix}\right] \det\left[\hat{X} + \begin{pmatrix} \bar{\psi} \\ 0 \end{pmatrix} \begin{pmatrix} \psi \\ \eta \end{pmatrix}^T\right]^s. \tag{8.28}$$

It is convenient to introduce the shorthand

$$(\hat{X})^{\text{trans}} \equiv \hat{X} + \begin{pmatrix} \bar{\psi} \\ 0 \end{pmatrix} \begin{pmatrix} \psi \\ \eta \end{pmatrix}^T \tag{8.29}$$

for the argument of \det .

Let us now assume that \hat{X} is an invertible real or complex matrix, and change variables from $\begin{pmatrix} \psi \\ \eta \end{pmatrix}$ to $\begin{pmatrix} \psi' \\ \eta' \end{pmatrix} = \hat{X}^{-T} \begin{pmatrix} \psi \\ \eta \end{pmatrix}$ with Jacobian $(\det \hat{X})^{-1}$. Dropping primes from the new variables, we observe that the expression for the translated matrix can be written as

$$(\hat{X})^{\text{trans}} = \hat{X} \left[I + \begin{pmatrix} \bar{\psi} \\ 0 \end{pmatrix} \begin{pmatrix} \psi \\ \eta \end{pmatrix}^T \right], \tag{8.30}$$

so that

$$\det(\hat{X})^{\text{trans}} = (\det \hat{X}) \det\left[I + \begin{pmatrix} \bar{\psi} \\ 0 \end{pmatrix} \begin{pmatrix} \psi \\ \eta \end{pmatrix}^T \right]. \tag{8.31}$$

Applying Lemma B.11 to the rightmost determinant yields

$$\det\left[I + \begin{pmatrix} \bar{\psi} \\ 0 \end{pmatrix} \begin{pmatrix} \psi \\ \eta \end{pmatrix}^T \right] = (1 - \bar{\psi}^T \psi)^{-1}, \tag{8.32}$$

so that we are left with the Grassmann-integral expression

$$\begin{aligned} \det(\hat{\partial}) \det(\hat{X})^s &= \det(\hat{X})^{s-1} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_{n-m}(\eta, \bar{\eta}) \\ &\quad \times \exp\left[\bar{\eta}^T B \hat{X}^T \begin{pmatrix} \psi \\ \eta \end{pmatrix}\right] (1 - \bar{\psi}^T \psi)^{-s} \end{aligned} \tag{8.33a}$$

$$\begin{aligned} &= \det(\hat{X})^{s-1} \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_{n-m}(\eta, \bar{\eta}) \\ &\quad \times \exp[\bar{\eta}^T B X^T \psi + \bar{\eta}^T B A^T \eta] (1 - \bar{\psi}^T \psi)^{-s}. \end{aligned} \tag{8.33b}$$

As the integrand depends on the Grassmann variables only through combinations of the form $\bar{\psi}_i \psi_j$, $\bar{\eta}_i \eta_j$ and $\bar{\eta}_i \psi_j$ (i.e. there is no $\bar{\psi}_i \eta_j$), we can drop all the terms $\bar{\eta}_i \psi_j$, as these terms would certainly remain unpaired in the expansion. This removes all the dependence on X in the integrand, and proves that $\det(\hat{\partial})$ is a Bernstein–Sato operator for $\det(\hat{X})$. We are left with the determination of the prefactor $b(s)$, which is given by

$$b(s) = \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_{n-m}(\eta, \bar{\eta}) \exp[\bar{\eta}^T B A^T \eta] (1 - \bar{\psi}^T \psi)^{-s}. \tag{8.34}$$

Integration over η and $\bar{\eta}$ gives a factor $\det(BA^T) = \det(AB^T)$, while the integration over ψ and $\bar{\psi}$ is identical to the one performed in the case of the ordinary Cayley identity [cf. (5.7)–(5.10)] and gives $s(s + 1) \cdots (s + m - 1)$. This proves (2.46) whenever \hat{X} is an invertible real or complex matrix, B is an arbitrary real or complex matrix, and s is a positive integer. Now, if A has rank $< n - m$, then both sides of (2.46) are identically zero; while if A has rank $n - m$, then $\hat{X} = \begin{pmatrix} X \\ A \end{pmatrix}$ is invertible for a nonempty open set of matrices X . The general validity of the identity therefore follows from Proposition 2.18. \square

8.3. Relation between product-parametrized and border-parametrized identities

Let us begin by recalling the product-parametrized Cayley identity (2.42) and the border-parametrized Cayley identity (2.46), writing the matrices A and B occurring in them as $A^{(0)}, B^{(0)}$ in the former identity and $A^{(1)}, B^{(1)}$ in the latter:

$$(2.42): \det(\partial B^{(0)})[\det(XA^{(0)})]^s = b(s) \det(A^{(0)T}B^{(0)})[\det(XA^{(0)})]^{s-1}, \tag{8.35}$$

$$(2.46): \det\left(\frac{\partial}{B^{(1)}}\right) \det\left(\frac{X}{A^{(1)}}\right)^s = b(s) \det(A^{(1)}B^{(1)T}) \det\left(\frac{X}{A^{(1)}}\right)^{s-1} \tag{8.36}$$

where $b(s) = s(s + 1) \cdots (s + m - 1)$. Here X is an $m \times n$ matrix, while $A^{(0)}$ and $B^{(0)}$ are $n \times m$ matrices, and $A^{(1)}$ and $B^{(1)}$ are $(n - m) \times n$ matrices. Note that $A^{(0)}$ and $B^{(0)}$ must have full rank m , otherwise (8.35) is identically zero; likewise, $A^{(1)}$ and $B^{(1)}$ must have full rank $n - m$.

We will construct the matrices $A^{(0)}$ and $A^{(1)}$ out of a larger $(n \times n)$ matrix A , as follows: Let A be an invertible $n \times n$ matrix, and define

$$A^{(0)} = A_{\star, [m]} = \text{first } m \text{ columns of } A, \tag{8.37a}$$

$$A^{(1)} = (A^{-1})_{[m]^c, \star} = \text{last } n - m \text{ rows of } A^{-1}. \tag{8.37b}$$

Likewise, let B be an invertible $n \times n$ matrix, and define $B^{(0)}$ and $B^{(1)}$ by the same procedure. We then have the following facts:

Lemma 8.2. *Let X and Y be $m \times n$ matrices, with $m \leq n$; let A and B be invertible $n \times n$ matrices; and define matrices $A^{(0)}, A^{(1)}, B^{(0)}, B^{(1)}$ as above. Then:*

- (a) $\det(XA^{(0)}) = (\det A) \det\left(\frac{X}{A^{(1)}}\right)$.
- (b) $\det(YB^{(0)}) = (\det B) \det\left(\frac{Y}{B^{(1)T}}\right)$.
- (c) $\det(A^{(0)T}B^{(0)}) = (\det A)(\det B) \det(A^{(1)}B^{(1)T})$.

Proof. (a) First we expand the left-hand side using the Cauchy–Binet identity (A.3):

$$\det(XA^{(0)}) = \sum_{|L|=m} (\det X_{\star L}) (\det A_{L\star}^{(0)}) = \sum_{|L|=m} (\det X_{\star L}) (\det A_{L, [m]}). \tag{8.38}$$

Next we expand the right-hand side using multi-row Laplace expansion (A.8) with row set $[m]$, followed by the Jacobi identity (A.7):

$$\begin{aligned} \det\left(\frac{X}{A^{(1)}}\right) &= \sum_{|L|=m} \epsilon(L)(\det X_{\star L})(\det A_{\star L}^{(1)}) \\ &= \sum_{|L|=m} \epsilon(L)(\det X_{\star L})(\det A_{L^c, [m]^c}^{-T}) \\ &= \sum_{|L|=m} \epsilon(L)(\det X_{\star L})\epsilon(L)(\det A)^{-1}(\det A_{L, [m]}). \end{aligned} \tag{8.39}$$

Comparing (8.38) and (8.39) proves (a); and (b) is of course identical.

(c) First we expand the left-hand side using Cauchy–Binet:

$$\det(A^{(0)T}B^{(0)}) = \sum_{|L|=m} (\det(A^{(0)T})_{\star L})(\det B_{L\star}^{(0)}) = \sum_{|L|=m} (\det A_{L, [m]})(\det B_{L, [m]}). \tag{8.40}$$

Next we expand the right-hand side using Cauchy–Binet and then using the Jacobi identity twice:

$$\begin{aligned} \det(A^{(1)}B^{(1)T}) &= \sum_{|K|=n-m} (\det A_{\star K}^{(1)})(\det(B^{(1)T})_{K\star}) \\ &= \sum_{|K|=n-m} (\det(A^{-T})_{K, [m]^c})(\det(B^{-T})_{K, [m]^c}) \\ &= \sum_{|K|=n-m} \epsilon(K)(\det A)^{-1}(\det A_{K^c, [m]})\epsilon(K)(\det B)^{-1}(\det B_{K^c, [m]}) \\ &= (\det A)^{-1}(\det B)^{-1} \sum_{|L|=m} (\det A_{L, [m]})(\det B_{L, [m]}). \end{aligned} \tag{8.41}$$

Comparing (8.40) and (8.41) proves (c). □

Using Lemma 8.2, we see immediately the equivalence of (8.35) and (8.36) whenever $A^{(0)}$ and $A^{(1)}$ are related by (8.37) and likewise for $B^{(0)}$ and $B^{(1)}$.

On the other hand, given any $n \times m$ matrix $A^{(0)}$ of rank m , it can be obviously completed to yield a nonsingular matrix A (which is invertible at least when the matrix elements take values in a field). Likewise, an $(n - m) \times n$ matrix $A^{(1)}$ of rank $n - m$ can be completed to yield a nonsingular matrix A^{-1} . So to each $A^{(0)}$ there corresponds a nonempty set of matrices $A^{(1)}$, and vice versa.

Remark. It is *not* in general true that every pair $(A^{(0)}, A^{(1)})$ arises from a matrix A . Consider, for instance, $m = 1$ and $n = 2$: an easy calculation shows that for arbitrary A we must have $A^{(1)}A^{(0)} = 0$.

9. Conjectures on minimality

Let us recall that any pair $Q(s, x, \partial/\partial x)$ and $b(s) \neq 0$ satisfying

$$Q(s, x, \partial/\partial x)P(x)^s = b(s)P(x)^{s-1} \tag{9.1}$$

is called a Bernstein–Sato pair for the polynomial $P(x)$. The minimal (with respect to factorization) monic polynomial $b(s)$ for which there exists such a Q is called the Bernstein–Sato polynomial (or b -function) of P . Our Cayley-type identities thus provide Bernstein–Sato pairs for certain polynomials arising from determinants. But are our polynomials $b(s)$ minimal?

For the ordinary Cayley identities (Theorems 2.1–2.9), it follows from the general theory of prehomogeneous vector spaces [52,60] that the polynomials $b(s)$ found here are indeed minimal, i.e. that the correct b -functions are

$$b(s) = \begin{cases} s(s+1) \cdots (s+n-1) & \text{for an } n \times n \text{ matrix,} \\ s(s+\frac{1}{2}) \cdots (s+\frac{n-1}{2}) & \text{for an } n \times n \text{ symmetric matrix,} \\ s(s+2) \cdots (s+2m-2) & \text{for a } 2m \times 2m \text{ antisymmetric matrix (pfaffian),} \\ (s-\frac{1}{2})s(s+\frac{1}{2}) \cdots (s+m-1) & \text{for a } 2m \times 2m \text{ antisymmetric matrix (determinant),} \\ s(s+1) \cdots (s+m-1)(s+n-m) \cdots (s+n-1) & \text{for a pair of } m \times n \text{ rectangular matrices,} \\ s(s+\frac{1}{2}) \cdots (s+\frac{m-1}{2})(s+\frac{n-m-1}{2}) \cdots (s+\frac{n-2}{2}) & \text{for an } m \times n \text{ rectangular matrix (symmetric),} \\ s(s+2) \cdots (s+2m-2)(s+2n-2m+1) \cdots (s+2n-1) & \text{for an } m \times n \text{ rectangular matrix (antisymmetric),} \\ \prod_{\alpha=1}^{\ell} \prod_{j=0}^{n_{\alpha}-1} (s+n_{\alpha}-n_1+j) & \text{for } \ell \text{ matrices of sizes } n_{\alpha} \times n_{\alpha+1}. \end{cases} \tag{9.2}$$

Indeed, the polynomials P occurring in these identities all correspond to relative invariants of prehomogeneous vector spaces: the ordinary, symmetric and antisymmetric Cayley identities [(2.1), (2.4) and (2.7)] correspond to cases (1), (2) and (3), respectively, in Kimura’s [60, Appendix] table of the irreducible reduced prehomogeneous vector spaces; the one-matrix rectangular symmetric and antisymmetric Cayley identities [(2.18) and (2.21)] correspond to cases (15) and (13) in the same table; while the two-matrix and multi-matrix rectangular Cayley identities [(2.16) and (2.23)] correspond to prehomogeneous vector spaces associated to equioriented quivers of type A [98]. Whenever P is a relative invariant of a prehomogeneous vector space, the general theory [52,60] allows the immediate identification of a suitable operator $Q(\partial/\partial x)$ – namely, the dual of P itself – and provides a proof that the corresponding $b(s)$ satisfies $\deg b = \deg P$ and is indeed (up to a constant factor) the Bernstein–Sato polynomial of P .³³

For the Laplacian-parametrized identities (Theorems 2.12 and 2.14), we conjecture that the polynomials $b(s)$ found here are also minimal, i.e. that the correct b -functions are

$$b(s) = \begin{cases} s(s+1) \cdots (s+n-2) & \text{for a Laplacian-parametrized } n \times n \text{ matrix,} \\ s(s+\frac{1}{2}) \cdots (s+\frac{n-2}{2}) & \text{for a Laplacian-parametrized } n \times n \text{ symmetric matrix.} \end{cases} \tag{9.3}$$

Perhaps these identities can also be interpreted within the framework of prehomogeneous vector spaces; or perhaps an alternate proof of minimality can be found.

It is even conceivable that the following general fact about Bernstein–Sato polynomials is true:

Conjecture 9.1. *Let $P(x_1, \dots, x_n) \not\equiv 0$ be a homogeneous polynomial in n variables with coefficients in a field K of characteristic 0, and let $b(s)$ be its Bernstein–Sato polynomial. Then $\deg b \geq \deg P$.*

Simple examples show that we need not have $\deg b \geq \deg P$ if P is not homogeneous: for instance, $P(x) = 1 - x^2$ has $b(s) = s$. Moreover, slightly more complicated examples show that one

³³ See [52, Corollary 6.1.1 and Theorem 6.1.1], [60, Proposition 2.22] for the first two points, and [52, Theorem 6.3.2] for the third. We are grateful to Nero Budur for explaining to us the connection between our results and the theory of prehomogeneous vector spaces, and in particular for pointing out that this connection provides a general proof of minimality.

can have $\deg b > \deg P$ even when P is homogeneous: for instance, the Bernstein–Sato polynomial of $P(x_1, x_2) = x_1 x_2 (x_1 + ax_2)$ with $a \neq 0$ is (in our “shifted” notation) $s^2(s - \frac{1}{3})(s + \frac{1}{3})$ [109, Corollary 4.14 and Remark 4.15], [84, 5.4].³⁴ However, no one that we have consulted seems to have any counterexample to Conjecture 9.1.

Let us remark that a *necessary* condition for a polynomial $b(s)$ to be a Bernstein–Sato polynomial is that its roots should be *rational* numbers < 1 : this is the content of a famous theorem of Kashiwara [56], [9, Chapter 6], [66, Proposition 2.11].³⁵ Our polynomials (9.2) and (9.3) satisfy this condition.

For the diagonal-parametrized Cayley identities (Theorems 2.10 and 2.11), a slightly more complicated situation arises. The polynomials $b(s)$ arising from the basic case $I = J = [n]$ of those theorems, namely

$$b(s) = \begin{cases} s(s+1) \cdots (s+n-1) & \text{for an } n \times n \text{ matrix,} \\ s(s + \frac{1}{2}) \cdots (s + \frac{n-1}{2}) & \text{for an } n \times n \text{ symmetric matrix} \end{cases} \quad (9.4)$$

are definitely *not* minimal. Indeed, as remarked already in Section 2.3, a lower-order Bernstein–Sato pair can be obtained by taking $I = J = [n] \setminus \{i_0\}$ for any fixed $i_0 \in [n]$:

$$b(s) = \begin{cases} s(s+1) \cdots (s+n-2) & \text{for an } n \times n \text{ matrix,} \\ s(s + \frac{1}{2}) \cdots (s + \frac{n-2}{2}) & \text{for an } n \times n \text{ symmetric matrix.} \end{cases} \quad (9.5)$$

We conjecture that these latter polynomials are indeed minimal, but we have no proof for general n .

It is curious that these polynomials are the *same* as we get for the Laplacian-parametrized Cayley identities. Furthermore, also the Q operators corresponding to these (conjecturally minimal) polynomials $b(s)$ of degree $n-1$ are somewhat similar in one respect: namely, for the diagonal-parametrized case the Q operator is given by any principal minor of size $n-1$ (i.e. $I = J = [n] \setminus \{i_0\}$) of the relevant matrix $D_{\alpha, \beta, s}$ of differential operators, while for the Laplacian-parametrized case the Q operator is a polynomial $\det(U + \partial) - \det(\partial)$ that is a *sum* over all minors (not necessarily principal) of size $n-1$. We do not know whether this resemblance is indicative of any deeper connection between these identities.

10. Note added

After this paper appeared in preprint form, Nero Budur (private communication, July 2011) informed us that he has a proof of our minimality conjectures for the Laplacian-parametrized and diagonal-parametrized Cayley identities.

Acknowledgments

We wish to thank Marco Polin for collaborating in early stages of this work, and for many helpful comments. We also wish to thank Anton Leykin for patiently answering our many naive questions about Bernstein–Sato polynomials; Malek Abdesselam for supplying historical references; and Alex Scott for valuable conversations. Finally, we are grateful to Nero Budur for explaining to us the connection between our results and the theory of prehomogeneous vector spaces, and for drawing our attention to the recent work of Sugiyama [98].

We wish to thank the Isaac Newton Institute for Mathematical Sciences, University of Cambridge, for generous support during the programme on Combinatorics and Statistical Mechanics (January–June 2008). One of us (A.D.S.) also thanks the Institut Henri Poincaré – Centre Emile Borel for

³⁴ More generally, this is the Bernstein–Sato polynomial for a homogeneous polynomial $P(x_1, x_2)$ of degree 3 with generic (e.g. random) coefficients [67].

³⁵ This is in our “shifted” notation (1.3). In the customary notation, the roots are rational numbers that are strictly negative. See footnote 7 above.

hospitality during the programmes on Interacting Particle Systems, Statistical Mechanics and Probability Theory (September–December 2008) and Statistical Physics, Combinatorics and Probability (September–December 2009). Finally, we thank the Laboratoire de Physique Théorique of the École Normale Supérieure (Paris) for hospitality during May–June 2010 and April–June 2011.

This research was supported in part by U.S. National Science Foundation grants PHY-0099393 and PHY-0424082.

Appendix A. Grassmann algebra and Gaussian integration

In this appendix we collect some needed information on Grassmann algebra (= exterior algebra) and Gaussian integration (both bosonic and fermionic). We begin by recalling the main properties of determinants, permanents, pfaffians and hafnians (Appendix A.1). We then recall the well-known properties of “bosonic” Gaussian integration, i.e. Gaussian integration over \mathbb{R}^n or \mathbb{C}^n (Appendix A.2). Next we define Grassmann algebra (Appendix A.3) and Grassmann–Berezin (“fermionic”) integration (Appendix A.4). Finally, we explain the formulae for fermionic Gaussian integration (Appendix A.5), which will play a central role in this paper. Our presentation in these latter three subsections is strongly indebted to Abdesselam [2, Section 2]; see also Zinn-Justin [117, Chapter 1] for a treatment aimed at physicists.

A.1. Determinants, permanents, pfaffians and hafnians

Notation. If A is an $m \times n$ matrix, then for subsets of indices $I \subseteq [m]$ and $J \subseteq [n]$ we denote by A_{IJ} the matrix A restricted to rows in I and columns in J , all kept in their original order. We also use the shorthand notation $A_{\star J} = A_{[m]J}$ when all the rows are kept, and $A_{I\star} = A_{I[n]}$ when all the columns are kept. Finally, if A is invertible, we denote by A^{-T} the matrix $(A^{-1})^T = (A^T)^{-1}$.

