Noncommutative determinants, Cauchy–Binet formulae, and Capelli-type identities

I. Generalizations of the Capelli and Turnbull identities

Sergio Caracciolo
Dipartimento di Fisica and INFN
Università degli Studi di Milano
via Celoria 16
I-20133 Milano, ITALY

Sergio.Caracciolo@mi.infn.it

Alan D. Sokal*
Department of Physics
New York University
4 Washington Place
New York, NY 10003 USA

sokal@nyu.edu

Andrea Sportiello
Dipartimento di Fisica and INFN
Università degli Studi di Milano
via Celoria 16
I-20133 Milano, ITALY

Andrea.Sportiello@mi.infn.it

Submitted: Sep 20, 2008; Accepted: Aug 3, 2009; Published: Aug 7, 2009
Mathematics Subject Classification: 15A15 (Primary); 05A19, 05A30, 05E15, 13A50, 15A24, 15A33, 15A72, 17B35, 20G05 (Secondary).

Abstract

We prove, by simple manipulation of commutators, two noncommutative generalizations of the Cauchy–Binet formula for the determinant of a product. As special cases we obtain elementary proofs of the Capelli identity from classical invariant theory and of Turnbull's Capelli-type identities for symmetric and antisymmetric matrices.

Key Words: Determinant, noncommutative determinant, row-determinant, column-determinant, Cauchy–Binet theorem, permanent, noncommutative ring, Capelli identity, Turnbull identity, Cayley identity, classical invariant theory, representation theory, Weyl algebra, right-quantum matrix, Cartier–Foata matrix, Manin matrix.

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

1 Introduction

Let R be a commutative ring, and let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix with elements in R. Define as usual the determinant

$$\det A \;:=\; \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)} \;. \tag{1.1}$$

One of the first things one learns about the determinant is the *multiplicative property*:

$$\det(AB) \;=\; (\det A)(\det B) \;. \tag{1.2}$$

More generally, if A and B are $m \times n$ matrices, and I and J are subsets of $[n] := \{1, 2, \ldots, n\}$ of cardinality $|I| = |J| = r$, then one has the *Cauchy–Binet formula*:

$$\det\,(A^{\mathrm{T}}B)_{IJ} \;=\; \sum_{\substack{L \subseteq [m] \\ |L| = r}} (\det\,(A^{\mathrm{T}})_{IL})(\det B_{LJ}) \tag{1.3a}$$

$$\;=\; \sum_{\substack{L \subseteq [m] \\ |L| = r}} (\det A_{LI})(\det B_{LJ}) \tag{1.3b}$$

where M_{IJ} denotes the submatrix of M with rows I and columns J (kept in their original order).

If one wants to generalize these formulae to matrices with elements in a *noncommutative* ring R, the first problem one encounters is that the definition (1.1) is ambiguous without an ordering prescription for the product. Rather, one can define numerous alternative "determinants": for instance, the *column-determinant*

$$\operatorname{col-det} A \;:=\; \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1)1}\, a_{\sigma(2)2} \,\cdots\, a_{\sigma(n)n} \tag{1.4}$$

and the *row-determinant*

$$\operatorname{row-det} A \;:=\; \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \,\cdots\, a_{n\sigma(n)} \;. \tag{1.5}$$

(Note that $\operatorname{col-det} A = \operatorname{row-det} A^{\mathrm{T}}$.) Of course, in the absence of commutativity these "determinants" need not have all the usual properties of the determinant.

Our goal here is to prove the analogues of (1.2)/(1.3) for a fairly simple noncommutative case: namely, that in which the elements of A are in a suitable sense "almost commutative" among themselves (see below) and/or the same for B, while the commutators $[x, y] := xy - yx$ of elements of A with those of B have the simple structure $[a_{ij}, b_{kl}] = -\delta_{ik} h_{jl}$.[1] More precisely, we shall need the following type of commutativity among the elements of A and/or B:

[1]The minus sign is inserted for future convenience. We remark that this formula makes sense even if the ring R lacks an identity element, as $\delta_{ik} h_{jl}$ is simply a shorthand for h_{jl} if $i = k$ and 0 otherwise.

Definition 1.1 *Let $M = (M_{ij})$ be a (not-necessarily-square) matrix with elements in a (not-necessarily-commutative) ring R. Then we say that M is* column-pseudo-commutative *in case*

$$[M_{ij}, M_{kl}] \;=\; [M_{il}, M_{kj}] \qquad \text{for all } i,j,k,l \tag{1.6}$$

and

$$[M_{ij}, M_{il}] \;=\; 0 \qquad \text{for all } i,j,l\;. \tag{1.7}$$

We say that M is row-pseudo-commutative *in case M^{T} is column-pseudo-commutative.*

In Sections 2 and 3 we will explain the motivation for this strange definition, and show that it really is the natural type of commutativity for formulae of Cauchy–Binet type.[2] Suffice it to observe now that column-pseudo-commutativity is a fairly weak condition: for instance, it is weaker than assuming that $[M_{ij}, M_{kl}] = 0$ whenever $j \neq l$. In many applications (though not all, see Example 3.6 below) we will actually have $[a_{ij}, a_{kl}] = [b_{ij}, b_{kl}] = 0$ for *all* i,j,k,l. Note also that (1.6) implies (1.7) if the ring R has the property that $2x = 0$ implies $x = 0$.

The main result of this paper is the following:

Proposition 1.2 (noncommutative Cauchy–Binet) *Let R be a (not-necessarily-commutative) ring, and let A and B be $m \times n$ matrices with elements in R. Suppose that*

$$[a_{ij}, b_{kl}] \;=\; -\delta_{ik} h_{jl} \tag{1.8}$$

where $(h_{jl})_{j,l=1}^{n}$ are elements of R. Then, for any $I, J \subseteq [n]$ of cardinality $|I| = |J| = r$:

(a) If A is column-pseudo-commutative, then

$$\sum_{\substack{L \subseteq [m] \\ |L| = r}} (\text{col-det}\,(A^{\mathrm{T}})_{IL})(\text{col-det}\,B_{LJ}) \;=\; \text{col-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\text{col}}] \tag{1.9}$$

where

$$(Q_{\text{col}})_{\alpha\beta} \;=\; (r-\beta)\, h_{i_\alpha j_\beta} \tag{1.10}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

(b) If B is column-pseudo-commutative, then

$$\sum_{\substack{L \subseteq [m] \\ |L| = r}} (\text{row-det}\,(A^{\mathrm{T}})_{IL})(\text{row-det}\,B_{LJ}) \;=\; \text{row-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\text{row}}] \tag{1.11}$$

where

$$(Q_{\text{row}})_{\alpha\beta} \;=\; (\alpha-1)\, h_{i_\alpha j_\beta} \tag{1.12}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

---

[2] Similar notions arose already two decades ago in Manin's work on quantum groups [38–40]. For this reason, some authors [15] call a row-pseudo-commutative matrix a *Manin matrix*; others [30–32] call it a *right-quantum matrix*. See the historical remarks at the end of Section 2.

In particular,

(c) *If* $[a_{ij}, a_{kl}] = 0$ *and* $[b_{ij}, b_{kl}] = 0$ *whenever* $j \neq l$*, then*

$$\sum_{\substack{L \subseteq [m] \\ |L| = r}} (\det (A^{\mathrm{T}})_{IL})(\det B_{LJ}) \;=\; \text{col-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}}] \qquad (1.13a)$$

$$\;=\; \text{row-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{row}}] \qquad (1.13b)$$

These identities can be viewed as a kind of "quantum analogue" of (1.3), with the matrices Q_{col} and Q_{row} supplying the "quantum correction". It is for this reason that we have chosen the letter h to designate the matrix arising in the commutator.

Please note that the hypotheses of Proposition 1.2 presuppose that $1 \leqslant r \leqslant n$ (otherwise I and J would be nonexistent or empty). But $r > m$ is explicitly allowed: in this case the left-hand side of (1.9)/(1.11)/(1.13) is manifestly zero (since the sum over L is empty), but Proposition 1.2 makes the nontrivial statement that the noncommutative determinant on the right-hand side is also zero.