A.1.1. Permanent and determinant

Let R be a commutative ring with identity; we shall consider matrices with entries in R . In particular, if $A = (a_{ij})_{i,j=1}^n$ is an $n \times n$ matrix with entries in R , we define its *permanent*

$$\text{per } A = \sum_{\sigma \in \mathcal{S}_n} a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \quad (\text{A.1})$$

and its *determinant*

$$\det A = \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}. \quad (\text{A.2})$$

Here the sums range over all permutations σ of $[n] \equiv \{1, \dots, n\}$, and $\text{sgn}(\sigma) = (-1)^{\#(\text{even cycles of } \sigma)}$ is the sign of the permutation σ . See [68] and [79] for basic information on permanents and determinants, respectively.

In this paper we shall need only a few of the most elementary properties of determinants:

Lemma A.1 (*Properties of the determinant*).

- (a) $\det I = 1$.
- (b) $\det(AB) = (\det A)(\det B)$.
- (c) (**Cauchy–Binet formula**) *More generally, let A be an $m \times n$ matrix, and let B be an $n \times m$ matrix. Then*

$$\det(AB) = \sum_{\substack{I \subseteq [n] \\ |I|=m}} (\det A_{\star I})(\det B_{I\star}). \quad (\text{A.3})$$

(d) Let A be an $n \times n$ matrix, and define the **adjugate matrix** $\text{adj } A$ by

$$(\text{adj } A)_{ij} = (-1)^{i+j} \det A_{\{j\}^c \{i\}^c} \tag{A.4}$$

(note the transpose between the left-hand and right-hand sides). Then

$$(\text{adj } A)A = A(\text{adj } A) = (\det A)I. \tag{A.5}$$

In particular, A is invertible in the ring $R^{n \times n}$ if and only if $\det A$ is invertible in the ring R (when R is a field, this means simply that $\det A \neq 0$); and in this case

$$A^{-1} = (\det A)^{-1}(\text{adj } A) \tag{A.6}$$

(Cramer’s rule).

(e) **(Jacobi’s identity)** More generally, if $I, J \subseteq [n]$ with $|I| = |J| = k$, then

$$\det((A^{-T})_{I,J}) = (\det A)^{-1} \epsilon(I, J) (\det A_{I^c J^c}) \tag{A.7}$$

where $\epsilon(I, J) = (-1)^{\sum_{i \in I} i + \sum_{j \in J} j}$.

(f) **(Multi-row Laplace expansion)** For any fixed set of rows $I \subseteq [n]$ with $|I| = k$, we have

$$\det A = \sum_{\substack{J \subseteq [n] \\ |J|=k}} \epsilon(I, J) (\det A_{I,J}) (\det A_{I^c J^c}). \tag{A.8}$$

A.1.2. Hafnian

Let $A = (a_{ij})_{i,j=1}^{2m}$ be a $2m \times 2m$ symmetric matrix with entries in R . We then define the *hafnian* [13]

$$\text{hf } A = \sum_{M \in \mathcal{M}_{2m}} \prod_{ij \in M} a_{ij}, \tag{A.9}$$

where the sum runs over all *perfect matchings* of the set $[2m]$, i.e. all partitions of the set $[2m]$ into m disjoint pairs. There are $(2m - 1)!! = (2m)! / (2^m m!)$ terms in this sum. We have, for example,

$$\text{hf} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = a_{12}, \tag{A.10a}$$

$$\text{hf} \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{12} & a_{22} & a_{23} & a_{24} \\ a_{13} & a_{23} & a_{33} & a_{34} \\ a_{14} & a_{24} & a_{34} & a_{44} \end{pmatrix} = a_{12}a_{34} + a_{13}a_{24} + a_{14}a_{23}. \tag{A.10b}$$

Note that the diagonal elements of A play no role in the hafnian.

Equivalently, we can identify matchings with a subclass of permutations by writing each pair $ij \in M$ in the order $i < j$ and then writing these pairs in increasing order of their first elements: we therefore have

$$\text{hf } A = \sum_{\sigma \in \mathcal{S}_{2m}^*} a_{\sigma(1)\sigma(2)} \cdots a_{\sigma(2m-1)\sigma(2m)}, \tag{A.11}$$

where the sum runs over all permutations σ of $[2m]$ satisfying $\sigma(1) < \sigma(3) < \cdots < \sigma(2m - 1)$ and $\sigma(2k - 1) < \sigma(2k)$ for $k = 1, \dots, m$.

Note that if we were to sum over all permutations, we would obtain each term in $\text{hf } A$ exactly $2^m m!$ times. Therefore, if the ring R contains the rationals, we can alternatively write

$$\text{hf } A = \frac{1}{2^m m!} \sum_{\sigma \in \mathcal{S}_{2m}} a_{\sigma(1)\sigma(2)} \cdots a_{\sigma(2m-1)\sigma(2m)}. \tag{A.12}$$

A.1.3. Pfaffian

Finally, let $A = (a_{ij})_{i,j=1}^{2m}$ be a $2m \times 2m$ antisymmetric matrix (i.e. $a_{ij} = -a_{ji}$ and $a_{ii} = 0$) with entries in R .³⁶ We then define the pfaffian by

$$\text{pf } A = \sum_{M \in \mathcal{M}_{2m}} \epsilon(\vec{M}, \vec{M}_0) \prod_{\substack{(i,j) \in \vec{M} \\ i < j}} a_{ij}. \tag{A.13}$$

Here the sum runs once again over all perfect matchings M of the set $[2m]$, and \vec{M} is an (arbitrarily chosen) oriented version of M , i.e. for each unordered pair $ij \in M$ one chooses an ordering (i, j) of the two elements. The value of the summand in (A.13) will be independent of the choice of \vec{M} because $\epsilon(\vec{M}, \vec{M}_0)$ will be odd under reorderings of pairs (see below), while A is antisymmetric. Here \vec{M}_0 is some fixed oriented perfect matching of $[2m]$ (we call it the “reference matching”). The sign $\epsilon(\vec{M}_1, \vec{M}_2)$ is defined as follows: If $\vec{M}_1 = \{(i_1, i_2), (i_3, i_4), \dots, (i_{2m-1}, i_{2m})\}$ and $\vec{M}_2 = \{(j_1, j_2), (j_3, j_4), \dots, (j_{2m-1}, j_{2m})\}$, then $\epsilon(\vec{M}_1, \vec{M}_2)$ is the sign of the permutation that takes $i_1 \cdots i_{2m}$ into $j_1 \cdots j_{2m}$. (This is well defined, i.e. independent of the order in which the ordered pairs of \vec{M}_1 and \vec{M}_2 are written, because interchanging two pairs is an even permutation.) This quantity is clearly odd under reorderings of pairs in \vec{M}_1 or \vec{M}_2 , and has the following properties:

- (a) $\epsilon(\vec{M}_1, \vec{M}_2) = \epsilon(\vec{M}_2, \vec{M}_1)$;
- (b) $\epsilon(\vec{M}, \vec{M}) = +1$;
- (c) $\epsilon(\vec{M}_1, \vec{M}_2)\epsilon(\vec{M}_2, \vec{M}_3) = \epsilon(\vec{M}_1, \vec{M}_3)$;
- (d) $\epsilon(\vec{M}_1, \vec{M}_2) = -1$ whenever \vec{M}_1 and \vec{M}_2 differ by reversal of the orientation of a single edge;
- (e) $\epsilon(\vec{M}_1, \vec{M}_2) = -1$ whenever \vec{M}_1 and \vec{M}_2 differ by changing directed edges $(a, b), (c, d)$ in \vec{M}_1 to $(b, c), (d, a)$ in \vec{M}_2 .

Indeed, it is not hard to show that $\epsilon(\vec{M}_1, \vec{M}_2)$ is the unique map from pairs of oriented perfect matchings into $\{\pm 1\}$ that has these five properties.³⁷ It follows from (a)–(c) that the oriented perfect matchings fall into two classes (call them A and B) such that $\epsilon(\vec{M}_1, \vec{M}_2)$ equals $+1$ if \vec{M}_1 and \vec{M}_2 belong to the same class and -1 if they belong to different classes. The choice of reference matching \vec{M}_0 really amounts, therefore, to choosing one of the two equivalence classes of matchings as the reference class, and thereby fixing the sign of the pfaffian.

³⁶ Such matrices are sometimes called *alternating* matrices, in order to emphasize that the condition $a_{ii} = 0$ is imposed. This latter condition is a consequence of $a_{ij} = -a_{ji}$ whenever R is an integral domain of characteristic $\neq 2$ (so that $2x = 0$ implies $x = 0$), but not in general otherwise. See e.g. [65, Section XV.9]. In this paper we use the term “antisymmetric” to denote $a_{ij} = -a_{ji}$ and $a_{ii} = 0$.

³⁷ This also implies that $\epsilon(\vec{M}_1, \vec{M}_2)$ can be given an equivalent (more graph-theoretic) definition as follows: Form the union $\vec{M}_1 \cup \vec{M}_2$. Ignoring orientations, it is a disjoint union of even-length cycles. Looking now at the orientations, let us call a cycle even (resp. odd) if it has an even (resp. odd) number of edges pointing in each of the two directions around the cycle. We then have $\epsilon(\vec{M}_1, \vec{M}_2) = (-1)^{\#(\text{even cycles})}$.

The choice of \vec{M}_0 can be encoded in an antisymmetric matrix J defined by

$$J_{ij} = \begin{cases} 1 & \text{if } (i, j) \in \vec{M}_0, \\ -1 & \text{if } (j, i) \in \vec{M}_0, \\ 0 & \text{otherwise} \end{cases} \tag{A.14}$$

and satisfying $\text{pf}(J) = 1$. The two most common conventions for the reference matching are

$$\vec{M}_0 = \{(1, 2), (3, 4), \dots, (2m-1, 2m)\}, \quad J = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & & & \\ & & \boxed{0 & 1} & & \\ & & \boxed{-1 & 0} & & \\ & & & & \ddots & \end{pmatrix} \tag{A.15}$$

and

$$\vec{M}_0 = \{(1, m+1), (2, m+2), \dots, (m, 2m)\}, \quad J = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}. \tag{A.16}$$

In this paper we shall adopt the convention (A.15). We thus have

$$\text{pf} \begin{pmatrix} 0 & a_{12} \\ -a_{12} & 0 \end{pmatrix} = a_{12}, \tag{A.17a}$$

$$\text{pf} \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{pmatrix} = a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23}. \tag{A.17b}$$

By identifying matchings $M \in \mathcal{M}_{2m}$ with permutations $\sigma \in \mathcal{S}_{2m}^*$ as was done for the hafnian, we can equivalently write

$$\text{pf } A = \text{sgn}(\sigma_0) \sum_{\sigma \in \mathcal{S}_{2m}^*} \text{sgn}(\sigma) a_{\sigma(1)\sigma(2)} \cdots a_{\sigma(2m-1)\sigma(2m)} \tag{A.18}$$

where $\sigma_0 \in \mathcal{S}_{2m}^*$ is the permutation corresponding to the reference matching M_0 . [For our choice (A.15), σ_0 is the identity permutation.] If the ring R contains the rationals, we can alternatively write

$$\text{pf } A = \frac{1}{2^m m!} \text{sgn}(\sigma_0) \sum_{\sigma \in \mathcal{S}_{2m}} \text{sgn}(\sigma) a_{\sigma(1)\sigma(2)} \cdots a_{\sigma(2m-1)\sigma(2m)}. \tag{A.19}$$

Let us now recall the following basic properties of pfaffians:

Lemma A.2 (Properties of the pfaffian). *Let A be an antisymmetric $2m \times 2m$ matrix with elements in a commutative ring R . Then:*

- (a) $\text{pf } J = 1$.
- (b) $(\text{pf } A)^2 = \det A$.
- (c) $\text{pf}(XAX^T) = (\det X)(\text{pf } A)$ for any $2m \times 2m$ matrix X .

(d) (**Minor summation formula for pfaffians [53,54]**) More generally, we have

$$\text{pf}(XAX^T) = \sum_{\substack{I \subseteq [2m] \\ |I|=2\ell}} (\det X_{*I})(\text{pf } A_{II}) \tag{A.20}$$

for any $2\ell \times 2m$ matrix X ($\ell \leq m$). Here X_{*I} denotes the submatrix of X with columns I (and all its rows).

(e) (**Jacobi’s identity for pfaffians**) If A is invertible, then $\text{pf}(A^{-T}) = (\text{pf } A)^{-1}$ and more generally

$$\text{pf}((A^{-T})_{II}) = \epsilon(I)(\text{pf } A)^{-1}(\text{pf } A_{I^c I^c}) \tag{A.21}$$

for any $I \subseteq [2m]$, where $\epsilon(I) = (-1)^{|I|(|I|-1)/2}(-1)^{\sum_{i \in I} i}$.

See [38,40,48,54,62,65,94] for further information on pfaffians.

Remark. In this paper we will not in fact use the minor summation formula for pfaffians; but we will rederive it using Grassmann–Berezin integration. See Theorem A.15 and the comments following it.

A.2. Bosonic Gaussian integration

We shall use the following notation: If $A = (a_{ij})$ is an $m \times n$ matrix, and $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_\ell)$ are sequences of indices (not necessarily distinct or ordered) in $[m]$ and $[n]$, respectively, then we denote by A_{IJ} the $k \times \ell$ matrix defined by

$$(A_{IJ})_{\alpha\beta} = a_{i_\alpha j_\beta}. \tag{A.22}$$

This generalizes our notation A_{IJ} for subsets $I \subseteq [m]$ and $J \subseteq [n]$, where a subset is identified with the sequence of its elements written in increasing order. We shall also use the corresponding notation for vectors: namely, if $\lambda = (\lambda_i)$ is an n -vector and $I = (i_1, \dots, i_k)$ is a sequence of indices (not necessarily distinct or ordered) in $[n]$, then we denote by λ_I the k -vector defined by $(\lambda_I)_\alpha = \lambda_{i_\alpha}$.

Let $\varphi = (\varphi_i)_{i=1}^n$ be real variables; we shall write

$$\mathcal{D}\varphi = \prod_{i=1}^n \frac{d\varphi_i}{\sqrt{2\pi}} \tag{A.23}$$

for Lebesgue measure on \mathbb{R}^n with a slightly unconventional normalization. Let $A = (a_{ij})_{i,j=1}^n$ be a real symmetric positive–definite $n \times n$ matrix. We then have the following fundamental facts about Gaussian integration on \mathbb{R}^n :

Theorem A.3 (Wick’s theorem for real bosons). Let $A = (a_{ij})_{i,j=1}^n$ be a real symmetric positive–definite $n \times n$ matrix. Then:

(a) For any vector $c = (c_i)_{i=1}^n$ in \mathbb{R}^n (or \mathbb{C}^n), we have

$$\int \mathcal{D}\varphi \exp\left(-\frac{1}{2}\varphi^T A \varphi + c^T \varphi\right) = (\det A)^{-1/2} \exp\left(\frac{1}{2}c^T A^{-1}c\right). \tag{A.24}$$

(b) For any sequence of indices $I = (i_1, \dots, i_r)$ in $[n]$, we have

$$\int \mathcal{D}\varphi \varphi_{i_1} \cdots \varphi_{i_r} \exp\left(-\frac{1}{2}\varphi^T A \varphi\right) = \begin{cases} 0 & \text{if } r \text{ is odd,} \\ (\det A)^{-1/2} \text{hf}((A^{-1})_{II}) & \text{if } r \text{ is even.} \end{cases} \tag{A.25}$$

(c) More generally, for any real or complex $r \times n$ matrix C , we have

$$\int \mathcal{D}\varphi \left(\prod_{\alpha=1}^r (C\varphi)_{\alpha} \right) \exp\left(-\frac{1}{2}\varphi^T A\varphi\right) = \begin{cases} 0 & \text{if } r \text{ is odd,} \\ (\det A)^{-1/2} \text{hf}(CA^{-1}C^T) & \text{if } r \text{ is even.} \end{cases} \quad (\text{A.26})$$

Historical remarks. Physicists call these formulae “Wick’s theorem” because Gian-Carlo Wick [113] proved the analogue of (A.25) for the correlation functions of a free quantum field (see e.g. [95]). These formulae are called “bosonic” because the functional-integral formulation for bosonic quantum fields (see e.g. [117]) leads to ordinary integrals over \mathbb{R}^n or \mathbb{C}^n (or infinite-dimensional generalizations thereof). By contrast, functional integrals for fermionic quantum fields lead to Grassmann–Berezin integrals, to be discussed in Appendices A.3–A.5.

The formula (A.25) for the moments of a mean-zero Gaussian measure goes back at least to Isserlis [55] in 1918. We thank Malek Abdesselam for drawing our attention to this reference.

Now let $\varphi = (\varphi_i)_{i=1}^n$ be complex variables; we denote complex conjugation by $\bar{}$ and shall write

$$\mathcal{D}(\varphi, \bar{\varphi}) = \prod_{i=1}^n \frac{(d \operatorname{Re} \varphi_i)(d \operatorname{Im} \varphi_i)}{\pi} \quad (\text{A.27})$$

for Lebesgue measure on \mathbb{C}^n with a slightly unconventional normalization. Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ complex matrix (not necessarily symmetric or hermitian) whose hermitian part $\frac{1}{2}(A + A^*)$ is positive-definite. We then have the following fundamental facts about Gaussian integration on \mathbb{C}^n :

Theorem A.4 (Wick’s theorem for complex bosons). *Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ complex matrix whose hermitian part $\frac{1}{2}(A + A^*)$ is positive-definite. Then:*

(a) For any vectors $b = (b_i)_{i=1}^n$ and $c = (c_i)_{i=1}^n$ in \mathbb{C}^n , we have

$$\int \mathcal{D}(\varphi, \bar{\varphi}) \exp(-\bar{\varphi}^T A\varphi + \bar{b}^T \varphi + \bar{\varphi}^T c) = (\det A)^{-1} \exp(\bar{b}^T A^{-1} c). \quad (\text{A.28})$$

(b) For any sequences of indices $I = (i_1, \dots, i_r)$ and $J = (j_1, \dots, j_s)$ in $[n]$, we have

$$\int \mathcal{D}(\varphi, \bar{\varphi}) \varphi_{i_1} \cdots \varphi_{i_r} \bar{\varphi}_{j_1} \cdots \bar{\varphi}_{j_s} \exp(-\bar{\varphi}^T A\varphi) = \begin{cases} 0 & \text{if } r \neq s, \\ (\det A)^{-1} \operatorname{per}((A^{-1})_{IJ}) & \text{if } r = s. \end{cases} \quad (\text{A.29})$$

(c) More generally, for any complex $r \times n$ matrix B and any complex $n \times s$ matrix C , we have

$$\begin{aligned} & \int \mathcal{D}(\varphi, \bar{\varphi}) \left(\prod_{\alpha=1}^r (B\varphi)_{\alpha} \right) \left(\prod_{\beta=1}^s (\bar{\varphi}^T C)_{\beta} \right) \exp(-\bar{\varphi}^T A\varphi) \\ &= \begin{cases} 0 & \text{if } r \neq s, \\ (\det A)^{-1} \operatorname{per}(BA^{-1}C) & \text{if } r = s. \end{cases} \end{aligned} \quad (\text{A.30})$$

Some final remarks. 1. In this article we shall use mainly the “source” versions of Wick’s theorem, i.e. part (a) of Theorems A.3 and A.4 and the corresponding theorems for fermions. The “correlation function” versions, i.e. parts (b) and (c), will be used only in Sections 4 and 8.1 and in the second proofs of Corollaries 5.2 and 5.4 (Section 5.2).

2. We have here presented bosonic Gaussian integration in an analytic context, i.e. integration on \mathbb{R}^n or \mathbb{C}^n . A combinatorial abstraction of bosonic Gaussian integration can be found in [1].

A.3. Grassmann algebra

Let R be a commutative ring. Every textbook on elementary abstract algebra defines the ring $R[x_1, \dots, x_n]$ of polynomials in commuting indeterminates x_1, \dots, x_n with coefficients in R , and studies its properties. Here we would like briefly to do the same for the ring $R[\chi_1, \dots, \chi_n]_{\text{Grass}}$ of polynomials in *anticommuting* indeterminates χ_1, \dots, χ_n : we call this ring the *Grassmann algebra* over R in generators χ_1, \dots, χ_n . (Of course, readers familiar with exterior algebra will recognize this as nothing other than the exterior algebra $\Lambda(R^n)$ built from the free R -module of dimension n .³⁸) To lighten the notation, we shall henceforth omit the subscripts $_{\text{Grass}}$, since it will always be clear by context whether we are referring to the Grassmann algebra or to the ordinary polynomial ring.

Here is the precise definition:

Definition A.5 (*Grassmann algebra*). Let R be a commutative ring with identity element,³⁹ and let χ_1, \dots, χ_n be a collection of letters. The Grassmann algebra $R[\chi_1, \dots, \chi_n]$ (or $R[\chi]$ for short) is the quotient of the ring $R\langle \chi_1, \dots, \chi_n \rangle$ of noncommutative polynomials in the letters χ_1, \dots, χ_n by the two-sided ideal generated by the expressions $\chi_i \chi_j + \chi_j \chi_i$ ($1 \leq i < j \leq n$) and χ_i^2 ($1 \leq i \leq n$). We can consider $R[\chi]$ as a ring and also as an R -algebra.

In other words, the generators χ_i of $R[\chi]$ satisfy the *anticommutation relations*

$$\chi_i \chi_j + \chi_j \chi_i = 0 \quad \text{for all } i, j \in [n] \tag{A.31}$$

as well as the relations

$$\chi_i^2 = 0 \quad \text{for all } i \in [n]. \tag{A.32}$$

Please note that the anticommutation relation (A.31) for $i = j$ states that $2\chi_i^2 = 0$; but this need not imply $\chi_i^2 = 0$ if the coefficient ring R does not contain an element $\frac{1}{2}$. For this reason we have explicitly adjoined the relations $\chi_i^2 = 0$. Of course, if the coefficient ring R contains an element $\frac{1}{2}$ (e.g. if it contains the rationals), then this extra relation could be replaced by the cases $i = j$ of $\chi_i \chi_j + \chi_j \chi_i = 0$.

The first important property of $R[\chi]$ is the following:

Proposition A.6. $R[\chi]$ is a free R -module with basis given by the 2^n monomials $\chi^l = \chi_{i_1} \cdots \chi_{i_p}$ where $l = \{i_1, \dots, i_p\} \subseteq [n]$ with $i_1 < \dots < i_p$.

It follows that each element $f \in R[\chi]$ can be written uniquely in the form

$$f = \sum_{l \subseteq [n]} f_l \chi^l \tag{A.33}$$

with $f_l \in R$. The term f_\emptyset that contains no factors χ_i is sometimes termed the *body* of f , and the rest $\sum_{l \neq \emptyset} f_l \chi^l$ is sometimes termed the *soul* of f .

³⁸ See e.g. [30, Section 6.4] or [65, Section XIX.1].

³⁹ Much of the elementary theory works also for coefficient rings without identity element. The main change is that polynomials and formal power series Φ must have constant term c_0 in R (and not in \mathbb{Z}) in order to be applied to elements $f \in R[\chi]_+$: see the paragraphs immediately after Proposition A.9. But this means that we can consider the most important case, namely $\Phi = \exp$, only when R has an identity element. For this reason it is convenient simply to make this assumption from the beginning.

Multiplication in the Grassmann algebra is of course R -bilinear, i.e.

$$\left(\sum_{I \subseteq [n]} f_I \chi^I\right) \left(\sum_{J \subseteq [n]} g_J \chi^J\right) = \sum_{I, J \subseteq [n]} f_I g_J \chi^I \chi^J \tag{A.34}$$

where

$$\chi^I \chi^J = \begin{cases} \sigma(I, J) \chi^{I \cup J} & \text{if } I \cap J = \emptyset, \\ 0 & \text{if } I \cap J \neq \emptyset, \end{cases} \tag{A.35}$$

and $\sigma(I, J)$ is the sign of the permutation that rearranges the sequence IJ into increasing order when I and J are each written in increasing order.

We define the *degree* of a monomial χ^I in the obvious way, namely, $\deg(\chi^I) = |I|$. The Grassmann algebra $R[\chi]$ then possesses a natural \mathbb{N} -grading

$$R[\chi] = \bigoplus_{p=0}^n R[\chi]_p \tag{A.36}$$

where $R[\chi]_p$ is generated, as an R -module, by the monomials of degree p . A coarser grading is the Z_2 -grading

$$R[\chi] = R[\chi]_{\text{even}} \oplus R[\chi]_{\text{odd}} \tag{A.37}$$

where

$$R[\chi]_{\text{even}} \stackrel{\text{def}}{=} \bigoplus_{p \text{ even}} R[\chi]_p, \tag{A.38}$$

$$R[\chi]_{\text{odd}} \stackrel{\text{def}}{=} \bigoplus_{p \text{ odd}} R[\chi]_p. \tag{A.39}$$

Note that $R[\chi]_{\text{even}}$ is a subalgebra of $R[\chi]$ (but $R[\chi]_{\text{odd}}$ is not). The *parity operator* P , defined by

$$P\left(\sum_{I \subseteq [n]} f_I \chi^I\right) = \sum_{I \subseteq [n]} (-1)^{|I|} f_I \chi^I, \tag{A.40}$$

is an involutive automorphism of $R[\chi]$: it acts as the identity on $R[\chi]_{\text{even}}$ and as minus the identity on $R[\chi]_{\text{odd}}$. A nonzero element f that belongs to either $R[\chi]_{\text{even}}$ or $R[\chi]_{\text{odd}}$ is said to be Z_2 -homogeneous, and its *parity* is $p(f) \stackrel{\text{def}}{=} 0$ in the first case and $p(f) \stackrel{\text{def}}{=} 1$ in the second. It is easy to see that Z_2 -homogeneous elements f, g satisfy the commutation/anticommutation relations

$$fg = (-1)^{p(f)p(g)} gf. \tag{A.41}$$

In other words, odd elements anticommute with other odd elements, while even elements commute with homogeneous elements of both types. One easily deduces the following two consequences:

Proposition A.7. *An even element $f \in R[\chi]_{\text{even}}$ commutes with the entire Grassmann algebra. (In particular, $R[\chi]_{\text{even}}$ is a commutative ring.)*

Proposition A.8. An odd element $f \in R[\chi]_{\text{odd}}$ is nilpotent of order 2, i.e. $f^2 = 0$.⁴⁰

A further consequence of the relations (A.31)/(A.32) and the finiteness of the number n of generators is that every element of

$$R[\chi]_+ \stackrel{\text{def}}{=} \bigoplus_{p=1}^n R[\chi]_p \tag{A.42}$$

(the set of elements with no term of degree 0) is nilpotent:

Proposition A.9. A “pure soul” element $f \in R[\chi]_+$ is nilpotent of order at most $M = \lfloor n/2 \rfloor + 2$, i.e. $f^M = 0$.

Proof. Write $f = f_0 + f_1$ with $f_0 \in R[\chi]_{\text{even}} \cap R[\chi]_+$ and $f_1 \in R[\chi]_{\text{odd}}$. Then f_0 and f_1 commute [by (A.41)] and $f_1^n = 0$ for $n \geq 2$ [by Proposition A.8], so we have $f^k = f_0^k + kf_0^{k-1}f_1$ for all k . And it is easy to see, using (A.31)/(A.32), that $f_0^k = 0$ for $k > n/2$. \square

If $I = \{i_1, \dots, i_p\}$ is a subset of $[n]$, the Grassmann algebra $R[\chi_I] \stackrel{\text{def}}{=} R[\chi_{i_1}, \dots, \chi_{i_p}]$ is naturally isomorphic to the subalgebra of $R[\chi] = R[\chi_1, \dots, \chi_n]$ generated by $\{\chi_i\}_{i \in I}$; we shall identify these two algebras and use the same notation $R[\chi_I]$ for both. In particular, the degree-zero subalgebra $R[\chi]_0$ is identified with the coefficient ring R .

If $\Phi(x) = \sum_{k=0}^N c_k x^k$ is any polynomial in a single indeterminate x with coefficients in either R or \mathbb{Z} , we can of course apply it to any $f \in R[\chi]$ to obtain $\Phi(f) = \sum_{k=0}^N c_k f^k \in R[\chi]$. Moreover, if $\Phi(x) = \sum_{k=0}^\infty c_k x^k$ is a formal power series with coefficients in R or \mathbb{Z} , we can apply it to any $f \in R[\chi]_+$ because f is nilpotent and the sum is therefore finite.

But more is true: suppose we have a formal power series $\Phi(x) = \sum_{k=0}^\infty (c_k/k!)x^k$ with coefficients in R or \mathbb{Z} (note the factorial denominators!). If the coefficient ring R contains the rationals as a subring (as it is usually convenient to assume), then of course $\Phi(f)$ is well defined for any $f \in R[\chi]_+$. But we claim that this expression has an unambiguous meaning for $f \in R[\chi]_+$ even if R does not contain the rationals. Indeed, let $f = \sum_{I \neq \emptyset} f_I \chi^I$ and consider the expansion of

$$f^k = \left(\sum_{I \neq \emptyset} f_I \chi^I \right)^k = \sum_{I_1, \dots, I_k \neq \emptyset} f_{I_1} \cdots f_{I_k} \chi^{I_1} \cdots \chi^{I_k}. \tag{A.43}$$

Whenever two or more of the sets I_1, \dots, I_k are equal (or indeed have any elements in common), we have $\chi^{I_1} \cdots \chi^{I_k} = 0$ by (A.31)/(A.32). When, by contrast, the sets I_1, \dots, I_k are all distinct, then there are $k!$ terms with the same coefficient $f_{I_1} \cdots f_{I_k}$ corresponding to the $k!$ different permutations of the factors $\chi^{I_1}, \dots, \chi^{I_k}$, and these terms are either all equal (if at most one of the sets I_α is odd) or else add to zero (if two or more of them are odd). It follows that

⁴⁰ **Proof of Proposition A.8.** Write $f = \sum_{I \text{ odd}} f_I \chi^I$; then

$$f^2 = \sum_{I \text{ odd}} f_I^2 (\chi^I)^2 + \sum_{\substack{I, J \text{ odd} \\ I < J}} f_I f_J (\chi^I \chi^J + \chi^J \chi^I)$$

where $I < J$ denotes that the distinct sets I and J are written in increasing lexicographic order. But $(\chi^I)^2 = 0$ and $\chi^I \chi^J + \chi^J \chi^I = 0$ by (A.31)/(A.32) and (A.41). \square

$$f^k = k! \sum_{\substack{I_1 < \dots < I_k \\ \text{at most one } |I_\alpha| \text{ odd}}} f_{I_1} \dots f_{I_k} \chi^{I_1} \dots \chi^{I_k} \tag{A.44}$$

where $I_1 < \dots < I_k$ denotes that the distinct sets I_1, \dots, I_k are written in increasing lexicographic order. So we can define

$$\Phi(f) = \sum_{k=0}^n c_k \sum_{\substack{I_1 < \dots < I_k \\ \text{at most one } |I_\alpha| \text{ odd}}} f_{I_1} \dots f_{I_k} \chi^{I_1} \dots \chi^{I_k} \tag{A.45}$$

even if the coefficient ring R does not contain the rationals.

The most important case is $\Phi = \exp$. Note that $\exp(f + g) = \exp(f)\exp(g)$ whenever $f, g \in R[\chi]_+ \cap R[\chi]_{\text{even}}$, but not in general otherwise. In this paper we will apply the exponential *only* to even elements.

We also need to define one other type of composition. Suppose first that $f \in R\langle \chi_1, \dots, \chi_n \rangle$ is a noncommutative polynomial in the letters χ_1, \dots, χ_n with coefficients in R , and let ξ_1, \dots, ξ_n be elements of some ring R' (not necessarily commutative) that contains R within its center. Then the composition $f(\xi_1, \dots, \xi_n)$, obtained by substituting each χ_i by the corresponding ξ_i , is a well-defined element of R' ; furthermore, this composition satisfies the obvious laws

$$(f + g)(\xi_1, \dots, \xi_n) = f(\xi_1, \dots, \xi_n) + g(\xi_1, \dots, \xi_n), \tag{A.46a}$$

$$(fg)(\xi_1, \dots, \xi_n) = f(\xi_1, \dots, \xi_n)g(\xi_1, \dots, \xi_n). \tag{A.46b}$$

Now suppose, instead, that f belongs to the Grassmann algebra $R[\chi_1, \dots, \chi_n] = R\langle \chi_1, \dots, \chi_n \rangle / \{\chi_i \chi_j + \chi_j \chi_i, \chi_i^2\}$ and that ξ_1, \dots, ξ_n are elements of R' that satisfy

$$\xi_i \xi_j + \xi_j \xi_i = 0, \tag{A.47a}$$

$$\xi_i^2 = 0 \tag{A.47b}$$

for all i, j . Then the composition $f(\xi_1, \dots, \xi_n)$ is again a well-defined element of R' , because the relations (A.47) guarantee that any representative of f in $R\langle \chi_1, \dots, \chi_n \rangle$ will give, after substituting each χ_i by ξ_i , the same element of R' ; moreover, the laws (A.46) continue to hold. In particular, we can take R' to be another Grassmann algebra over R (which may or may not contain some of the χ_i as generators, and which may or may not contain additional generators) and ξ_1, \dots, ξ_n to be arbitrary *odd* elements of this Grassmann algebra; the relations (A.47) hold by virtue of (A.41) and Proposition A.8, respectively. We shall exploit this type of composition in Propositions A.10 and A.12 below.

A.4. Grassmann–Berezin (fermionic) integration

Thus far we have simply been recalling standard facts about exterior algebra. But now we go on to introduce a process called *Grassmann–Berezin integration*, which has become a standard tool of theoretical physicists over the last 40 years but is still surprisingly little known among mathematicians. As we shall see, the term “integration” is a misnomer, because the construction is purely algebraic/combinatorial. But the term is nevertheless felicitous, because Grassmann–Berezin integration behaves in many ways *analogously* to ordinary integration over \mathbb{R}^n or \mathbb{C}^n , and this analogy is heuristically very fruitful.

We start by defining, for each $i \in [n]$, the derivation $\partial_i = \frac{\partial}{\partial \chi_i}$ (acting to the right) as the R -linear map $\partial_i : R[\chi] \rightarrow R[\chi]$ defined by the following action on monomials $\chi_{i_1} \dots \chi_{i_p}$ with $i_1 < \dots < i_p$:

$$\partial_i \chi_{i_1} \cdots \chi_{i_p} \stackrel{\text{def}}{=} \begin{cases} (-1)^{\alpha-1} \chi_{i_1} \cdots \chi_{i_{\alpha-1}} \chi_{i_{\alpha+1}} \cdots \chi_{i_p} & \text{if } i = i_\alpha, \\ 0 & \text{if } i \notin \{i_1, \dots, i_p\}. \end{cases} \tag{A.48}$$

(Of course, in the former instance there is a *unique* index $\alpha \in [p]$ for which $i = i_\alpha$, so this definition is unambiguous.) It is then easy to see that (A.48) holds also when the indices i_1, \dots, i_p are not necessarily ordered, provided that they are all distinct. Clearly ∂_i is a map of degree -1 , i.e. $\partial_i : R[\chi]_p \rightarrow R[\chi]_{p-1}$; moreover, it takes values in $R[\chi_{\{i\}^c}]$, i.e. the subalgebra generated by $\{\chi_j\}_{j \neq i}$. Furthermore, the maps ∂_i satisfy

$$\partial_i^2 = 0, \tag{A.49}$$

$$\partial_i \partial_j + \partial_j \partial_i = 0 \tag{A.50}$$

as well as the modified Leibniz rule

$$\partial_i (fg) = (\partial_i f)g + (Pf)(\partial_i g) \tag{A.51}$$

where P is the parity operator (A.40).

We now make a surprising definition: *integration is the same as differentiation*. That is, we define

$$\int d\chi_i f = \partial_i f. \tag{A.52}$$

We always write the operator of integration *to the left* of the integrand, just as we do for the (completely equivalent) operator of differentiation. To lighten the notation we refrain from repeating the \int sign in iterated integrals, so that $\int d\chi_{i_1} \cdots d\chi_{i_k} f$ is a shorthand for $\int d\chi_{i_1} \cdots \int d\chi_{i_k} f$ and hence we have

$$\int d\chi_{i_1} \cdots d\chi_{i_k} f = \partial_{i_1} \cdots \partial_{i_k} f. \tag{A.53}$$

Note by (A.50) that changing the order of integration changes the sign:

$$\int d\chi_{i_{\sigma(1)}} \cdots d\chi_{i_{\sigma(k)}} f = \text{sgn}(\sigma) \int d\chi_{i_1} \cdots d\chi_{i_k} f \tag{A.54}$$

for any permutation $\sigma \in S_k$. For instance, reversing the order of integrations gives

$$\int d\chi_{i_k} \cdots d\chi_{i_1} f = (-1)^{k(k-1)/2} \int d\chi_{i_1} \cdots d\chi_{i_k} f. \tag{A.55}$$

In particular, when we integrate $f = \sum_{I \subseteq [n]} f_I \chi^I$ with respect to *all* the generators, we have

$$\int d\chi_n \cdots d\chi_1 f = f_{[n]}. \tag{A.56}$$

That is, integration with respect to all the generators simply picks out the coefficient of the “top” monomial $I = [n]$, provided that we write these integrations in reverse order.⁴¹ We adopt the shorthand

⁴¹ Of course, since the integrations are performed from right to left, this actually corresponds to performing $\int d\chi_1$ first and $\int d\chi_n$ last.

$$\mathcal{D}\chi \stackrel{\text{def}}{=} d\chi_n \cdots d\chi_1, \tag{A.57}$$

and we sometimes write $\mathcal{D}_n(\chi)$ if we wish to stress the number of generators.

An important special case arises when n is even, say $n = 2m$, and the generators χ_1, \dots, χ_n are divided into two sets ψ_1, \dots, ψ_m and $\bar{\psi}_1, \dots, \bar{\psi}_m$, where we think of each ψ_i as paired with its corresponding $\bar{\psi}_i$. In this case we adopt the shorthand notation

$$\mathcal{D}(\psi, \bar{\psi}) \stackrel{\text{def}}{=} d\psi_1 d\bar{\psi}_1 \cdots d\psi_m d\bar{\psi}_m \tag{A.58}$$

and we sometimes write $\mathcal{D}_m(\psi, \bar{\psi})$ if we wish to stress the number of pairs of generators. Please note that since each pair $d\psi_i d\bar{\psi}_i$ is Grassmann-even, we can also write

$$\mathcal{D}(\psi, \bar{\psi}) \stackrel{\text{def}}{=} \prod_{i=1}^m d\psi_i d\bar{\psi}_i \tag{A.59}$$

where the terms in the product can be taken in any order. Note also that

$$\mathcal{D}(\psi, \bar{\psi}) = (-1)^{m(m-1)/2} \mathcal{D}\bar{\psi} \mathcal{D}\psi. \tag{A.60}$$

The notation $\bar{}$ is intended to be suggestive of complex conjugation, but we stress that it has nothing to do with complex numbers: it merely denotes the extra combinatorial structure on the index set $[2m]$ that arises from the splitting of $[2m]$ into two sets of cardinality m and the fixing of a bijection between the two sets.⁴² In particular, the coefficient ring R is still completely arbitrary. The general case (χ) and special case $(\psi, \bar{\psi})$ are nevertheless known in the physics literature as “real fermions” and “complex fermions”, respectively. We shall retain this terminology but shall always put the adjectives “real” and “complex” in quotation marks in order to warn the reader that they are potentially misleading.

Now let $A = (A_{ij})_{i,j=1}^n$ be an $n \times n$ matrix, and define new Grassmann variables by $\xi_i = \sum_{j=1}^n A_{ij} \chi_j$. We then have

$$\begin{aligned} \int d\chi_n \cdots d\chi_1 \xi_1 \cdots \xi_n &= \int d\chi_n \cdots d\chi_1 \sum_{j_1=1}^n A_{1j_1} \chi_{j_1} \cdots \sum_{j_n=1}^n A_{nj_n} \chi_{j_n} \\ &= \sum_{\sigma} \epsilon(\sigma) A_{1\sigma(1)} \cdots A_{n\sigma(n)} \\ &= \det A. \end{aligned} \tag{A.61}$$

This result can be reformulated as follows:

Proposition A.10 (*Linear change of variables in Grassmann–Berezin integration*). *Let $f \in R[\xi_1, \dots, \xi_n]$, and define $F \in R[\chi_1, \dots, \chi_n]$ by*

$$F(\chi_1, \dots, \chi_n) = f(\xi_1, \dots, \xi_n)|_{\xi_i = \sum_{j=1}^n A_{ij} \chi_j} \tag{A.62}$$

⁴² We can view this as dividing the $2m$ individuals into m males and m females and then pairing those individuals into m heterosexual couples.

[note that this substitution is well defined by virtue of the discussion in the last paragraph of Appendix A.3]. Then

$$\int d\chi_n \cdots d\chi_1 F = (\det A) \int d\xi_n \cdots d\xi_1 f. \tag{A.63}$$

Please note that this is the reverse of the change-of-variables formula for ordinary multivariate integrals, i.e. in an ordinary integral the factor $\det(A)$ would appear on the other side.

We will also need:

Proposition A.11 (Fubini theorem for Grassmann–Berezin integration). Let $I = \{i_1, \dots, i_p\}$ with $i_1 < \dots < i_p$ be a subset of $[n]$, and let $I^c = \{j_1, \dots, j_{n-p}\}$ with $j_1 < \dots < j_{n-p}$. Then for any elements $f \in R[\chi_I]$ and $g \in R[\chi_{I^c}]$ we have

$$\int \mathcal{D}\chi_I \mathcal{D}\chi_{I^c} fg = (-1)^{p(n-p)} \left(\int \mathcal{D}\chi_I f \right) \left(\int \mathcal{D}\chi_{I^c} g \right) \tag{A.64}$$

where $\mathcal{D}\chi_I$ (resp. $\mathcal{D}\chi_{I^c}$) is shorthand for $d\chi_{i_p} \cdots d\chi_{i_1}$ (resp. $d\chi_{j_{n-p}} \cdots d\chi_{j_1}$).