Note also that the hypothesis in part (c) — what we shall call *column-commutativity*, see Section 2 — is sufficient to make the determinants of A and B well-defined without any ordering prescription. We have therefore written det (rather than col-det or row-det) for these determinants.

Replacing A and B by their transposes and interchanging m with n in Proposition 1.2, we get the following "dual" version in which the commutator $-\delta_{ik} h_{jl}$ is replaced by $-h_{ik}\delta_{jl}$:

Proposition 1.2′ *Let* R *be a (not-necessarily-commutative) ring, and let* A *and* B *be* $m \times n$ *matrices with elements in* R*. Suppose that*

$$[a_{ij}, b_{kl}] \;=\; -h_{ik}\delta_{jl} \qquad (1.14)$$

where $(h_{ik})_{i,k=1}^{m}$ *are elements of* R*. Then, for any* $I, J \subseteq [m]$ *of cardinality* $|I| = |J| = r$*:*

(a) *If* A *is row-pseudo-commutative, then*

$$\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\text{col-det}\, A_{IL})(\text{col-det}\, (B^{\mathrm{T}})_{LJ}) \;=\; \text{col-det}[(AB^{\mathrm{T}})_{IJ} + Q_{\mathrm{col}}] \qquad (1.15)$$

where Q_{col} *is defined in (1.10).*

(b) *If* B *is row-pseudo-commutative, then*

$$\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\text{row-det}\, A_{IL})(\text{row-det}\, (B^{\mathrm{T}})_{LJ}) \;=\; \text{row-det}[(AB^{\mathrm{T}})_{IJ} + Q_{\mathrm{row}}] \qquad (1.16)$$

where Q_{row} *is defined in (1.12).*

In particular,

(c) If $[a_{ij}, a_{kl}] = 0$ *and* $[b_{ij}, b_{kl}] = 0$ *whenever* $i \neq k$*, then*

$$\begin{aligned} \sum_{\substack{L \subseteq [n] \\ |L| = r}} (\det A_{IL})(\det (B^{\mathrm{T}})_{LJ}) &= \text{col-det}[(AB^{\mathrm{T}})_{IJ} + Q_{\mathrm{col}}] && (1.17a) \\ &= \text{row-det}[(AB^{\mathrm{T}})_{IJ} + Q_{\mathrm{row}}] && (1.17b) \end{aligned}$$

When the commutator has the special form $[a_{ij}, b_{kl}] = -h\delta_{ik}\delta_{jl}$, then *both* Propositions 1.2 and 1.2′ apply, and by summing (1.13)/(1.17) over $I = J$ of cardinality r, we obtain:

Corollary 1.3 *Let* R *be a (not-necessarily-commutative) ring, and let* A *and* B *be* $m \times n$ *matrices with elements in* R*. Suppose that*

$$\begin{aligned} [a_{ij}, a_{kl}] &= 0 && (1.18a) \\ [b_{ij}, b_{kl}] &= 0 && (1.18b) \\ [a_{ij}, b_{kl}] &= -h\delta_{ik}\delta_{jl} && (1.18c) \end{aligned}$$

where $h \in R$*. Then, for any positive integer* r*, we have*

$$\begin{aligned} \sum_{\substack{I \subseteq [m] \\ |I| = r}} \sum_{\substack{L \subseteq [n] \\ |L| = r}} (\det A_{IL})(\det B_{IL}) &= \sum_{\substack{I \subseteq [n] \\ |I| = r}} \text{col-det}[(A^{\mathrm{T}}B)_{II} + Q_{\mathrm{col}}] && (1.19a) \\ &= \sum_{\substack{I \subseteq [n] \\ |I| = r}} \text{row-det}[(A^{\mathrm{T}}B)_{II} + Q_{\mathrm{row}}] && (1.19b) \\ &= \sum_{\substack{I \subseteq [m] \\ |I| = r}} \text{col-det}[(AB^{\mathrm{T}})_{II} + Q_{\mathrm{col}}] && (1.19c) \\ &= \sum_{\substack{I \subseteq [m] \\ |I| = r}} \text{row-det}[(AB^{\mathrm{T}})_{II} + Q_{\mathrm{row}}] && (1.19d) \end{aligned}$$

where

$$\begin{aligned} Q_{\mathrm{col}} &= h \operatorname{diag}(r-1, r-2, \ldots, 0) && (1.20a) \\ Q_{\mathrm{row}} &= h \operatorname{diag}(0, 1, \ldots, r-1) && (1.20b) \end{aligned}$$