Proof. Expanding f and g in monomials, we see that the only terms contributing to the integrals on either side are the “top” monomials χ^I and χ^{I^c} in f and g , respectively; so we can assume without loss of generality that $f = \chi^I$ and $g = \chi^{I^c}$. Now use the fact that integration is the same as differentiation, and successively apply the operators ∂_i for $i \in I^c$ to the product fg using the Leibniz rule (A.51). The differentiations hit only g , and we have $Pf = (-1)^p f$. It follows that

$$\int \mathcal{D}\chi_{I^c} fg = (-1)^{p(n-p)} f \left(\int \mathcal{D}\chi_{I^c} g \right). \tag{A.65}$$

The result then follows by integrating both sides with $\mathcal{D}\chi_I$. \square

Let us remark that the same formula (A.64) would hold for any choice of the orderings in defining $\mathcal{D}\chi_I$ and $\mathcal{D}\chi_{I^c}$, provided only that we use the same orderings on both sides of the equation. We have chosen to write $\mathcal{D}\chi_I = d\chi_{i_p} \cdots d\chi_{i_1}$ for compatibility with our convention (A.57) that $\mathcal{D}\chi = d\chi_n \cdots d\chi_1$.

Finally, the following proposition shows that a Grassmann–Berezin integral over $d\chi_i$ is invariant under translation by an arbitrary odd element of the Grassmann algebra that does not involve the variable χ_i . More generally, one can consider integration over a set $I \subseteq [n]$ of generators:

Proposition A.12 (Invariance under translation). Let $I = \{i_1, \dots, i_p\} \subseteq [n]$ and let $\xi_1, \dots, \xi_n \in R[\chi_{I^c}]_{\text{odd}}$ satisfy $\xi_j = 0$ whenever $j \notin I$. Then

$$\int d\chi_{i_p} \cdots d\chi_{i_1} f(\chi + \xi) = \int d\chi_{i_p} \cdots d\chi_{i_1} f(\chi) \tag{A.66}$$

where $f(\chi + \xi)$ denotes the substitution defined at the end of Appendix A.3. [Recall that the oddness of the ξ_i is required for this substitution to make sense.]

Proof. The formula (A.66) can be rewritten as

$$\partial_{i_p} \cdots \partial_{i_1} f(\chi + \xi) = \partial_{i_p} \cdots \partial_{i_1} f(\chi). \tag{A.67}$$

To prove this relation, it suffices to consider the cases in which $f(\chi) = \chi^J$ for some $J \subseteq [n]$. Now for any $i \in I$ and $j \in [n]$ we have

$$\partial_i(\chi_j + \xi_j) = \partial_i \chi_j = \delta_{ij} \tag{A.68}$$

because $\xi_j \in R[\chi_{I^c}]$. Using this relation together with the Leibniz rule, we see that $\partial_i(\chi + \xi)^J$ equals the same object in which ξ_i has been replaced by zero. Doing this successively for $\partial_{i_1}, \dots, \partial_{i_p}$, we set ξ_i to zero for all $i \in I$. But by hypothesis these are the only nonzero ξ_i . \square

A.5. Fermionic Gaussian integration

We are now ready to state the fundamental formulae (“Wick’s theorem”) for fermionic Gaussian integration, which are analogues of Theorems A.3 and A.4 for bosonic Gaussian integration. The main difference is that fermionic integration is a purely algebraic/combinatorial construction, so that it works over an arbitrary (commutative) coefficient ring R and does not require any positive-definiteness condition on the matrix A .

We begin with the formula for a pure Gaussian integral, i.e. the integral of the exponential of a quadratic form. So let R be a commutative ring with identity element, let χ_1, \dots, χ_n be the generators of a Grassmann algebra, and let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ antisymmetric matrix (i.e. $a_{ij} = -a_{ji}$ and $a_{ii} = 0$) with entries in R .⁴³ We use the notation

$$\frac{1}{2} \chi^T A \chi \stackrel{\text{def}}{=} \sum_{1 \leq i < j \leq n} \chi_i a_{ij} \chi_j \tag{A.69}$$

and observe that the right-hand side makes sense even if the coefficient ring R does not contain an element $\frac{1}{2}$. We then have the following formula, which shows that a Gaussian fermionic integral equals a pfaffian:

Proposition A.13 (Gaussian integral for “real” fermions). *Let A be an $n \times n$ antisymmetric matrix with coefficients in R . Then*

$$\int d\chi_n \cdots d\chi_1 e^{\frac{1}{2} \chi^T A \chi} = \int d\chi_1 \cdots d\chi_n e^{-\frac{1}{2} \chi^T A \chi} = \begin{cases} \text{pf } A & \text{if } n \text{ is even,} \\ 0 & \text{if } n \text{ is odd.} \end{cases} \tag{A.70}$$

Proof. We expand the exponential using (A.45) with $\Phi = \exp$ and then integrate $\int d\chi_n \cdots d\chi_1$. When n is odd the integral vanishes, and when n is even (say, $n = 2m$) the only contribution comes from $k = m$ in (A.45), yielding

$$\begin{aligned} & \int d\chi_n \cdots d\chi_1 e^{\frac{1}{2} \chi^T A \chi} \\ &= \int d\chi_n \cdots d\chi_1 \sum_{\substack{I_1 < \cdots < I_m \\ I_k = \{i_k, j_k\} \text{ with } i_k < j_k}} a_{i_1 j_1} \cdots a_{i_m j_m} \chi^{I_1} \cdots \chi^{I_m} \end{aligned} \tag{A.71a}$$

$$= \sum_{\substack{\sigma: \sigma(2i-1) < \sigma(2i) \text{ and} \\ \sigma(1) < \sigma(3) < \cdots < \sigma(2m-1)}} \epsilon(\sigma) a_{\sigma(1)\sigma(2)} \cdots a_{\sigma(2m-1)\sigma(2m)} \tag{A.71b}$$

$$= \text{pf } A. \tag{A.71c}$$

(Note that this holds whether or not the coefficient ring R contains the rational numbers.)

⁴³ See footnote 36 above concerning our use of the term “antisymmetric”.

For the variant in which A is replaced by $-A$ and the order of integration is reversed, it suffices to observe that

$$\text{pf}(-A) = (-1)^{n/2} \text{pf}(A) \tag{A.72}$$

[this is an immediate consequence of the definition (A.18)] and use (A.55) to get

$$\int d\chi_1 \cdots d\chi_n e^{-\frac{1}{2}\chi^T A \chi} = (-1)^{n(n-1)/2} \text{pf}(-A) = (-1)^{n^2/2} \text{pf}(A) = \text{pf}(A) \tag{A.73}$$

because n is even. \square

Remark. Eq. (A.70) shows that there exist two equally natural conventions for Gaussian integrals with “real” fermions: either we write our quadratic forms as $e^{\frac{1}{2}\chi^T A \chi}$ and our integrals as $\int d\chi_n \cdots d\chi_1$, or alternatively we write our quadratic forms as $e^{-\frac{1}{2}\chi^T A \chi}$ and our integrals as $\int d\chi_1 \cdots d\chi_n$. In this paper we have adopted the first convention [cf. (A.57)].

In the special case in which $n = 2m$ and the generators χ_1, \dots, χ_n are divided into two sets ψ_1, \dots, ψ_m and $\bar{\psi}_1, \dots, \bar{\psi}_m$, we have

Proposition A.14 (Gaussian integral for “complex” fermions). *Let A be an $m \times m$ matrix with coefficients in R . Then*

$$\int d\psi_1 d\bar{\psi}_1 \cdots d\psi_m d\bar{\psi}_m e^{\bar{\psi}^T A \psi} = \det A. \tag{A.74}$$

Please note that here A is an arbitrary matrix; no condition of symmetry or antisymmetry need be imposed on it.

Proof of Proposition A.14. By the change of variables $(\psi', \bar{\psi}') = (A\psi, \bar{\psi})$, we have from Proposition A.10

$$\int \mathcal{D}(\psi, \bar{\psi}) e^{\bar{\psi}^T A \psi} = (\det A) \int \mathcal{D}(\psi', \bar{\psi}') e^{\bar{\psi}'^T \psi'} \tag{A.75a}$$

$$= (\det A) \int \mathcal{D}(\psi, \bar{\psi}) \prod_{i=1}^m (1 + \bar{\psi}_i \psi_i). \tag{A.75b}$$

When we expand the product, only the term $\bar{\psi}_1 \psi_1 \cdots \bar{\psi}_m \psi_m$ has a nonzero integral, and the integral of this term is 1. \square

Here is an alternate proof, which treats “complex” fermions as a special case of “real” fermions and invokes Proposition A.13:

Alternate proof of Proposition A.14. Let us write $(\chi_1, \dots, \chi_{2m}) = (\psi_1, \dots, \psi_m, \bar{\psi}_1, \dots, \bar{\psi}_m)$. Then by (A.59) we have

$$\mathcal{D}(\psi, \bar{\psi}) = (-1)^m d\bar{\psi}_m d\psi_m \cdots d\bar{\psi}_1 d\psi_1 \tag{A.76a}$$

$$= (-1)^{m(m+1)/2} d\bar{\psi}_m \cdots d\bar{\psi}_1 d\psi_m \cdots d\psi_1 \tag{A.76b}$$

$$= (-1)^{m(m+1)/2} d\chi_{2m} \cdots d\chi_1 \tag{A.76c}$$

and

$$\bar{\psi}^T A \psi = \frac{1}{2} \chi^T K \chi, \tag{A.77}$$

where K is the $2m \times 2m$ matrix

$$K = \begin{pmatrix} 0 & -A^T \\ A & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} 0 & A^T \\ I & 0 \end{pmatrix}. \tag{A.78}$$

Then by Lemma A.2(c) we have

$$\text{pf } K = \det \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix} \cdot \text{pf} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = (-1)^m (\det A) \cdot (-1)^{m(m-1)/2}, \tag{A.79}$$

so that

$$\int \mathcal{D}(\psi, \bar{\psi}) e^{\bar{\psi}^T A \psi} = (-1)^{m(m+1)/2} \text{pf } K = \det A. \quad \square \tag{A.80}$$

Let us make one further observation concerning Gaussian integrals (both “real” and “complex”), which will be useful in proving certain aspects of Wick’s theorem. If $I = (i_1, \dots, i_k)$ is an arbitrary sequence of indices in $[n]$ – not necessarily distinct or ordered – then the “real” Gaussian integral (A.70) can be generalized to

$$\int d\chi_{i_k} \dots d\chi_{i_1} e^{\frac{1}{2} \chi_I^T A_{II} \chi_I} = \begin{cases} \text{pf } A_{II} & \text{if } k \text{ is even,} \\ 0 & \text{if } k \text{ is odd.} \end{cases} \tag{A.81}$$

To prove this, it suffices to observe, first of all, that if the indices i_1, \dots, i_k are distinct, then this formula is merely (A.70) after a relabeling of indices: the point is that by our definition (A.22), the variables arise in the *same* order in the integration measure and in the matrix A_{II} . On the other hand, if the sequence i_1, \dots, i_k contains any repeated index, then the left-hand side vanishes because $d\chi_i d\chi_i = 0$ by (A.49), while $\text{pf } A_{II} = 0$ in such a case because the pfaffian changes sign under simultaneous permutations of rows and columns [this is a special case of Lemma A.2(c)]. In a similar way, the “complex” Gaussian integral (A.74) can be generalized to

$$\int d\psi_{i_1} d\bar{\psi}_{j_1} \dots d\psi_{i_k} d\bar{\psi}_{j_k} e^{\bar{\psi}_J^T A_{JI} \psi_I} = \det A_{JI} \tag{A.82}$$

for arbitrary sequences $I = (i_1, \dots, i_k)$ and $J = (j_1, \dots, j_k)$.

We are now ready to state the full Wick’s theorem for fermions. Since the “source” version of Wick’s theorem for fermions will involve a *fermionic* source λ [cf. (A.83)], this means that we will be working (at least when discussing this equation) in an extended Grassmann algebra $R[\chi, \theta]$ with generators χ_1, \dots, χ_n and $\theta_1, \dots, \theta_N$ for some $N \geq 1$, and the sources λ_i will belong to the odd part of the Grassmann subalgebra $R[\theta]$.

We assume without further ado that n is *even*, i.e. $n = 2m$. Recall also from (A.22) the notation A_{IJ} for arbitrary sequences of indices I and J .

Theorem A.15 (Wick’s theorem for “real” fermions). *Let R be a commutative ring with identity element, and let $A = (a_{ij})_{i,j=1}^{2m}$ be a $2m \times 2m$ antisymmetric matrix with elements in R . Then:*

(a) If the matrix A is invertible, we have

$$\int \mathcal{D}\chi \exp\left(\frac{1}{2}\chi^T A \chi + \lambda^T \chi\right) = (\text{pf } A) \exp\left(\frac{1}{2}\lambda^T A^{-1} \lambda\right) \tag{A.83}$$

whenever $\lambda = (\lambda_i)_{i=1}^{2m}$ are odd elements of the Grassmann algebra that do not involve χ , i.e. $\lambda_i \in R[\theta]_{\text{odd}}$.

(b) For any subset $I = \{i_1, \dots, i_r\} \subseteq [2m]$ with $i_1 < \dots < i_r$, we have

$$\int \mathcal{D}\chi \chi_{i_1} \cdots \chi_{i_r} \exp\left(\frac{1}{2}\chi^T A \chi\right) = \begin{cases} 0 & \text{if } r \text{ is odd,} \\ \in(I) \text{ pf } A_{I^c I^c} & \text{if } r \text{ is even.} \end{cases} \tag{A.84}$$

(c) For any sequence of indices $I = (i_1, \dots, i_r)$ in $[2m]$, if the matrix A is invertible we have

$$\int \mathcal{D}\chi \chi_{i_1} \cdots \chi_{i_r} \exp\left(\frac{1}{2}\chi^T A \chi\right) = \begin{cases} 0 & \text{if } r \text{ is odd,} \\ (\text{pf } A) \text{ pf}((A^{-T})_{II}) & \text{if } r \text{ is even.} \end{cases} \tag{A.85}$$

(d) More generally, for any $r \times 2m$ matrix C with entries in R , we have

$$\int \mathcal{D}\chi (C\chi)_1 \cdots (C\chi)_r \exp\left(\frac{1}{2}\chi^T A \chi\right) = \begin{cases} 0 & \text{if } r \text{ is odd,} \\ \sum_{|I|=r} (\det C_{\star I}) \in(I) \text{ pf}(A_{I^c I^c}) & \text{if } r \text{ is even,} \end{cases} \tag{A.86}$$

where C_{\star} denotes the submatrix of C with columns in I ; and if the matrix A is invertible,

$$\int \mathcal{D}\chi (C\chi)_1 \cdots (C\chi)_r \exp\left(\frac{1}{2}\chi^T A \chi\right) = \begin{cases} 0 & \text{if } r \text{ is odd,} \\ (\text{pf } A) \text{ pf}(CA^{-T}C^T) & \text{if } r \text{ is even.} \end{cases} \tag{A.87}$$

Note that (A.84) and (A.85) are equivalent by virtue of Jacobi’s identity for pfaﬃans [cf. (A.21)], and that (A.86) and (A.87) are equivalent by virtue of Jacobi’s identity and the minor summation formula for pfaﬃans [cf. (A.20)]. However, we shall give *independent* proofs of all these formulae; as a consequence, our argument provides *ab initio* proofs of these pfaﬃan identities by means of Grassmann–Berezin integration.

Proof of Theorem A.15. To prove (a), we perform the translation $\chi' = \chi + A^{-1}\lambda$ and use Proposition A.12. We have

$$\begin{aligned} \frac{1}{2}\chi'^T A \chi' + \lambda^T \chi' &= \frac{1}{2}(\chi^T + \lambda^T A^{-T})A(\chi + A^{-1}\lambda) + \lambda^T(\chi + A^{-1}\lambda) \\ &= \frac{1}{2}\chi^T A \chi + \frac{1}{2}\lambda^T A^{-1} \lambda \end{aligned} \tag{A.88}$$

since $A^{-T} = -A^{-1}$ and $\chi^T \lambda = -\lambda^T \chi$. Therefore (A.83) is an immediate consequence of Proposition A.13.

For (b)–(d), let us first remark that if r is odd, then the integral must vanish, because the expansion of the exponential can only give an even number of factors of χ , so the “top” monomial in χ cannot be generated (recall that the total number of generators is even, i.e. $n = 2m$).

To prove (b), we compute

$$\begin{aligned}
 \int \mathcal{D}\chi \chi^I \exp\left(\frac{1}{2} \chi^T A \chi\right) &= \int \mathcal{D}\chi \chi^I \exp\left(\frac{1}{2} \chi_{I^c}^T A_{I^c I^c} \chi_{I^c}\right) \\
 &= \epsilon(I) \int \mathcal{D}\chi_I \mathcal{D}\chi_{I^c} \chi^I \exp\left(\frac{1}{2} \chi_{I^c}^T A_{I^c I^c} \chi_{I^c}\right) \\
 &= \epsilon(I) (-1)^{r(2m-r)} \left(\int \mathcal{D}\chi_I \chi^I \right) \left(\int \mathcal{D}\chi_{I^c} \exp\left(\frac{1}{2} \chi_{I^c}^T A_{I^c I^c} \chi_{I^c}\right) \right) \\
 &= \epsilon(I) \text{pf}(A_{I^c I^c}).
 \end{aligned} \tag{A.89}$$

The first equality holds because any factor ξ_i with $i \in I$ arising from the expansion of the exponential would be annihilated by the prefactor χ^I ; the second equality is simply a reordering of the integration variables; the third equality is Fubini’s theorem (Proposition A.11); and the last equality is simply the evaluation of the two integrals (using Proposition A.13 for the second) together with the fact that r is even.

(c) We specialize result (a) to the case where the “sources” λ_i are generators of an extended Grassmann algebra, then differentiate (or equivalently integrate!) with respect to the sources $\lambda_{i_1}, \dots, \lambda_{i_r}$, and finally replace all the λ_i by zero⁴⁴:

$$\begin{aligned}
 \int \mathcal{D}\chi \chi_{i_1} \cdots \chi_{i_r} \exp\left(\frac{1}{2} \chi^T A \chi\right) &= (\text{pf } A) \frac{\partial}{\partial \lambda_{i_1}} \cdots \frac{\partial}{\partial \lambda_{i_r}} \exp\left(\frac{1}{2} \lambda^T A^{-1} \lambda\right) \Big|_{\lambda=0} \\
 &= (\text{pf } A) \int d\lambda_{i_1} \cdots d\lambda_{i_r} \exp\left(\frac{1}{2} \lambda_I^T (A^{-1})_{II} \lambda_I\right) \\
 &= (\text{pf } A) (-1)^{r(r-1)/2} \int \mathcal{D}\lambda_I \exp\left(\frac{1}{2} \lambda_I^T (A^{-1})_{II} \lambda_I\right) \\
 &= (\text{pf } A) (-1)^{r(r-1)/2} \text{pf}((A^{-1})_{II}) \\
 &= (\text{pf } A) (-1)^{r(r-1)/2} (-1)^{r/2} \text{pf}((-A^{-1})_{II}) \\
 &= (\text{pf } A) \text{pf}((A^{-T})_{II}).
 \end{aligned} \tag{A.90}$$

Here the third equality involved reordering the integration variables from increasing to decreasing order; the fourth equality performed the Gaussian integral using (A.81); the fifth equality used (A.72); and the final equality used $-A^{-1} = A^{-T}$ and $(-1)^{r^2/2} = 1$ (which holds since r is even).

(d) We have

$$\begin{aligned}
 &\int \mathcal{D}\chi (C\chi)_1 \cdots (C\chi)_r \exp\left(\frac{1}{2} \chi^T A \chi\right) \\
 &= \sum_{i_1, \dots, i_r} c_{1i_1} \cdots c_{ri_r} \int \mathcal{D}\chi \chi_{i_1} \cdots \chi_{i_r} \exp\left(\frac{1}{2} \chi^T A \chi\right),
 \end{aligned} \tag{A.91}$$

but the only nonvanishing contributions in the sum come when i_1, \dots, i_r are all distinct; so we can first require $i_1 < \dots < i_r$ and then sum over permutations, yielding

⁴⁴ Setting $\lambda \rightarrow 0$ can be interpreted as extracting the monomials that do not involve λ , or alternatively as a special case of the substitution discussed at the end of Appendix A.3 (since 0 is odd).

$$\sum_{|I|=r} \sum_{\sigma \in S_r} C_{1i_{\sigma(1)}} \cdots C_{ri_{\sigma(r)}} \operatorname{sgn}(\sigma) \int \mathcal{D}\chi \chi^I \exp\left(\frac{1}{2} \chi^T A \chi\right) = \sum_{|I|=r} (\det C_{\star I}) \in(I) \operatorname{pf}(A_{I^c I^c}) \quad (\text{A.92})$$

by part (b). If A is invertible, we can write

$$\begin{aligned} \int \mathcal{D}\chi (C\chi)_1 \cdots (C\chi)_r \exp\left(\frac{1}{2} \chi^T A \chi\right) &= \frac{\partial}{\partial \lambda_{i_1}} \cdots \frac{\partial}{\partial \lambda_{i_r}} \int \mathcal{D}\chi \exp\left(\frac{1}{2} \chi^T A \chi + \lambda^T C\chi\right) \Big|_{\lambda=0} \\ &= \frac{\partial}{\partial \lambda_{i_1}} \cdots \frac{\partial}{\partial \lambda_{i_r}} (\operatorname{pf} A) \exp\left(\frac{1}{2} \lambda^T C A^{-1} C^T \lambda\right) \Big|_{\lambda=0} \end{aligned} \quad (\text{A.93})$$

by using (a). By the same reasoning as in (A.90) we can see that this equals $(\operatorname{pf} A) \operatorname{pf}(C A^{-1} C^T)$. \square

Next we state Wick’s theorem for “complex” fermions. Once again, when discussing the “source” version of this theorem [cf. (A.94)], we will work in an extended Grassmann algebra $R[\psi, \bar{\psi}, \theta]$ in which the sources $\lambda_i, \bar{\lambda}_i$ belong to the odd part of the Grassmann subalgebra $R[\theta]$.

Theorem A.16 (Wick’s theorem for “complex” fermions). *Let R be a commutative ring with identity element, and let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix with elements in R . Then:*

(a) *If the matrix A is invertible, we have*

$$\int \mathcal{D}(\psi, \bar{\psi}) \exp(\bar{\psi}^T A \psi + \bar{\lambda}^T \psi + \bar{\psi}^T \lambda) = (\det A) \exp(-\bar{\lambda}^T A^{-1} \lambda) \quad (\text{A.94})$$

whenever $\lambda = (\lambda_i)_{i=1}^n$ and $\bar{\lambda} = (\bar{\lambda}_i)_{i=1}^n$ are odd elements of the Grassmann algebra that do not involve ψ and $\bar{\psi}$, i.e. $\lambda_i, \bar{\lambda}_i \in R[\theta]_{\text{odd}}$.

(b) *For any subsets $I = \{i_1, \dots, i_r\}$ and $J = \{j_1, \dots, j_r\}$ of $[n]$ having the same cardinality r , with $i_1 < \dots < i_r$ and $j_1 < \dots < j_r$, we have*

$$\int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r \bar{\psi}_{i_\alpha} \psi_{j_\alpha} \right) \exp(\bar{\psi}^T A \psi) = \in(I, J) (\det A_{I^c J^c}). \quad (\text{A.95})$$

[If there is an unequal number of factors ψ and $\bar{\psi}$, then the integral is zero.]

(c) *For any sequences of indices $I = (i_1, \dots, i_r)$ and $J = (j_1, \dots, j_r)$ in $[n]$ of the same length r , if the matrix A is invertible we have*

$$\int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r \bar{\psi}_{i_\alpha} \psi_{j_\alpha} \right) \exp(\bar{\psi}^T A \psi) = (\det A) \det((A^{-1})_{IJ}). \quad (\text{A.96})$$

[Again, if there is an unequal number of factors ψ and $\bar{\psi}$, then the integral is zero.]

(d) *More generally, for any $r \times n$ matrix B and $n \times r$ matrix C with entries in R , we have*

$$\begin{aligned} &\int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r (\bar{\psi} C)_\alpha (B \psi)_\alpha \right) \exp(\bar{\psi}^T A \psi) \\ &= \sum_{|I|=|J|=r} \in(I, J) (\det B_{\star J}) (\det A_{I^c J^c}) (\det C_{I\star}), \end{aligned} \quad (\text{A.97})$$

and if the matrix A is invertible,

$$\int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r (\bar{\psi} C)_{\alpha} (B \psi)_{\alpha} \right) \exp(\bar{\psi}^T A \psi) = (\det A) \det(BA^{-1}C). \tag{A.98}$$

Note that (A.95) and (A.96) are equivalent by virtue of Jacobi’s identity (A.7), and that (A.97) and (A.98) are equivalent by virtue of Jacobi’s identity together with the Cauchy–Binet identity (A.3). However, we shall give *independent* proofs of all these formulae; as a consequence, our argument provides an *ab initio* proof of Jacobi’s identity by means of Grassmann–Berezin integration.

The proof of Theorem A.16 follows closely the pattern used in the proof of Theorem A.15, but with slightly different combinatorics.

Proof of Theorem A.16. To prove (a), we perform the translations $\psi' = \psi - A^{-1}\lambda$ and $\bar{\psi}' = \bar{\psi} - A^{-T}\bar{\lambda}$ and use Proposition A.12. We have

$$\begin{aligned} & \bar{\psi}'^T A \psi' + \bar{\lambda}^T \psi' + \bar{\psi}'^T \lambda \\ &= (\bar{\psi}^T - \bar{\lambda}^T A^{-1}) A (\psi - A^{-1}\lambda) + \bar{\lambda}^T (\psi - A^{-1}\lambda) + (\bar{\psi}^T - \bar{\lambda}^T A^{-1}) \lambda \\ &= \bar{\psi}^T A \psi - \bar{\lambda}^T A^{-1} \lambda. \end{aligned} \tag{A.99}$$

Therefore (A.94) is an immediate consequence of Proposition A.14.

To prove (b), we begin by observing as before that

$$\begin{aligned} & \int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r \bar{\psi}_{i_{\alpha}} \psi_{j_{\alpha}} \right) \exp(\bar{\psi}^T A \psi) \\ &= \int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r \bar{\psi}_{i_{\alpha}} \psi_{j_{\alpha}} \right) \exp(\bar{\psi}_{I^c}^T A_{I^c J^c} \psi_{J^c}). \end{aligned} \tag{A.100}$$

We now proceed to reorder the integration measure. We have $I = \{i_1, \dots, i_r\}$ with $i_1 < \dots < i_r$, and let us write $I^c = \{i'_1, \dots, i'_{n-r}\}$ with $i'_1 < \dots < i'_{n-r}$; and likewise for J . Then $\mathcal{D}(\psi, \bar{\psi})$ is the product of factors $d\psi_i d\bar{\psi}_i$ taken in arbitrary order; we choose the order to be II^c (i.e., $i_1 \dots i_r i'_1 \dots i'_{n-r}$). We now leave the factors $\bar{\psi}$ in place, but reorder the factors ψ to be in order JJ^c rather than II^c : this produces a sign $\epsilon(I, J)$. Using the notation

$$\mathcal{D}(\psi_J, \bar{\psi}_I) = d\psi_{j_1} d\bar{\psi}_{i_1} \dots d\psi_{j_r} d\bar{\psi}_{i_r} \tag{A.101}$$

and likewise for the complementary sets, we have proven that

$$\mathcal{D}(\psi, \bar{\psi}) = \epsilon(I, J) \mathcal{D}(\psi_J, \bar{\psi}_I) \mathcal{D}(\psi_{J^c}, \bar{\psi}_{I^c}). \tag{A.102}$$

This in turn can be trivially rewritten as $\epsilon(I, J) \mathcal{D}(\psi_{J^c}, \bar{\psi}_{I^c}) \mathcal{D}(\psi_J, \bar{\psi}_I)$ since the two factors \mathcal{D} are Grassmann-even. We can now apply this measure to the integrand in (A.100); integrating the monomial against $\mathcal{D}(\psi_J, \bar{\psi}_I)$ gives 1, and integrating the exponential against $\mathcal{D}(\psi_{J^c}, \bar{\psi}_{I^c})$ gives $\det A_{I^c J^c}$ by Proposition A.14.

(c) As in the case of “real” fermions, we specialize result (a) to the case where the “sources” $\lambda_i, \bar{\lambda}_i$ are generators of an extended Grassmann algebra, and then differentiate with respect to them:

$$\begin{aligned}
 & \int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r \bar{\psi}_{i_\alpha} \psi_{j_\alpha} \right) \exp(\bar{\psi}^T A \psi) \\
 &= (\det A) \left(\prod_{\alpha=1}^r -\frac{\partial}{\partial \lambda_{i_\alpha}} \frac{\partial}{\partial \bar{\lambda}_{j_\alpha}} \right) \exp(-\bar{\lambda}^T A^{-1} \lambda) \Big|_{\bar{\lambda}=\lambda=0} \\
 &= (\det A) (-1)^r \int \mathcal{D}(\lambda_I, \bar{\lambda}_J) \exp(-\bar{\lambda}_J^T (A^{-1})_{JJ} \lambda_I) \\
 &= (\det A) (-1)^r \det((-A^{-1})_{JJ}) \\
 &= (\det A) \det((A^{-T})_{JJ}). \tag{A.103}
 \end{aligned}$$

Here the minus sign in the first equality comes from the fact that differentiation of the source term $\bar{\psi}^T \lambda$ with respect to λ_i yields $-\bar{\psi}_i$ according to (A.51); the second equality says that differentiation is the same as integration; and the third equality uses (A.82).

(d) We have

$$\begin{aligned}
 & \int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r (\bar{\psi} C)_\alpha (B \psi)_\alpha \right) \exp(\bar{\psi}^T A \psi) \\
 &= \sum_{i_1, \dots, i_r} \sum_{j_1, \dots, j_r} \left(\prod_{\alpha=1}^r C_{i_\alpha, \alpha} B_{\alpha, j_\alpha} \right) \int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r \bar{\psi}_{i_\alpha} \psi_{j_\alpha} \right) \exp(\bar{\psi}^T A \psi), \tag{A.104}
 \end{aligned}$$

but the only nonvanishing contributions in the sum come when i_1, \dots, i_r , and also j_1, \dots, j_r , are all distinct; so we can first require $i_1 < \dots < i_r$, and $j_1 < \dots < j_r$ and then sum over permutations, yielding

$$\begin{aligned}
 & \sum_{|I|=r} \sum_{|J|=r} \sum_{\sigma, \tau \in S_r} C_{i_{\sigma(1)}, 1} \cdots C_{i_{\sigma(r)}, r} B_{1, j_{\tau(1)}} \cdots B_{r, j_{\tau(r)}} \\
 & \quad \times \operatorname{sgn}(\sigma) \operatorname{sgn}(\tau) \int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r \bar{\psi}_{i_\alpha} \psi_{j_\alpha} \right) \exp(\bar{\psi}^T A \psi) \\
 &= \sum_{|I|=r} \sum_{|J|=r} (\det B_{\star J}) \epsilon(I, J) (\det A_{J^c J^c}) (\det C_{I \star}) \tag{A.105}
 \end{aligned}$$

by part (b). If A is invertible, we can write

$$\begin{aligned}
 & \int \mathcal{D}(\psi, \bar{\psi}) \left(\prod_{\alpha=1}^r (\bar{\psi} C)_\alpha (B \psi)_\alpha \right) \exp(\bar{\psi}^T A \psi) \\
 &= \left(\prod_{i=1}^r -\frac{\partial}{\partial \lambda_i} \frac{\partial}{\partial \bar{\lambda}_i} \right) \int \mathcal{D}(\psi, \bar{\psi}) \exp(\bar{\psi}^T A \psi + \bar{\lambda}^T B \psi + \bar{\psi}^T C \lambda) \Big|_{\bar{\lambda}=\lambda=0} \\
 &= \left(\prod_{i=1}^r -\frac{\partial}{\partial \lambda_i} \frac{\partial}{\partial \bar{\lambda}_i} \right) (\det A) \exp(-\bar{\lambda}^T B A^{-1} C \lambda) \Big|_{\bar{\lambda}=\lambda=0} \tag{A.106}
 \end{aligned}$$

by using (a). Then by the same reasoning as in (A.103) we see that this equals $(\det A) \det(BA^{-1}C)$. \square

Table 1

Summary of the four types of Gaussian integration. Here A is the matrix appearing in the quadratic form in the exponential. See text for details.

Type of variables	Combinatorial conditions	Analytic conditions	Value of Gaussian int.	Contractions in $2k$ -point correlation
Real bosons	A symmetric	A pos.–def.	$(\det A)^{-1/2}$	$2k \times 2k$ hafnian
Complex bosons	none	$A + A^*$ pos.–def.	$(\det A)^{-1}$	$k \times k$ permanent
“Real” fermions	A antisymm.	none	$\text{pf } A$	$2k \times 2k$ pfaffian
“Complex” fermions	none	none	$\det A$	$k \times k$ determinant

A.6. Summary of bosonic and fermionic Gaussian integration

The four types of Gaussian integration and their basic characteristics are summarized in Table 1. Let us stress the following facts:

- Fermionic integration is a purely algebraic/combinatorial operation: no analytic conditions on the matrix A (such as positive–definiteness) are needed, and in fact fermionic integration makes sense over an arbitrary (commutative) coefficient ring R . Bosonic integration, by contrast, is an analytic operation (at least as we have defined it here) and requires positive–definiteness of A or its hermitian part.
- For complex bosons and “complex” fermions, no symmetry or antisymmetry conditions are imposed on the matrix A . By contrast, for real bosons and “real” fermions, the matrix A must be symmetric or antisymmetric, respectively.

Appendix B. Some useful identities

In this appendix we collect some auxiliary results that will be used at various places in this paper: identities for sums of products of binomial coefficients (Appendix B.1) and for determinants and pfaffians (Appendix B.2), lemmas on matrix factorization (Appendix B.3), and an identity that we call the “dilation–translation formula” (Appendix B.4).

B.1. Binomial identities

We collect here a few identities for sums of products of binomial coefficients that will be needed in the proof of the rectangular Cayley identities (Sections 5.5 and 5.6). We use the standard convention [47] for the definition of binomial coefficients:

$$\binom{r}{k} = \begin{cases} r(r-1)\cdots(r-k+1)/k! & \text{for integer } k > 0, \\ 1 & \text{for } k = 0, \\ 0 & \text{for integer } k < 0, \end{cases} \tag{B.1}$$

where r is an indeterminate and k is always an integer. Multinomial coefficients

$$\binom{a_1 + \cdots + a_k}{a_1, \dots, a_k} = \frac{(a_1 + \cdots + a_k)!}{a_1! \cdots a_k!} \tag{B.2}$$

will, by contrast, be used only when all the a_i are nonnegative integers.

We now state some easy combinatorial lemmas involving binomial coefficients. They are all either contained in the textbook of Graham, Knuth and Patashnik [47, Chapter 5] or derived in the same fashion.

Lemma B.1 (Parallel summation).

$$\sum_{k \leq m} \binom{r+k}{k} = \binom{r+m+1}{m} \text{ for } m \text{ integer.} \tag{B.3}$$

Lemma B.2 (Chu–Vandermonde convolution).

$$\sum_k \binom{w}{k} \binom{m}{p-k} = \binom{w+m}{p}. \tag{B.4}$$

Proof. For positive integers w and m , this is the number of ways of selecting p people out of a group of w women and m men. Since, for any fixed integer p , both sides are polynomials in w and m , the identity holds as a polynomial identity. \square

Lemma B.3.

$$\sum_{\substack{k,h,l \geq 0 \\ k+h+l=m}} (-1)^k \binom{a}{h} \binom{b}{k} \binom{h+l}{h} = \binom{a-b+m}{m}. \tag{B.5}$$

Proof. The left-hand side equals

$$\sum_{k=0}^m \sum_{h=0}^{m-k} (-1)^k \binom{a}{h} \binom{b}{k} \binom{m-k}{m-k-h} = \sum_{k=0}^m (-1)^k \binom{a+m-k}{m-k} \binom{b}{k} \tag{B.6a}$$

$$= \sum_{k=0}^m (-1)^m \binom{-a-1}{m-k} \binom{b}{k} \tag{B.6b}$$

$$= (-1)^m \binom{b-a-1}{m} \tag{B.6c}$$

$$= \binom{a-b+m}{m} \tag{B.6d}$$

where the first and third equalities use the Chu–Vandermonde convolution. \square

Lemma B.4.

$$\sum_{\substack{k,h,l \geq 0 \\ k+h+l \leq m}} (-1)^k \binom{a}{h} \binom{b}{k} \binom{h+l}{h} = \binom{a-b+m+1}{m}. \tag{B.7}$$

Proof. Sum (B.5) and use (B.3). \square

B.2. Determinant and pfaffian identities

We collect here some identities for determinants and pfaffians that will be needed in Section 5: these concern the determinant or pfaffian of a partitioned matrix and the change of determinant under low-rank perturbation. These identities are well known, but for completeness we will give compact proofs using Grassmann–Berezin integration. We will also give a (possibly new) “fermionic” analogue of the low-rank perturbation formula, which will play a crucial role throughout Section 5.

Let us begin with the formula for the determinant of a partitioned matrix, due to Schur [87, Hilfssatz, pp. 216–217]:

Proposition B.5 (Schur’s formula for the determinant of a partitioned matrix). Consider a partitioned matrix of the form

$$M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) \tag{B.8}$$

where A, B, C, D are matrices of sizes $m \times m, m \times n, n \times m$ and $n \times n$, respectively, with elements in a commutative ring with identity.

- (a) If A is invertible, then $\det M = (\det A) \det(D - CA^{-1}B)$.
- (b) If D is invertible, then $\det M = (\det D) \det(A - BD^{-1}C)$.

The matrix $D - CA^{-1}B$ is called the Schur complement of A in the partitioned matrix M ; see [24, 31,76,116] for reviews. One well-known proof of Schur’s formula is based on the identity

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) = \left(\begin{array}{c|c} A & 0 \\ \hline C & I_n \end{array} \right) \left(\begin{array}{c|c} I_m & A^{-1}B \\ \hline 0 & D - CA^{-1}B \end{array} \right). \tag{B.9}$$

Let us give a quick proof of Schur’s formula using Grassmann–Berezin integration:

Proof of Proposition B.5. Let R be the commutative ring with identity in which the elements of A, B, C, D take values. We introduce Grassmann variables $\psi_i, \bar{\psi}_i$ ($1 \leq i \leq m$) and $\eta_j, \bar{\eta}_j$ ($1 \leq j \leq n$) and work in the Grassmann algebra $R[\psi, \bar{\psi}, \eta, \bar{\eta}]$. We have

$$\begin{aligned} \det \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) &= \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \exp \left[(\bar{\psi}^T, \bar{\eta}^T) \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right) \begin{pmatrix} \psi \\ \eta \end{pmatrix} \right] \\ &= \int \mathcal{D}_m(\psi, \bar{\psi}) \mathcal{D}_n(\eta, \bar{\eta}) \exp[\bar{\psi}^T A \psi + \bar{\psi}^T B \eta + \bar{\eta}^T C \psi + \bar{\eta}^T D \eta]. \end{aligned} \tag{B.10}$$

If A is invertible, we can perform the integration over $\psi, \bar{\psi}$ using Wick’s theorem for “complex” fermions (Theorem A.16), yielding

$$(\det A) \int \mathcal{D}_n(\eta, \bar{\eta}) \exp(\bar{\eta}^T D \eta - \bar{\eta}^T C A^{-1} B \eta). \tag{B.11}$$

Then performing the integration over $\eta, \bar{\eta}$ yields $(\det A) \det(D - CA^{-1}B)$. This proves (a); and the proof of (b) is identical. □

Here are some important special cases of Proposition B.5:

Corollary B.6. Let U, V be $m \times n$ matrices with elements in a commutative ring with identity. Then

$$\det(UV^T) = \det\left(\begin{array}{c|c} 0_m & U \\ -V^T & I_n \end{array}\right), \quad (\text{B.12})$$

where 0_m is the $m \times m$ zero matrix and I_n is the $n \times n$ identity matrix.

Corollary B.7 (Matrix determinant lemma). Let A be an invertible $m \times m$ matrix, let W be an invertible $n \times n$ matrix, and let U, V be $m \times n$ matrices, all with elements in a commutative ring with identity. Then

$$\det(A + UWV^T) = (\det A)(\det W) \det(W^{-1} + V^T A^{-1} U). \quad (\text{B.13})$$

In particular, if we take $W = I_n$, then

$$\det(A + UV^T) = (\det A) \det(I_n + V^T A^{-1} U). \quad (\text{B.14})$$

If in addition we take $A = I_m$, then

$$\det(I_m + UV^T) = \det(I_n + V^T U). \quad (\text{B.15})$$

Corollary B.7 is sometimes known as the “matrix determinant lemma”, and the special case (B.15) is sometimes known as “Sylvester’s theorem for determinants”. When $n \ll m$, we can interpret (B.13)–(B.15) as formulae for the change of determinant under a low-rank perturbation: see Lemma B.11 below for an explicit statement.