The cognoscenti will of course recognize Corollary 1.3 as (an abstract version of) the Capelli identity [6–8] of classical invariant theory. In Capelli's identity, the ring R is the Weyl algebra $A_{m\times n}(K)$ over some field K of characteristic 0 (e.g. $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$) generated by an $m \times n$ collection $X = (x_{ij})$ of commuting indeterminates ("positions") and the corresponding collection $\partial = (\partial/\partial x_{ij})$ of differential operators (proportional to "momenta"); we then take $A = X$ and $B = \partial$, so that (1.18) holds with $h = 1$.

The Capelli identity has a beautiful interpretation in the theory of group representations [23]: Let $K = \mathbb{R}$ or $\mathbb{C}$, and consider the space $K^{m\times n}$ of $m \times n$ matrices with elements in K, parametrized by coordinates $X = (x_{ij})$. The group $GL(m) \times GL(n)$ acts on $K^{m\times n}$ by

$$(M, N)X \;=\; M^{\mathrm{T}} X N \tag{1.21}$$

where $M \in GL(m)$, $N \in GL(n)$ and $X \in K^{m\times n}$. Then the infinitesimal action associated to (1.21) gives a faithful representation of the Lie algebra $\mathfrak{gl}(m) \oplus \mathfrak{gl}(n)$ by vector fields on $K^{m\times n}$ with linear coefficients:

$$\mathfrak{gl}(m)\colon \quad L_{ij} \;:=\; \sum_{l=1}^{n} x_{il} \frac{\partial}{\partial x_{jl}} \;=\; (X\partial^{\mathrm{T}})_{ij} \qquad \text{for } 1 \leqslant i, j \leqslant m \tag{1.22a}$$

$$\mathfrak{gl}(n)\colon \quad R_{ij} \;:=\; \sum_{l=1}^{m} x_{li} \frac{\partial}{\partial x_{lj}} \;=\; (X^{\mathrm{T}}\partial)_{ij} \qquad \text{for } 1 \leqslant i, j \leqslant n \tag{1.22b}$$

These vector fields have the commutation relations

$$[L_{ij}, L_{kl}] \;=\; \delta_{jk} L_{il} \;-\; \delta_{il} L_{kj} \tag{1.23a}$$

$$[R_{ij}, R_{kl}] \;=\; \delta_{jk} R_{il} \;-\; \delta_{il} R_{kj} \tag{1.23b}$$

$$[L_{ij}, R_{kl}] \;=\; 0 \tag{1.23c}$$

characteristic of $\mathfrak{gl}(m) \oplus \mathfrak{gl}(n)$. Furthermore, the action (L, R) extends uniquely to a homomorphism from the universal enveloping algebra $\mathcal{U}(\mathfrak{gl}(m) \oplus \mathfrak{gl}(n))$ into the Weyl algebra $A_{m\times n}(K)$ [which is isomorphic to the algebra $\mathcal{PD}(K^{m\times n})$ of polynomial-coefficient differential operators on $K^{m\times n}$]. As explained in [23, secs. 1 and 11.1], it can be shown abstractly that any element of the Weyl algebra that commutes with both L and R must be the image via L of some element of the center of $\mathcal{U}(\mathfrak{gl}(m))$, and also the image via R of some element of the center of $\mathcal{U}(\mathfrak{gl}(n))$. The Capelli identity (1.19) with $A = X$ and $B = \partial$ gives an explicit formula for the generators Γ_r $[1 \leqslant r \leqslant \min(m, n)]$ of this subalgebra, from which it is manifest from (1.19a or b) that Γ_r belongs to the image under R of $\mathcal{U}(\mathfrak{gl}(n))$ and commutes with the image under L of $\mathcal{U}(\mathfrak{gl}(m))$, and from (1.19c or d) the reverse fact. See [21–23,25,35,54,55,58] for further discussion of the role of the Capelli identity in classical invariant theory and representation theory, as well as for proofs of the identity.