Analogues of Proposition B.5 and Corollary B.6 exist also for pfaffians:

Proposition B.8 (Pfaffian of a partitioned matrix). Consider a partitioned matrix of the form

$$M = \left(\begin{array}{c|c} A & B \\ -B^T & D \end{array}\right) \quad (\text{B.16})$$

where A, B, D are matrices of sizes $2m \times 2m$, $2m \times 2n$ and $2n \times 2n$, respectively, with elements in a commutative ring with identity, and A and D are antisymmetric.

- (a) If A is invertible, then $\text{pf } M = (\text{pf } A) \text{pf}(D + B^T A^{-1} B)$.
 (b) If D is invertible, then $\text{pf } M = (\text{pf } D) \text{pf}(A + BD^{-1} B^T)$.

Note that the matrices $D + B^T A^{-1} B$ and $A + BD^{-1} B^T$ appearing here are just the usual Schur complements $D - CA^{-1} B$ and $A - BD^{-1} C$ specialized to $C = -B^T$.

Corollary B.9. Let U be a $2m \times 2n$ matrix with elements in a commutative ring with identity. Then

$$\text{pf}(U J_{2n} U^T) = (-1)^m \text{pf}(-U J_{2n} U^T) = (-1)^m \text{pf}\left(\begin{array}{c|c} 0_{2m} & U \\ -U^T & J_{2n} \end{array}\right) \quad (\text{B.17})$$

where J_{2n} is the standard $2n \times 2n$ symplectic form (A.15).

Proof of Proposition B.8. This time we introduce “real” Grassmann variables θ_i ($1 \leq i \leq 2m$) and λ_i ($1 \leq i \leq 2n$). We have

$$\text{pf } M = \int \mathcal{D}_{2m}(\theta) \mathcal{D}_{2n}(\lambda) \exp \left[\frac{1}{2} \theta^T A \theta + \theta^T B \lambda + \frac{1}{2} \lambda^T D \lambda \right]. \tag{B.18}$$

If A is invertible, we can perform the integration over θ using Wick’s theorem for “real” fermions (Theorem A.15), yielding

$$(\text{pf } A) \int \mathcal{D}_{2n}(\lambda) \exp \left(\frac{1}{2} \lambda^T D \lambda + \frac{1}{2} \lambda^T B^T A^{-1} B \lambda \right). \tag{B.19}$$

Then performing the integration over λ yields $(\text{pf } A) \text{pf}(D + B^T A^{-1} B)$. This proves (a); and the proof of (b) is identical. \square

We next wish to prove an analogue of Corollary B.7 when the entries in the various matrices belong, not to a commutative ring, but to a Grassmann algebra. More precisely, the entries in A and W will be *even* elements of the Grassmann algebra, while the entries in U and V will be *odd* elements of the Grassmann algebra.

Proposition B.10 (Fermionic matrix determinant lemma). *Let \mathcal{G} be a Grassmann algebra over a commutative ring with identity; let A be an invertible $m \times m$ matrix and W an invertible $n \times n$ matrix, whose elements belong to $\mathcal{G}_{\text{even}}$; and let U, V be $m \times n$ matrices whose elements belong to \mathcal{G}_{odd} . Then⁴⁵*

$$\det(A + U W V^T) = (\det A)(\det W)^{-1} \det(W^{-1} + V^T A^{-1} U)^{-1}. \tag{B.20}$$

In particular, if we take $W = I_n$, then

$$\det(A + U V^T) = (\det A) \det(I_n + V^T A^{-1} U)^{-1}. \tag{B.21}$$

If in addition we take $A = I_m$, then

$$\det(I_m + U V^T) = \det(I_n + V^T U)^{-1}. \tag{B.22}$$

Note that (B.20)–(B.22) differ from (B.13)–(B.15) by replacing some determinants with their inverse.

Proof of Proposition B.10. Let us first observe that since A and W are invertible, we can rewrite (B.20) as

$$\det(I_m + A^{-1} U W V^T) = \det(I_n + W V^T A^{-1} U)^{-1}, \tag{B.23}$$

which is equivalent to

$$\det(I_m + \tilde{U} \tilde{V}^T) = \det(I_n + \tilde{V}^T \tilde{U})^{-1} \tag{B.24}$$

under the (invertible) change of variables $\tilde{U} = A^{-1} U$, $\tilde{V} = V W^T$. So it suffices to prove (B.22).

⁴⁵ Note that all the matrix elements in all these determinants belong to the commutative ring $\mathcal{G}_{\text{even}}$; therefore, these determinants are unambiguously defined.

Let us first prove (B.22) when the ring R is \mathbb{R} . We augment the Grassmann algebra by introducing generators $\eta_j, \bar{\eta}_j$ ($1 \leq j \leq n$), and we also introduce bosonic variables $\varphi_i, \bar{\varphi}_i$ ($1 \leq i \leq m$). We wish to consider the mixed bosonic–fermionic integral

$$\int \mathcal{D}_m(\varphi, \bar{\varphi}) \mathcal{D}_n(\eta, \bar{\eta}) \exp[-\bar{\varphi}^T \varphi + \bar{\varphi}^T U \eta + \bar{\eta}^T V^T \varphi + \bar{\eta}^T \eta]. \tag{B.25}$$

Please note that the quantities $U\eta, \bar{\eta}^T V^T$ and $\bar{\eta}^T \eta$ all belong to the even “pure soul” part of the augmented Grassmann algebra, and in particular are nilpotent; therefore the integrand can be interpreted as

$$\exp(-\bar{\varphi}^T \varphi) \sum_{k=0}^{\infty} \frac{1}{k!} (\bar{\varphi}^T U \eta + \bar{\eta}^T V^T \varphi + \bar{\eta}^T \eta)^k \tag{B.26}$$

where the sum over k is in fact *finite*. The Gaussian integration over $\varphi, \bar{\varphi}$ can thus be interpreted as separate Gaussian integrations for the coefficients (which belong to \mathbb{R}) of each monomial in the augmented Grassmann algebra; this makes perfect analytic sense. Let us now evaluate (B.25) in two ways: Integrating first over $\varphi, \bar{\varphi}$ and then over $\eta, \bar{\eta}$, we get

$$\int \mathcal{D}_n(\eta, \bar{\eta}) \exp[\bar{\eta}^T \eta + \bar{\eta}^T V^T U \eta] = \det(I_n + V^T U). \tag{B.27}$$

On the other hand, integrating first over $\eta, \bar{\eta}$ and then over $\varphi, \bar{\varphi}$, we get

$$\int \mathcal{D}_m(\varphi, \bar{\varphi}) \exp[-\bar{\varphi}^T \varphi - \bar{\varphi}^T U V^T \varphi] = \det(I_m + U V^T)^{-1}. \tag{B.28}$$

This proves (B.22) when the ring R is \mathbb{R} .

Let us finally give an abstract argument showing that (B.22) holds for arbitrary commutative rings R . The point is that the matrix elements of U and V belong to a Grassmann algebra over some finite set of generators χ_1, \dots, χ_N , and hence can be written as

$$U_{ij} = \sum_{K \text{ odd}} \alpha_{ij;K} \chi^K, \tag{B.29a}$$

$$V_{ij} = \sum_{K \text{ odd}} \beta_{ij;K} \chi^K \tag{B.29b}$$

for some coefficients $\alpha_{ij;K}, \beta_{ij;K} \in R$. Now, both sides of (B.22) are of the form $\sum_{L \text{ even}} \gamma_L \chi^L$ where the coefficients γ_L are polynomials in $\{\alpha_{ij;K}, \beta_{ij;K}\}$ with integer coefficients.⁴⁶ But we have just shown that these two polynomials coincide whenever $\{\alpha_{ij;K}, \beta_{ij;K}\}$ are replaced by any set of specific values in \mathbb{R} . Therefore, they must coincide as *polynomials in the indeterminates* $\{\alpha_{ij;K}, \beta_{ij;K}\}$. But this implies that they are equal when $\{\alpha_{ij;K}, \beta_{ij;K}\}$ are replaced by specific values in *any* commutative ring R . \square

It is convenient, for our applications, to rephrase (B.15) and (B.22) as explicit formulae for the change of determinant under a low-rank perturbation. So let u_1, \dots, u_n and v_1, \dots, v_n be vectors of length m whose entries are elements of a Grassmann algebra and are either all Grassmann-even or

⁴⁶ The determinant on the right-hand side is an element of the Grassmann algebra whose “body” term is 1; therefore, it is invertible in the Grassmann algebra and the coefficients of its inverse are polynomials (with integer coefficients) in its own coefficients.

all Grassmann-odd. We call these cases $\epsilon = +1$ and $\epsilon = -1$, respectively. We then have the following formula for the determinant of a rank- n perturbation of the identity matrix:

Lemma B.11 (Low-rank perturbation lemma). *Let u_1, \dots, u_n and v_1, \dots, v_n be as above, and define the $n \times n$ matrix M by*

$$M_{\alpha\beta} = v_\alpha \cdot u_\beta \equiv \sum_{i=1}^m (v_\alpha)_i (u_\beta)_i. \tag{B.30}$$

Then we have

$$\det \left(I_m + \sum_{\alpha=1}^n u_\alpha v_\alpha^T \right) = \det(I_n + M)^\epsilon \tag{B.31}$$

where $\epsilon = \pm 1$ is as above.

Here is one special case of the low-rank perturbation lemma that will be useful in Sections 5.3 and 7.2 in treating symmetric Cayley identities:

Corollary B.12. *Let $\eta = (\eta_1, \dots, \eta_m)$ and $\bar{\eta} = (\bar{\eta}_1, \dots, \bar{\eta}_m)$ be Grassmann variables, and let A be an invertible $m \times m$ symmetric matrix whose elements are Grassmann-even (hence commute with everything). Then*

$$\det(I_m + \bar{\eta}\eta^T - A^{-1}\eta\bar{\eta}^T A) = (1 - \bar{\eta}^T \eta)^{-2}. \tag{B.32}$$

Proof. Applying Lemma B.11 with $u_1 = \bar{\eta}$, $v_1 = \eta$, $u_2 = A^{-1}\eta$ and $v_2 = -A^T \bar{\eta} = -A\bar{\eta}$ gives

$$\det(I_m + \bar{\eta}\eta^T - A^{-1}\eta\bar{\eta}^T A) = \det^{-1} \begin{pmatrix} 1 + \eta^T \bar{\eta} & \eta^T A^{-1} \eta \\ -\bar{\eta}^T A \bar{\eta} & 1 - \bar{\eta}^T \eta \end{pmatrix}. \tag{B.33}$$

Since A and A^{-1} are symmetric, the off-diagonal elements vanish, which gives the result. \square

Another special case of the low-rank perturbation lemma will arise in Section 5.9: here the $m \times m$ matrix $I_m + UV^T$ occurring on the left-hand side on (B.31) will be written as a product of rectangular matrices, each of which is a rank-1 perturbation of the corresponding rectangular pseudo-identity matrix. (The $m \times n$ pseudo-identity matrix \hat{I}_{mn} has matrix elements $(\hat{I}_{mn})_{ij} = \delta_{ij}$.) Direct application of the low-rank perturbation lemma to such a product matrix yields a rather messy result, but after some row operations we can obtain a fairly neat alternative formula:

Corollary B.13. *Fix integers $\ell \geq 1$ and $n_1, \dots, n_\ell \geq 1$ with $n_\alpha \geq n_1$ for $2 \leq \alpha \leq \ell$, and write $n_{\ell+1} = n_1$. Let x_1, \dots, x_ℓ and y_1, \dots, y_ℓ be vectors, where x_α is of length n_α and y_α is of length $n_{\alpha+1}$, whose entries are elements of a Grassmann algebra and are either all commuting ($\epsilon = +1$) or else all anticommuting ($\epsilon = -1$). Then we have*

$$\det \left(\prod_{\alpha=1}^{\ell} (\hat{I}_{n_\alpha n_{\alpha+1}} - x_\alpha y_\alpha^T) \right) = (\det N)^\epsilon \tag{B.34}$$

where the product is read from left ($\alpha = 1$) to right ($\alpha = \ell$), and the $\ell \times \ell$ matrix N is defined by

$$N_{\alpha\beta} = \begin{cases} \sum_{i=n_1+1}^{n_{\alpha+1,\beta}} y_i^\alpha x_i^\beta & \text{if } \alpha < \beta, \\ \delta_{\alpha\beta} - \sum_{i=1}^{n_1} y_i^\alpha x_i^\beta & \text{if } \alpha \geq \beta, \end{cases} \tag{B.35}$$

where $n_{\alpha,\beta} = \min_{\alpha \leq \gamma \leq \beta} n_\gamma$.

Proof. Note first that $\hat{I}_{n_\alpha n_{\alpha+1}} \hat{I}_{n_{\alpha+1} n_{\alpha+2}} \cdots \hat{I}_{n_{\beta-1} n_\beta}$ is an $n_\alpha \times n_\beta$ matrix whose ij element is 1 if $i = j \leq n_{\alpha,\beta}$ and 0 otherwise. In particular, $\hat{I}_{n_1 n_2} \hat{I}_{n_2 n_3} \cdots \hat{I}_{n_{\alpha-1} n_\alpha} = \hat{I}_{n_1 n_\alpha}$. For v a vector of length $m \geq n_1$, define \bar{v} as the vector restricted to the first n_1 components. So we can expand the matrix on the left-hand side of (B.34) as

$$\begin{aligned} & \prod_{\alpha=1}^{\ell} (\hat{I}_{n_\alpha n_{\alpha+1}} - x_\alpha y_\alpha^T) \\ &= I_{n_1} - \sum_{\alpha=1}^{\ell} \hat{I}_{n_1 n_2} \hat{I}_{n_2 n_3} \cdots \hat{I}_{n_{\alpha-1} n_\alpha} x_\alpha y_\alpha^T (\hat{I}_{n_{\alpha+1} n_{\alpha+2}} - x_{\alpha+1} y_{\alpha+1}^T) \cdots (\hat{I}_{n_\ell n_{\ell+1}} - x_\ell y_\ell^T) \\ &= I_{n_1} - \sum_{\alpha=1}^{\ell} \bar{x}_\alpha y_\alpha^T (\hat{I}_{n_{\alpha+1} n_{\alpha+2}} - x_{\alpha+1} y_{\alpha+1}^T) \cdots (\hat{I}_{n_\ell n_{\ell+1}} - x_\ell y_\ell^T). \end{aligned} \tag{B.36}$$

In this form, we are ready to apply Lemma B.11 with vectors $u_\alpha = -\bar{x}_\alpha$ and $v_\alpha^T = y_\alpha^T (\hat{I}_{n_{\alpha+1} n_{\alpha+2}} - x_{\alpha+1} y_{\alpha+1}^T) \cdots (\hat{I}_{n_\ell n_{\ell+1}} - x_\ell y_\ell^T)$ for $\alpha = 1, \dots, \ell$. This gives

$$\det \left(\prod_{\alpha=1}^{\ell} (\hat{I}_{n_\alpha n_{\alpha+1}} - x_\alpha y_\alpha^T) \right) = (\det N^{(0)})^\epsilon \tag{B.37}$$

with $N_{\alpha\beta}^{(0)} = \delta_{\alpha\beta} + v_\alpha^T u_\beta$. Now observe that

$$v_\alpha^T = \bar{y}_\alpha^T - \sum_{\beta > \alpha} (y_\alpha^T \hat{I}_{n_{\alpha+1} n_{\alpha+2}} \cdots \hat{I}_{n_{\beta-1} n_\beta} x_\beta) v_\beta^T \tag{B.38}$$

and call $c_{\alpha\beta} = y_\alpha^T \hat{I}_{n_{\alpha+1} n_{\alpha+2}} \cdots \hat{I}_{n_{\beta-1} n_\beta} x_\beta = \sum_{i=1}^{n_{\alpha+1,\beta}} y_i^\alpha x_i^\beta$. Thus, defining the $\ell \times \ell$ upper-triangular matrix \hat{C} as

$$\hat{C}_{\alpha\beta} = \begin{cases} c_{\alpha\beta} & \text{if } \alpha < \beta, \\ \delta_{\alpha\beta} & \text{if } \alpha = \beta, \\ 0 & \text{if } \alpha > \beta, \end{cases} \tag{B.39}$$

we have

$$\sum_{\beta} \hat{C}_{\alpha\beta} v_\beta^T = \bar{y}_\alpha^T. \tag{B.40}$$

Clearly $\det \hat{C} = 1$, so if we define $N = CN^{(0)}$ we have $\det N = \det N^{(0)}$. It is easy to see, using (B.40), that N is exactly the matrix given in (B.35). \square

Please note that all the entries of the matrix N are polynomials of degree at most two in the variables x and y – unlike the matrix $N^{(0)}$ coming from the bare application of the low-rank perturbation lemma, which contains terms of degree as high as 2ℓ .

Corollary B.6 is in fact the case $\ell = 2$ of a more general lemma that holds for $\ell \geq 2$, and will be needed in Section 5.9:

Lemma B.14. Fix integers $\ell \geq 2$ and $n_1, \dots, n_\ell \geq 1$, and write $n_{\ell+1} = n_1$. Let U_1, \dots, U_ℓ be matrices with elements in a commutative ring with identity, U_α being of dimension $n_\alpha \times n_{\alpha+1}$. Define

$$M_\ell(U_1, \dots, U_\ell) := \begin{pmatrix} 0_{n_1} & -U_1 & 0 & \dots & 0 \\ 0 & I_{n_2} & -U_2 & & 0 \\ 0 & 0 & I_{n_3} & \ddots & \vdots \\ \vdots & & & \ddots & -U_{\ell-1} \\ U_\ell & 0 & \dots & 0 & I_{n_\ell} \end{pmatrix}, \tag{B.41}$$

where 0_n is the $n \times n$ zero matrix and I_n is the $n \times n$ identity matrix. Then

$$\det M_\ell(U_1, \dots, U_\ell) = \det(U_1 \cdots U_\ell). \tag{B.42}$$

Proof. We prove this by induction on ℓ . The case $\ell = 2$ is already proven by Corollary B.6. For $\ell \geq 2$ we use the Grassmann representation of $\det(M_\ell)$:

$$\begin{aligned} \det M_\ell(U_1, \dots, U_\ell) &= \int \mathcal{D}_{n_1}(\psi^1, \bar{\psi}^1) \cdots \mathcal{D}_{n_\ell}(\psi^\ell, \bar{\psi}^\ell) \\ &\quad \times \exp \left[\sum_{\alpha=2}^{\ell} (\bar{\psi}^\alpha \psi^\alpha - \bar{\psi}^{\alpha-1} U_{\alpha-1} \psi^\alpha) + \bar{\psi}^\ell U_\ell \psi^1 \right]. \end{aligned} \tag{B.43}$$

Now perform the integration over $(\psi^\ell, \bar{\psi}^\ell)$: highlighting the factors in the integrand that involve these fields, we see that

$$\int \mathcal{D}_{n_\ell}(\psi^\ell, \bar{\psi}^\ell) \exp[\bar{\psi}^\ell \psi^\ell - \bar{\psi}^{\ell-1} U_{\ell-1} \psi^\ell + \bar{\psi}^\ell U_\ell \psi^1] = \exp(\bar{\psi}^{\ell-1} U_{\ell-1} U_\ell \psi^1). \tag{B.44}$$

Comparing this with (B.43), we see that

$$\det M_\ell(U_1, \dots, U_\ell) = \det M_{\ell-1}(U_1, \dots, U_{\ell-2}, U_{\ell-1} U_\ell), \tag{B.45}$$

which provides the required inductive step. \square

Remark. Just as Corollary B.6 is a specialization of the more general Proposition B.5, so Lemma B.14 has a similar generalization, proven through an identical procedure (of which the details are left to the reader), namely:

Lemma B.15. Fix integers $\ell \geq 2$ and $n_1, \dots, n_\ell \geq 1$, and write $n_{\ell+1} = n_1$. Let B_1, \dots, B_ℓ be matrices with elements in a commutative ring with identity, B_α being of dimension $n_\alpha \times n_{\alpha+1}$. Let A_1, \dots, A_ℓ be square matrices with elements in the same commutative ring, A_α being of dimension $n_\alpha \times n_\alpha$. Assume that A_2, \dots, A_ℓ are invertible. Define

$$M(A_1, \dots, A_\ell; B_1, \dots, B_\ell) := \begin{pmatrix} A_1 & -B_1 & 0 & \cdots & 0 \\ 0 & A_2 & -B_2 & & 0 \\ 0 & 0 & A_3 & \ddots & \vdots \\ \vdots & & & \ddots & -B_{\ell-1} \\ B_\ell & 0 & \cdots & 0 & A_\ell \end{pmatrix}. \quad (\text{B.46})$$

Then

$$\det M(A_1, \dots, A_\ell; B_1, \dots, B_\ell) = \det(A_1 + B_1 A_2^{-1} B_2 \cdots A_\ell^{-1} B_\ell) \prod_{j=2}^{\ell} \det A_j. \quad (\text{B.47})$$

If also A_1 is invertible, then we can obtain an expression with a form of cyclic symmetry:

$$\det M(A_1, \dots, A_\ell; B_1, \dots, B_\ell) = \det(I_{n_1} + A_1^{-1} B_1 A_2^{-1} B_2 \cdots A_\ell^{-1} B_\ell) \prod_{j=1}^{\ell} \det A_j. \quad (\text{B.48})$$

B.3. Matrix factorization lemmas

In Sections 5.4, 5.5, 5.6, 5.7 and 5.9 we shall need some matrix factorization lemmas having the general form:

For any matrix X of the form there exists a matrix A of the form such that $\Phi(X, A) = 0$ [where Φ denotes a specified collection of polynomial or rational functions];

or the multi-matrix generalization thereof:

For any matrices X, Y, \dots of the form there exist matrices A, B, \dots of the form such that $\Phi(X, Y, \dots, A, B, \dots) = 0$.