Let us remark that Proposition 1.2′ also contains Itoh's [25] Capelli-type identity for the generators of the left action of $\mathfrak{o}(m)$ on $m \times n$ matrices (see Example 3.6 below).

Let us also mention one important (and well-known) application of the Capelli identity: namely, it provides a simple proof of the "Cayley" identity[3] for $n \times n$ matrices,

$$\det(\partial)\,(\det X)^s \;=\; s(s+1)\cdots(s+n-1)\,(\det X)^{s-1}\,. \tag{1.24}$$

[3] The identity (1.24) is conventionally attributed to Arthur Cayley (1821–1895); the generalization to arbitrary minors [see (A.17) below] is sometimes attributed to Alfredo Capelli (1855–1910). The trouble is, neither of these formulae occurs anywhere — as far as we can tell — in the *Collected Papers* of Cayley [14]. Nor are we able to find these formulae in any of the relevant works of Capelli [5–9]. The

To derive (1.24), one simply applies both sides of the Capelli identity (1.19) to $(\det X)^s$: the "polarization operators" $L_{ij} = (X\partial^{\mathrm{T}})_{ij}$ and $R_{ij} = (X^{\mathrm{T}}\partial)_{ij}$ act in a very simple way on $\det X$, thereby allowing $\text{col-det}(X\partial^{\mathrm{T}} + Q_{\text{col}})\,(\det X)^s$ and $\text{col-det}(X^{\mathrm{T}}\partial + Q_{\text{col}})\,(\det X)^s$ to be computed easily; they both yield $\det X$ times the right-hand side of (1.24).[4] In fact, by a similar method we can use Proposition 1.2 to prove a generalized "Cayley" identity that lives in the Weyl algebra (rather than just the polynomial algebra) and from which the standard "Cayley" identity can be derived as an immediate corollary: see Proposition A.1 and Corollaries A.3 and A.4 in the Appendix. See also [11] for alternate combinatorial proofs of a variety of Cayley-type identities.

Since the Capelli identity is widely viewed as "mysterious" [2, p. 324] but also as a "powerful formal instrument" [58, p. 39] and a "relatively deep formal result" [52, p. 40], it is of interest to provide simpler proofs. Moreover, since the statement (1.19)/(1.20) of the Capelli identity is purely algebraic/combinatorial, it is of interest to give a purely algebraic/combinatorial proof, independent of the apparatus of representation theory. Such a combinatorial proof was given a decade ago by Foata and Zeilberger [20] for the case $m = n = r$, but their argument was unfortunately somewhat intricate, based on the construction of a sign-reversing involution. The principal goal of the present paper is to provide an extremely short and elementary algebraic proof of Proposition 1.2 and hence of the Capelli identity, based on simple manipulation of commutators. We give this proof in Section 3.

In 1948 Turnbull [53] proved a Capelli-type identity for *symmetric* matrices (see also [57]), and Foata and Zeilberger [20] gave a combinatorial proof of this identity as well. Once again we prove a generalization:

Proposition 1.4 (noncommutative Cauchy–Binet, symmetric version) *Let R be a (not-necessarily-commutative) ring, and let A and B be $n \times n$ matrices with elements in R. Suppose that*

$$[a_{ij}, b_{kl}] \;=\; -h\,(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{1.25}$$

where h is an element of R.

(a) Suppose that A is column-pseudo-commutative and symmetric; and if $n = 2$, suppose further that either

(i) the ring R has the property that $2x = 0$ implies $x = 0$, or

(ii) $[a_{12}, h] = 0$.