The prototype for such matrix decomposition lemmas is the well-known Cholesky factorization [44, Theorem 4.2.5]:

Lemma B.16 (Cholesky factorization). *Let X be a real symmetric positive-definite $n \times n$ matrix. Then there exists a unique lower-triangular real matrix A with strictly positive diagonal entries, such that $X = AA^T$.*

A similar but less well-known result is the following factorization for antisymmetric matrices [6,12,71,112,115]:

Lemma B.17 (AJA^T factorization of an antisymmetric matrix). *Let X be a (real or complex) antisymmetric $2m \times 2m$ matrix. Then there exists a (real or complex, respectively) $2m \times 2m$ matrix A such that $X = AJA^T$, where J is defined in (A.15). In particular, if X is nonsingular, then $A \in GL(2m)$. [The form of A can be further restricted in various ways, but we shall not need this.]*

In our applications we shall not need the uniqueness of A , but merely its existence. Nor shall we need any particular structure of A (e.g. triangularity) beyond lying in $GL(n)$ or $O(n)$ or $Sp(2n)$ as the case may be. Finally, and most importantly, we shall not need the existence of A for all matrices X of a given type, but only for those in some nonempty open set (for instance, a small neighborhood of the identity matrix). We shall therefore give easy existence proofs using the implicit function theorem. It is an interesting open question whether our decompositions actually extend to arbitrary matrices X in the given classes.

More precisely, we shall need the following decomposition lemmas in addition to Lemma B.17. The matrices \hat{I}_{mn} are defined in (5.64).

Lemma B.18. Let X and Y be a (real or complex) $m \times n$ matrices ($m \leq n$) of rank m that are sufficiently close to the matrix \hat{I}_{mn} . Then there exist matrices $P, R \in GL(m)$ and $Q \in GL(n)$ such that $X = P\hat{I}_{mn}Q$ and $Y = R\hat{I}_{mn}Q^{-T}$.

Lemma B.19. Let X be a (real or complex) $m \times n$ matrix ($m \leq n$) of rank m that is sufficiently close to the matrix \hat{I}_{mn} . Then there exist matrices $P \in GL(m)$ and $Q \in O(n)$ such that $X = P\hat{I}_{mn}Q$.

Lemma B.20. Let X be a (real or complex) $2m \times 2n$ matrix ($m \leq n$) of rank $2m$ that is sufficiently close to the matrix $\hat{I}_{2m,2n}$. Then there exist matrices $P \in GL(2m)$ and $Q \in Sp(2n)$ such that $X = P\hat{I}_{2m,2n}Q$.

Lemma B.21. Let $\ell \geq 1$ and $n_1, \dots, n_{\ell+1} \geq 1$; and let $\{X_\alpha\}_{1 \leq \alpha \leq \ell}$ be (real or complex) $n_\alpha \times n_{\alpha+1}$ matrices of rank $\min(n_\alpha, n_{\alpha+1})$ that are sufficiently close to the matrix $\hat{I}_{n_\alpha n_{\alpha+1}}$. Then there exist matrices $\{P_\alpha\}_{1 \leq \alpha \leq \ell+1}$ with $P_\alpha \in GL(n_\alpha)$ such that $X_\alpha = P_\alpha \hat{I}_{n_\alpha n_{\alpha+1}} P_{\alpha+1}^{-1}$.

Lemmas B.17–B.21 will be needed in Sections 5.4, 5.5, 5.6, 5.7 and 5.9, respectively. In addition, Cholesky factorization could be used in Section 5.3 but we were able to avoid it; see the Remark at the end of that section.

The proofs of these lemmas will all follow the same pattern. First we find an explicit pair X_0, A_0 (or the multi-matrix generalization) with the needed properties. Then we linearize the functions Φ in a neighborhood of (X_0, A_0) , and we show that the tangent space for A at A_0 is mapped onto the full tangent space for X at X_0 . The required existence of A for X in a neighborhood of X_0 then follows from the implicit function theorem.⁴⁷ For completeness we will also show how Lemmas B.16 and B.17, which are known to hold globally, have simple proofs in this “infinitesimal” setting. Nearly all these proofs will be easy; only the last (Lemma B.21) turns out to be slightly tricky.

Proof of Lemma B.16 for X near I . Linearizing $X = AA^T$ in a neighborhood of $(X_0, A_0) = (I, I)$ by writing $X = I + X'$ and $A = I + A'$, we have $X' = A' + (A')^T$ (plus higher-order corrections that we always drop). Then an explicit solution is given by the lower-triangular matrix

$$A'_{ij} = \begin{cases} X'_{ij}/2 & \text{if } i = j, \\ X'_{ij} & \text{if } i > j, \\ 0 & \text{if } i < j. \end{cases} \tag{B.49}$$

The rest follows from the implicit function theorem.⁴⁸ □

⁴⁷ We will use the implicit function theorem in the following form: Let $U \subseteq \mathbb{R}^N$ and $V \subseteq \mathbb{R}^p$ be open sets, and let $f : U \times V \rightarrow \mathbb{R}^N$ be a C^k function ($k \geq 1$). Let $u_0 \in U$ and $v_0 \in V$ satisfy $f(u_0, v_0) = 0$, with $(\partial f / \partial u)(u_0, v_0)$ nonsingular. Then there exist neighborhoods $U' \ni u_0$ and $V' \ni v_0$ such that for all $v \in V'$ there exists a unique $u \in U'$ satisfying $f(u, v) = 0$; moreover, the map $v \mapsto u$ is C^k .

⁴⁸ For completeness let us make explicit how the implicit function theorem is used in this case; the analogous reasoning for the remaining lemmas can be supplied by the reader.

If $M = (m_{ij})_{i,j=1}^n$ is an $n \times n$ matrix, let us write $[M]_{\text{LT}} = (m_{ij})_{1 \leq i \leq j \leq n}$ to denote its lower-triangular part. We then use the implicit function theorem (see the preceding footnote) as follows: Let $u = A$ be a generic lower-triangular matrix $(a_{ij})_{1 \leq i \leq j \leq n}$, $v = [X]_{\text{LT}}$, $u_0 = v_0 = [I]_{\text{LT}}$, $f(u, v) = [AA^T - X]_{\text{LT}}$. (Since $AA^T - X$ is manifestly symmetric, it vanishes if and only if its lower-triangular part does.) Then $(\partial f / \partial u)(u_0, v_0)$ is the linear map $A' \mapsto A' + (A')^T$; in other words, we have

$$\frac{\partial f_{ij}}{\partial u_{kl}}(u_0, v_0) = \delta_{(ij),(kl)}(1 + \delta_{kl}),$$

which is a diagonal matrix with nonzero entries (namely, 1 and 2), hence nonsingular.

Proof of Lemma B.17 for X near J. Linearizing $X = AJA^T$ in a neighborhood of $(X_0, A_0) = (J, I)$ by writing $X = J + X'$ and $A = I - A'J$, we have $X' = A' - (A')^T$. Then one explicit solution is given by the strictly lower-triangular matrix

$$A'_{ij} = \begin{cases} X'_{ij} & \text{if } i > j, \\ 0 & \text{if } i \leq j. \end{cases} \tag{B.50}$$

Another explicit solution is given by the antisymmetric matrix $A' = X'/2$. \square

Proof of Lemma B.18. Linearizing $X = P\hat{I}_{mn}Q$ and $Y = R\hat{I}_{mn}Q^{-T}$ in a neighborhood of $(X_0, Y_0, P_0, R_0, Q_0) = (\hat{I}_{mn}, \hat{I}_{mn}, I_m, I_m, I_n)$ by writing $X = \hat{I}_{mn} + X'$ and so forth, we obtain $X' = P'\hat{I}_{mn} + \hat{I}_{mn}Q'$ and $Y' = R'\hat{I}_{mn} - \hat{I}_{mn}(Q')^T$. In terms of the block decompositions $\hat{I}_{mn} \equiv (I_m, 0_{m \times (n-m)})$, $X' = (X'_1, X'_2)$, $Y' = (Y'_1, Y'_2)$ and $Q' = \begin{pmatrix} Q'_{11} & Q'_{12} \\ Q'_{21} & Q'_{22} \end{pmatrix}$, we have

$$X'_1 = P' + Q'_{11}, \tag{B.51a}$$

$$X'_2 = Q'_{12}, \tag{B.51b}$$

$$Y'_1 = R' - (Q'_{11})^T, \tag{B.51c}$$

$$Y'_2 = -(Q'_{21})^T. \tag{B.51d}$$

We can choose Q'_{11} and Q'_{22} arbitrarily; then the remaining unknowns P', R', Q'_{12}, Q'_{21} are uniquely determined. \square

We shall actually prove the following generalization of Lemma B.19:

Lemma B.19'. Let X be a (real or complex) $m \times n$ matrix ($m \leq n$) of rank m that is sufficiently close to the matrix \hat{I}_{mn} , and let Y be a (real or complex) $n \times n$ symmetric matrix that is sufficiently close to the identity matrix I_n . Then there exist matrices $P \in GL(m)$ and $Q \in GL(n)$ such that $X = P\hat{I}_{mn}Q$ and $Y = Q^TQ$.

When $Y = I_n$ this reduces to Lemma B.19.

Proof of Lemma B.19'. Linearizing $X = P\hat{I}_{mn}Q$ and $Y = Q^TQ$ in a neighborhood of $(X_0, Y_0, P_0, Q_0) = (\hat{I}_{mn}, I_n, I_m, I_n)$ by writing $X = \hat{I}_{mn} + X'$ and so forth, we obtain $X' = P'\hat{I}_{mn} + \hat{I}_{mn}Q'$ and $Y' = Q' + (Q')^T$. In terms of the block decompositions $\hat{I}_{mn} \equiv (I_m, 0_{m \times (n-m)})$, $X' = (X'_1, X'_2)$, $Y' = \begin{pmatrix} Y'_{11} & Y'_{12} \\ (Y'_{12})^T & Y'_{22} \end{pmatrix}$ with Y'_{11} and Y'_{22} symmetric, and $Q' = \begin{pmatrix} Q'_{11} & Q'_{12} \\ Q'_{21} & Q'_{22} \end{pmatrix}$, we have

$$X'_1 = P' + Q'_{11}, \tag{B.52a}$$

$$X'_2 = Q'_{12}, \tag{B.52b}$$

$$Y'_{11} = Q'_{11} + (Q'_{11})^T, \tag{B.52c}$$

$$Y'_{22} = Q'_{22} + (Q'_{22})^T, \tag{B.52d}$$

$$Y'_{12} = Q'_{12} + (Q'_{21})^T. \tag{B.52e}$$

We can choose arbitrarily the antisymmetric parts of Q'_{11} and Q'_{22} (e.g. by taking Q'_{11} and Q'_{22} lower-triangular, or alternatively by taking Q'_{11} and Q'_{22} symmetric); then the remaining unknowns are uniquely determined. \square

Similarly, let us prove the following generalization of Lemma B.20:

Lemma B.20'. Let X be a (real or complex) $2m \times 2n$ matrix ($m \leq n$) of rank $2m$ that is sufficiently close to the matrix $\hat{I}_{2m,2n}$, and let Y be a (real or complex) $2n \times 2n$ antisymmetric matrix that is sufficiently close to J_{2n} . Then there exist matrices $P \in GL(2m)$ and $Q \in GL(2n)$ such that $X = P\hat{I}_{2m,2n}Q$ and $Y = Q^T J Q$.

When $Y = J_{2n}$ this reduces to Lemma B.20.

Proof of Lemma B.20'. Linearizing $X = P\hat{I}_{2m,2n}Q$ and $Y = Q^T J Q$ in a neighborhood of $(X_0, Y_0, P_0, Q_0) = (\hat{I}_{2m,2n}, J_{2n}, I_{2m}, I_{2n})$ by writing $X = \hat{I}_{2m,2n} + X'$, $Y = J_{2n} + Y'$, $P = I_{2m} + P'$ and $Q = I_{2n} - J_{2n}Q'$, we obtain $X' = P'\hat{I}_{2m,2n} - \hat{I}_{2m,2n}J_{2n}Q'$ and $Y = Q' - (Q')^T$. In terms of the block decompositions $\hat{I}_{mn} \equiv (I_m, 0_{m \times (n-m)})$, $X' = (X'_1, X'_2)$, $Y' = \begin{pmatrix} Y'_{11} & Y'_{12} \\ -(Y'_{12})^T & Y'_{22} \end{pmatrix}$ with Y'_{11} and Y'_{22} antisymmetric, and $Q' = \begin{pmatrix} Q'_{11} & Q'_{12} \\ Q'_{21} & Q'_{22} \end{pmatrix}$, we have

$$X'_1 = P' - J_{2m}Q'_{11}, \tag{B.53a}$$

$$X'_2 = -J_{2m}Q'_{12}, \tag{B.53b}$$

$$Y'_{11} = Q'_{11} - (Q'_{11})^T, \tag{B.53c}$$

$$Y'_{22} = Q'_{22} - (Q'_{22})^T, \tag{B.53d}$$

$$Y'_{12} = Q'_{12} - (Q'_{21})^T. \tag{B.53e}$$

We can choose arbitrarily the symmetric parts of Q'_{11} and Q'_{22} ; then the remaining unknowns are uniquely determined. \square

In preparation for the proof of Lemma B.21, it is convenient to introduce some simple notation for decomposing rectangular matrices. Given an $m \times n$ matrix Y , we define the strictly lower-triangular $m \times m$ matrix $\mathcal{L}(Y)$ by taking the strictly lower-triangular part of Y and either deleting the last $n - m$ columns (if $m < n$) or appending $m - n$ columns of zeros (if $m > n$). Likewise, we define the upper-triangular $n \times n$ matrix $\mathcal{U}(Y)$ by taking the upper-triangular part of Y and either deleting the last $m - n$ rows (if $m > n$) or appending $n - m$ rows of zeros (if $m < n$). It follows immediately from these definitions that

$$Y = \mathcal{L}(Y)\hat{I}_{mn} + \hat{I}_{mn}\mathcal{U}(Y). \tag{B.54}$$

Proof of Lemma B.21. Linearizing $X_\alpha = P_\alpha \hat{I}_{n_\alpha n_{\alpha+1}} P_{\alpha+1}^{-1}$ in a neighborhood of $X_\alpha = \hat{I}_{n_\alpha n_{\alpha+1}}$ and $P_\alpha = I_{n_\alpha}$ by writing $X_\alpha = \hat{I}_{n_\alpha n_{\alpha+1}} + X'_\alpha$ and $P_\alpha = I_{n_\alpha} + P'_\alpha$, we obtain $X'_\alpha = P'_\alpha \hat{I}_{n_\alpha n_{\alpha+1}} - \hat{I}_{n_\alpha n_{\alpha+1}} P'_{\alpha+1}$.

Let us decompose each P'_α as a sum of its strictly lower-triangular part L_α and its upper-triangular part U_α . We therefore need to solve the equations

$$X'_\alpha = (L_\alpha + U_\alpha)\hat{I}_{n_\alpha n_{\alpha+1}} - \hat{I}_{n_\alpha n_{\alpha+1}}(L_{\alpha+1} + U_{\alpha+1}), \tag{B.55}$$

where X_1, \dots, X_ℓ are considered as parameters and $L_1, \dots, L_{\ell+1}, U_1, \dots, U_{\ell+1}$ are considered as unknowns. But let us prove a bit more, namely that the matrices U_1 and $L_{\ell+1}$ can be considered as parameters (i.e. can be chosen arbitrarily). Note first that the system (B.55) can be solved by

$$L_\alpha = \mathcal{L}(X'_\alpha - U_\alpha \hat{I}_{n_\alpha n_{\alpha+1}} + \hat{I}_{n_\alpha n_{\alpha+1}} L_{\alpha+1}), \tag{B.56a}$$

$$U_{\alpha+1} = \mathcal{U}(X'_\alpha - U_\alpha \hat{I}_{n_\alpha n_{\alpha+1}} + \hat{I}_{n_\alpha n_{\alpha+1}} L_{\alpha+1}) \tag{B.56b}$$

in the sense that any solution of (B.56) provides a solution of (B.55). Eqs. (B.56) appear at first glance to be entangled, i.e. L_α depends on U_α and vice versa. But this is only an appearance, because the operators \mathcal{L} and \mathcal{U} “see”, respectively, only the strictly-lower-triangular and upper-triangular parts of the matrix on which they act. Therefore, the system (B.56) can be rewritten as

$$L_\alpha = \mathcal{L}(X'_\alpha + \hat{I}_{n_\alpha n_{\alpha+1}} L_{\alpha+1}), \tag{B.57a}$$

$$U_{\alpha+1} = \mathcal{U}(X'_\alpha - U_\alpha \hat{I}_{n_\alpha n_{\alpha+1}}). \tag{B.57b}$$

But these latter equations can manifestly be solved sequentially for L_ℓ, \dots, L_1 (given $L_{\ell+1}$) and for $U_2, \dots, U_{\ell+1}$ (given U_1). \square

Remark. It is instructive to count parameters and variables in Lemma B.21. The parameters are X_1, \dots, X_ℓ , and their number is $N_p = \sum_{\alpha=1}^{\ell} n_\alpha n_{\alpha+1}$. (The matrices U_1 and $L_{\ell+1}$, which can be chosen arbitrarily, do not count as extra parameters because they merely redefine the matrices X_1 and X_ℓ , respectively.) The variables are L_1, \dots, L_ℓ and $U_2, \dots, U_{\ell+1}$, and their number is

$$N_v = \frac{n_1(n_1 - 1)}{2} + n_2^2 + \dots + n_\ell^2 + \frac{n_{\ell+1}(n_{\ell+1} + 1)}{2}. \tag{B.58}$$

Therefore

$$N_v - N_p = \sum_{\alpha=1}^{\ell} \frac{(n_\alpha - n_{\alpha+1})(n_\alpha - n_{\alpha+1} - 1)}{2} \geq 0. \tag{B.59}$$

The $N_v - N_p$ extra variables were fixed by our choice of the operators \mathcal{L} and \mathcal{U} : we decided to append $m - n$ columns of zeros to \mathcal{L} when $m > n$, and $n - m$ rows of zeros to \mathcal{U} when $m < n$, but we could equally well have inserted an arbitrary strictly-lower-triangular matrix of size $m - n$ into \mathcal{L} and an arbitrary upper-triangular matrix of size $n - m$ into \mathcal{U} . This gives $(m - n)(m - n - 1)/2$ additional variables in both cases, which precisely accounts for $N_v - N_p$.

B.4. Dilation–translation formula

In Sections 6.1 and 6.2 we will need the following well-known generalization of the translation formula (5.1):

Lemma B.22 (Dilation–translation formula). *Let $P(z)$ be a polynomial in a single indeterminate z , with coefficients in a commutative ring R containing the rationals, and let a and b be indeterminates. Then*

$$\exp\left((a + bz) \frac{\partial}{\partial z}\right) P(z) = P\left(e^b z + \frac{e^b - 1}{b} a\right) \tag{B.60}$$

as an identity in the ring $R[[z, a]][[b]]$ of formal power series in b whose coefficients are polynomials in z and a . [Here the exponential is defined by its Taylor series, as are e^b and $(e^b - 1)/b$.]

In particular, this identity can be evaluated at any nilpotent element $b \in R$, as both sides then reduce to finite sums.

Remark. If b is nilpotent of order 2 (i.e. $b^2 = 0$), then the formula simplifies further to

$$\exp\left((a + bz) \frac{\partial}{\partial z}\right) P(z) = P\left((1 + b)z + \left(1 + \frac{b}{2}\right)a\right). \tag{B.61}$$

In our applications in Sections 6.1 and 6.2 we will have $b^2 = 0$ and also $ba = 0$, in which case the identity holds even when the ring R does not contain the rationals.

Proof of Lemma B.22. When $a = 0$, the formula (B.60) states the well-known fact that the operator $z \partial / \partial z$ generates dilations; it is easily checked by applying both sides to z^n . The general case is handled by the change of variables $w = z + a/b$. \square

Remark. The formula (B.60) is a special case of a more general formula for operators of the form $\exp[tg(z) \partial / \partial z]$:

$$\exp\left(tg(z) \frac{\partial}{\partial z}\right) P(z) = P(\tilde{z}(t; z)) \quad (\text{B.62})$$

where $\tilde{z}(t; z)$ is the solution of the differential equation $d\tilde{z}(t; z)/dt = g(\tilde{z}(t; z))$ with initial condition $\tilde{z}(0; z) = z$.

References

- [1] A. Abdesselam, Feynman diagrams in algebraic combinatorics, *Sém. Lothar. Combin.* B49 (2003), article B49c, arXiv:math/0212121 [math.CO].
- [2] A. Abdesselam, Grassmann–Berezin calculus and theorems of the matrix-tree type, *Adv. Appl. Math.* 33 (2004) 51–70, arXiv:math/0306396 [math.CO].
- [3] A. Abdesselam, T. Crilly, A.D. Sokal, The tangled history of the “Cayley” identity $\det(\partial)(\det X)^s = s(s+1)\cdots(s+n-1)(\det X)^{s-1}$, in preparation.
- [4] H. Aslaksen, Quaternionic determinants, *Math. Intelligencer* 18 (1996) 57–65.
- [5] M.F. Atiyah, Resolution of singularities and division of distributions, *Comm. Pure Appl. Math.* 23 (1970) 145–150.
- [6] P. Benner, R. Byers, H. Fassbender, V. Mehrmann, D. Watkins, Cholesky-like factorizations of skew-symmetric matrices, *Electron. Trans. Numer. Anal.* 11 (2000) 85–93; Addendum: *Electron. Trans. Numer. Anal.* 11 (2000) 93A.
- [7] I.N. Bernšteĭn, The analytic continuation of generalized functions with respect to a parameter, *Funkcional. Anal. i Priložen.* 6 (1972) 26–40; *Funct. Anal. Appl.* 6 (1972) 273–285.
- [8] I.N. Bernšteĭn, S.I. Gelfand, Meromorphy of the function P^λ , *Funkcional. Anal. i Priložen.* 3 (1969) 84–85; *Funct. Anal. Appl.* 3 (1969) 68–69.
- [9] J.-E. Björk, *Rings of Differential Operators*, North-Holland, Amsterdam, Oxford, New York, 1979.
- [10] P.M. Blekher, Integration of functions in a space with a complex number of dimensions, *Teoret. Mat. Fiz.* 50 (1982) 370–382; English translation in: *Theoret. and Math. Phys.* 50 (1982) 243–251.
- [11] B. Blind, Distributions homogènes sur une algèbre de Jordan, *Bull. Soc. Math. France* 125 (1997) 493–528.
- [12] J.R. Bunch, A note on the stable decomposition of skew-symmetric matrices, *Math. Comp.* 38 (1982) 475–479.
- [13] E.R. Caianiello, *Combinatorics and Renormalization in Quantum Field Theory*, Benjamin, Reading, MA, London, Amsterdam, 1973.
- [14] E.R. Canfield, S.G. Williamson, Hook length products and Cayley operators of classical invariant theory, *Linear Multilinear Algebra* 9 (1981) 289–297.
- [15] A. Capelli, Fondamenti di una teoria generale delle forme algebriche, *Atti Reale Accad. Lincei, Mem. Classe Sci. Fis. Mat. Nat. (Serie 3)* 12 (1882) 529–598.
- [16] A. Capelli, Ueber die Zurückführung der Cayley’schen Operation Ω auf gewöhnliche Polar-Operationen, *Math. Ann.* 29 (1887) 331–338.
- [17] A. Capelli, Ricerca delle operazioni invariantive fra più serie di variabili permutabili con ogni altra operazione invariantiva fra le stesse serie, *Atti Reale Accad. Sci. Fis. Mat. Napoli (Serie 2)* 1 (1888) 1–17.
- [18] A. Capelli, Sur les opérations dans la théorie des formes algébriques, *Math. Ann.* 37 (1890) 1–37.
- [19] A. Capelli, *Lezioni sulla Teoria delle Forme Algebriche*, Pellerano, Napoli, 1902.
- [20] S. Caracciolo, J.L. Jacobsen, H. Saleur, A.D. Sokal, A. Sportiello, Fermionic field theory for trees and forests, *Phys. Rev. Lett.* 93 (2004) 080601, arXiv:cond-mat/0403271.
- [21] S. Caracciolo, A.D. Sokal, A. Sportiello, Grassmann integral representation for spanning hyperforests, *J. Phys. A* 40 (2007) 13799–13835, arXiv:0706.1509 [math-ph].
- [22] S. Caracciolo, A.D. Sokal, A. Sportiello, Noncommutative determinants, Cauchy–Binet formulae, and Capelli-type identities. I. Generalizations of the Capelli and Turnbull identities, *Electron. J. Combin.* 16 (1) (2009), #R103, arXiv:0809.3516 [math.CO].
- [23] S. Caracciolo, A.D. Sokal, A. Sportiello, in preparation.
- [24] D. Carlson, What are Schur complements, anyway?, *Linear Algebra Appl.* 74 (1986) 257–275.
- [25] A. Cayley, On linear transformations, *Cambridge and Dublin Math. J.* 1 (1846) 104–122; Also in: *The Collected Mathematical Papers of Arthur Cayley*, vol. 1, Cambridge University Press, Cambridge, 1889–1897, pp. 95–112.

- [26] A. Cayley, *The Collected Mathematical Papers of Arthur Cayley*, 13 vols., Cambridge University Press, Cambridge, 1889–1897; Also republished by Johnson Reprint Corp., New York, 1963.
- [27] S. Chaiken, A combinatorial proof of the all minors matrix-tree theorem, *SIAM J. Alg. Disc. Meth.* 3 (1982) 319–329.
- [28] A. Clebsch, Ueber symbolische Darstellung algebraischer Formen, *J. Reine Angew. Math.* 59 (1861) 1–62.
- [29] A. Clebsch, *Theorie der binären algebraischen Formen*, B.G. Teubner, Leipzig, 1872.
- [30] P.M. Cohn, *Basic Algebra: Groups, Rings and Fields*, Springer-Verlag, London, Berlin, Heidelberg, 2003.
- [31] R.W. Cottle, Manifestations of the Schur complement, *Linear Algebra Appl.* 8 (1974) 189–211.
- [32] S.C. Coutinho, *A Primer of Algebraic D-Modules*, London Math. Soc. Stud. Texts, vol. 33, Cambridge University Press, Cambridge, 1995.
- [33] M. Creutz, On invariant integration over $SU(N)$, *J. Math. Phys.* 19 (1978) 2043–2046.
- [34] I. Dolgachev, *Lectures on Invariant Theory*, London Math. Soc. Lecture Note Ser., vol. 296, Cambridge University Press, Cambridge, 2003.
- [35] P. Etingof, Note on dimensional regularization, in: P. Deligne, et al. (Eds.), *Quantum Fields and Strings: A Course for Mathematicians*, vol. 1, American Mathematical Society, Providence, RI, 1999, pp. 597–607.
- [36] J. Faraut, A. Korányi, *Analysis on Symmetric Cones*, Oxford University Press, Oxford, New York, 1994 (Chapter VII).
- [37] D. Foata, D. Zeilberger, Combinatorial proofs of Capelli's and Turnbull's identities from classical invariant theory, *Electron. J. Combin.* 1 (1994), #R1.
- [38] M. Fulmek, Graphical condensation, overlapping pfaffians and superpositions of matchings, *Electron. J. Combin.* 17 (2010), #R83.
- [39] W. Fulton, J. Harris, *Representation Theory: A First Course*, Springer-Verlag, New York, Heidelberg, 1991 (Appendix F).
- [40] W. Fulton, P. Pragacz, *Schubert Varieties and Degeneracy Loci*, Lecture Notes in Math., vol. 1689, Springer-Verlag, Berlin, Heidelberg, 1998 (Appendix D).
- [41] L. Gårding, Extension of a formula by Cayley to symmetric determinants, *Proc. Edinb. Math. Soc.* 8 (1948) 73–75.
- [42] I. Gelfand, Some aspects of functional analysis and algebra, in: *Proceedings of the International Congress of Mathematicians*, vol. 1, Amsterdam, 1954, Noordhoff/North-Holland, Groningen/Amsterdam, 1957, pp. 253–276.
- [43] I.M. Gel'fand, G.E. Shilov, *Generalized Functions*, vol. 1, Academic Press, New York, London, 1964.
- [44] G.H. Golub, C.F. Van Loan, *Matrix Computations*, 3rd edition, The Johns Hopkins University Press, Baltimore, 1996.
- [45] P. Gordan, Ueber Combinanten, *Math. Ann.* 5 (1872) 95–122.
- [46] J.H. Grace, A. Young, *The Algebra of Invariants*, Cambridge University Press, Cambridge, 1903, pp. 259–260.
- [47] R.L. Graham, D.E. Knuth, O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd edition, Addison-Wesley, Reading, MA, 1994.
- [48] A.M. Hamel, Pfaffian identities: a combinatorial approach, *J. Combin. Theory Ser. A* 94 (2001) 205–217.
- [49] H. Hironaka, Resolution of singularities of an algebraic variety over a field of characteristic zero, *Ann. Math.* 79 (1964) 109–326.
- [50] R. Howe, Remarks on classical invariant theory, *Trans. Amer. Math. Soc.* 313 (1989) 539–570; Erratum: *Trans. Amer. Math. Soc.* 318 (1990) 823.
- [51] R. Howe, T. Umeda, The Capelli identity, the double commutant theorem, and multiplicity-free actions, *Math. Ann.* 290 (1991) 565–619.
- [52] J.-i. Igusa, *An Introduction to the Theory of Local Zeta Functions*, AMS/IP Stud. Adv. Math., vol. 14, American Mathematical Society, Providence, RI, 2000.
- [53] M. Ishikawa, M. Wakayama, Minor summation formula of pfaffians, *Linear Multilinear Algebra* 39 (1995) 285–305.
- [54] M. Ishikawa, M. Wakayama, Applications of minor summation formula. III. Plücker relations, lattice paths and pfaffian identities, *J. Combin. Theory Ser. A* 113 (2006) 113–155.
- [55] L. Isserlis, On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables, *Biometrika* 12 (1918) 134–139.
- [56] M. Kashiwara, B -functions and holonomic systems. Rationality of roots of B -functions, *Invent. Math.* 38 (1976/1977) 33–53.
- [57] S.P. Khèkalo, Riesz potentials in the space of rectangular matrices, and the iso-Huygens deformation of the Cayley–Laplace operator, *Dokl. Akad. Nauk* 376 (2001) 168–170; English translation in: *Dokl. Math.* 63 (2001) 35–37.
- [58] S.P. Khèkalo, The Cayley–Laplace differential operator on the space of rectangular matrices, *Izv. Ross. Akad. Nauk Ser. Mat.* 69 (2005) 195–224; English translation in: *Izv. Math.* 69 (2005) 191–219.
- [59] T. Kimura, The b -functions and holonomy diagrams of irreducible regular prehomogeneous vector spaces, *Nagoya Math. J.* 85 (1982) 1–80.
- [60] T. Kimura, *Introduction to Prehomogeneous Vector Spaces*, Transl. Math. Monogr., vol. 215, American Mathematical Society, Providence, RI, 2003.
- [61] K. Kinoshita, M. Wakayama, Explicit Capelli identities for skew symmetric matrices, *Proc. Edinb. Math. Soc.* 45 (2002) 449–465.
- [62] D.E. Knuth, Overlapping pfaffians, *Electron. J. Combin.* 3 (2) (1996), #R5.
- [63] B. Kostant, S. Sahi, The Capelli identity, tube domains, and the generalized Laplace transform, *Adv. Math.* 87 (1991) 71–92.
- [64] G.R. Krause, T.H. Lenagan, *Growth of Algebras and Gelfand–Kirillov Dimension*, revised edition, American Mathematical Society, Providence, RI, 2000.
- [65] S. Lang, *Algebra*, revised 3rd edition, Springer-Verlag, New York, 2002.
- [66] A. Leykin, Constructibility of the set of polynomials with a fixed Bernstein–Sato polynomial: an algorithmic approach, *J. Symbolic Comput.* 32 (2001) 663–675.
- [67] A. Leykin, private communication, 2006.
- [68] H. Minc, *Permanents*, *Encyclopedia Math. Appl.*, vol. 6, Addison-Wesley, Reading, MA, 1978.

- [69] J.W. Moon, Some determinant expansions and the matrix-tree theorem, *Discrete Math.* 124 (1994) 163–171.
- [70] T. Muir, *The Theory of Determinants in the Historical Order of Development*, 4 vols., Macmillan, London, 1906–1923.
- [71] F.D. Murnaghan, A. Wintner, A canonical form for real matrices under orthogonal transformations, *Proc. Natl. Acad. Sci. USA* 17 (1931) 417–420.
- [72] M. Muro, Singular invariant hyperfunctions on the space of real symmetric matrices, *Tohoku Math. J.* 51 (1999) 329–364.
- [73] M. Muro, Singular invariant hyperfunctions on the space of complex and quaternion Hermitian matrices, *J. Math. Soc. Japan* 53 (2001) 589–602.
- [74] M. Muro, Singular invariant hyperfunctions on the square matrix space and the alternating matrix space, *Nagoya Math. J.* 169 (2003) 19–75.
- [75] P.J. Olver, *Classical Invariant Theory*, London Math. Soc. Stud. Texts, vol. 44, Cambridge University Press, Cambridge, 1999.
- [76] D.V. Ouellette, Schur complements and statistics, *Linear Algebra Appl.* 36 (1981) 187–295.
- [77] I. Ozeki, On the microlocal structure of a regular prehomogeneous vector space associated with $GL(8)$, *Proc. Japan Acad. Ser. A Math. Sci.* 56 (1980) 18–21.
- [78] I. Ozeki, On the microlocal structure of the regular prehomogeneous vector space associated with $SL(5) \times GL(4)$, *Publ. Res. Inst. Math. Sci. Kyoto Univ.* 26 (1990) 539–584.
- [79] V.V. Prasolov, *Problems and Theorems in Linear Algebra*, Transl. Math. Monogr., vol. 134, American Mathematical Society, Providence, RI, 1994.
- [80] M. Raïs, Distributions homogènes sur des espaces de matrices, *Mém. Soc. Math. Fr.* 30 (1972) 3–109.
- [81] F. Ricci, E.M. Stein, Homogeneous distributions on spaces of Hermitean matrices, *J. Reine Angew. Math.* 368 (1986) 142–164.
- [82] H. Rubenthaler, G. Schiffmann, Opérateurs différentiels de Shimura et espaces préhomogènes, *Invent. Math.* 90 (1987) 409–442.
- [83] B. Rubin, Riesz potentials and integral geometry in the space of rectangular matrices, *Adv. Math.* 205 (2006) 549–598.
- [84] M. Saito, Multiplier ideals, b -function, and spectrum of a hypersurface singularity, *Compos. Math.* 143 (2007) 1050–1068.
- [85] M. Sato, M. Kashiwara, T. Kimura, T. Oshima, Micro-local analysis of prehomogeneous vector spaces, *Invent. Math.* 62 (1980) 117–179.
- [86] F. Sato, K. Sugiyama, Multiplicity one property and the decomposition of b -functions, *Internat. J. Math.* 17 (2006) 195–229.
- [87] J. [Isaai] Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, *J. Reine Angew. Math.* 147 (1917) 205–232; English translation: On power series which are bounded in the interior of the unit circle. I, in: I. Gohberg (Ed.), *I. Schur Methods in Operator Theory and Signal Processing*, Birkhäuser, Basel, 1986, pp. 31–59.
- [88] I. Schur, *Vorlesungen über Invariantentheorie*, Springer-Verlag, Berlin, Heidelberg, New York, 1968.
- [89] G. Shimura, On differential operators attached to certain representations of classical groups, *Invent. Math.* 77 (1984) 463–488.
- [90] A.D. Sokal, The multivariate Tutte polynomial (alias Potts model) for graphs and matroids, in: B.S. Webb (Ed.), *Surveys in Combinatorics, 2005*, Cambridge University Press, Cambridge, New York, 2005, pp. 173–226, arXiv:math/0503607 [math.CO].
- [91] E.R. Speer, Dimensional and analytic renormalization, in: G. Velo, A.S. Wightman (Eds.), *Renormalization Theory, Proceedings of the NATO Advanced Study Institute, Erice, 1975*, in: NATO Adv. Study Inst. Ser. C: Math. Phys. Sci., vol. 23, Reidel, Dordrecht, 1976, pp. 25–93.
- [92] R.P. Stanley, *Enumerative Combinatorics*, vol. 1, Cambridge University Press, Cambridge, New York, 1997.
- [93] P.R. Stein, On an identity from classical invariant theory, *Linear Multilinear Algebra* 11 (1982) 39–44.
- [94] J.R. Stembridge, Nonintersecting paths, pfaffians, and plane partitions, *Adv. Math.* 83 (1990) 96–131 (Section 2 and Lemma 4.2).
- [95] R.F. Streater, A.S. Wightman, *PCT, Spin and Statistics, and All That*, Benjamin/Cummings, Reading, MA, 1964; 2nd edition, Benjamin/Cummings, Reading, MA, 1978.
- [96] K. Sugiyama, b -Functions of regular simple prehomogeneous vector spaces, *Tsukuba J. Math.* 26 (2002) 407–451.
- [97] K. Sugiyama, b -Function of a prehomogeneous vector space with no regular component, *Comment. Math. Univ. St. Pauli* 54 (2005) 99–119.
- [98] K. Sugiyama, b -Functions associated with quivers of type A, *Transform. Groups* 16 (2011) 1183–1222, arXiv:1005.3596 [math.RT].
- [99] U. Tamm, The determinant of a Hessenberg matrix and some applications in discrete mathematics, <http://www.mathematik.uni-bielefeld.de/ahlschwede/pub/tamm/hessen.ps>, undated.
- [100] U. Tamm, Majorization in lattice path enumeration and creating order, University of Bielefeld preprint 00-108, <http://www.math.uni-bielefeld.de/sfb343/preprints/pr00108.ps.gz>, 2000.
- [101] H.W. Turnbull, *The Theory of Determinants, Matrices, and Invariants*, Blackie & Son, London, Glasgow, 1928, pp. 114–116; 2nd edition, Blackie & Son, London, Glasgow, 1945; 3rd edition, Dover, New York, 1960.
- [102] H.W. Turnbull, Symmetric determinants and the Cayley and Capelli operators, *Proc. Edinb. Math. Soc.* 8 (1948) 76–86.
- [103] H.W. Turnbull, Note upon the generalized Cayleyan operator, *Canad. J. Math.* 1 (1949) 48–56.
- [104] K. Ukai, b -Functions of prehomogeneous vector spaces of Dynkin–Kostant type for exceptional groups, *Compos. Math.* 135 (2003) 49–101.
- [105] T. Umeda, The Capelli identities, a century after, *Amer. Math. Soc. Transl.* 183 (1998) 51–78.
- [106] G. Vivanti, Alcune formole relative all'operazione Ω , *Rend. Circ. Mat. Palermo* 4 (1890) 261–268.
- [107] S. Wakatsuki, b -Functions of regular 2-simple prehomogeneous vector spaces associated to the symplectic group and the orthogonal group, *Comment. Math. Univ. St. Pauli* 53 (2004) 121–137.
- [108] A.H. Wallace, A note on the Capelli operators associated with a symmetric matrix, *Proc. Edinb. Math. Soc.* 9 (1953) 7–12.

- [109] U. Walther, Bernstein–Sato polynomial versus cohomology of the Milnor fiber for generic hyperplane arrangements, *Compos. Math.* 141 (2005) 121–145.
- [110] R. Weitzenböck, *Invariantentheorie*, P. Noordhoff, Groningen, 1923, pp. 15–16.
- [111] H. Weyl, *The Classical Groups, Their Invariants and Representations*, 2nd edition, Princeton University Press, Princeton, NJ, 1946.
- [112] H. Weyl, *The Theory of Groups and Quantum Mechanics*, Dover, New York, 1950 (Appendix 3).
- [113] G.C. Wick, The evaluation of the collision matrix, *Phys. Rev.* 80 (1950) 268–272.
- [114] S.G. Williamson, Generic common minor expansions, *Linear Multilinear Algebra* 20 (1987) 253–279 (Section 4).
- [115] H. Xu, An SVD-like matrix decomposition and its applications, *Linear Algebra Appl.* 368 (2003) 1–24.
- [116] F. Zhang (Ed.), *The Schur Complement and Its Applications*, Numer. Methods Algorithms, vol. 4, Springer-Verlag, New York, 2005.
- [117] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, 4th edition, Clarendon Press, Oxford, 2002.