---

operator $\Omega = \det(\partial)$ was indeed introduced by Cayley on the second page of his famous 1846 paper on invariants [13]; it became known as Cayley's Ω-process and went on to play an important role in classical invariant theory (see e.g. [18, 21, 35, 47, 51, 58]). But we strongly doubt that Cayley ever knew (1.24). See [1, 11] for further historical discussion.

[4]See e.g. [54, p. 53] or [23, pp. 569–570] for derivations of this type.

Then, for any $I, J \subseteq [n]$ of cardinality $|I| = |J| = r$, we have

$$\begin{aligned}\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\text{col-det}\, A_{LI})(\text{col-det}\, B_{LJ}) &= \text{col-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\text{col}}] &\quad (1.26\text{a}) \\ &= \text{col-det}[(AB)_{IJ} + Q_{\text{col}}] &\quad (1.26\text{b})\end{aligned}$$

where

$$(Q_{\text{col}})_{\alpha\beta} = (r - \beta)\, h\delta_{i_\alpha j_\beta} \tag{1.27}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

(b) Suppose that B is column-pseudo-commutative and symmetric; and if $n = 2$, suppose further that either

(i) the ring R has the property that $2x = 0$ implies $x = 0$, or

(ii) $[b_{12}, h] = 0$.

Then, for any $I, J \subseteq [n]$ of cardinality $|I| = |J| = r$, we have

$$\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\text{row-det}\, A_{LI})(\text{row-det}\, B_{LJ}) = \text{row-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\text{row}}] \tag{1.28}$$

where

$$(Q_{\text{row}})_{\alpha\beta} = (\alpha - 1)\, h\delta_{i_\alpha j_\beta} \tag{1.29}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

Turnbull [53] and Foata–Zeilberger [20] proved their identity for a specific choice of matrices $A = X^{\text{sym}}$ and $B = \partial^{\text{sym}}$ in a Weyl algebra, but it is easy to see that their proof depends only on the commutation properties and symmetry properties of A and B. Proposition 1.4 therefore generalizes their work in three principal ways: they consider only the case $r = n$, while we prove a general identity of Cauchy–Binet type[5]; they assume that *both* A and B are symmetric, while we show that it suffices for *one* of the two to be symmetric; and they assume that *both* $[a_{ij}, a_{kl}] = 0$ and $[b_{ij}, b_{kl}] = 0$, while we show that only *one* of these plays any role and that it moreover can be weakened to column-pseudo-commutativity.[6] We prove Proposition 1.4 in Section 4.[7]

[5]See also Howe and Umeda [23, sec. 11.2] for a formula valid for general r, but involving a *sum* over minors analogous to (1.19).

[6]This last weakening is, however, much less substantial than it might appear at first glance, because a matrix M that is column-pseudo-commutative *and* symmetric necessarily satisfies $2[M_{ij}, M_{kl}] = 0$ for *all* i, j, k, l (see Lemma 2.5 for the easy proof). In particular, in a ring R in which $2x = 0$ implies $x = 0$, column-pseudo-commutativity plus symmetry implies full commutativity.

[7]In the first preprint version of this paper we mistakenly failed to include the extra hypotheses (i) or (ii) in Proposition 1.4 when $n = 2$. For further discussion, see Section 4 and in particular Example 4.2.

Finally, Howe and Umeda [23, eq. (11.3.20)] and Kostant and Sahi [33] independently discovered and proved a Capelli-type identity for *antisymmetric* matrices.[8] Unfortunately, Foata and Zeilberger [20] were unable to find a combinatorial proof of the Howe–Umeda–Kostant–Sahi identity; and we too have been (thus far) unsuccessful. We shall discuss this identity further in Section 5.

Both Turnbull [53] and Foata–Zeilberger [20] also considered a different (and admittedly less interesting) antisymmetric analogue of the Capelli identity, which involves a generalization of the *permanent* of a matrix A,

$$\operatorname{per} A \;:=\; \sum_{\sigma \in \mathcal{S}_n} \prod_{i=1}^{n} a_{i\sigma(i)} \;, \tag{1.30}$$

to matrices with elements in a noncommutative ring R. Since the definition (1.30) is ambiguous without an ordering prescription for the product, we consider the *column-permanent*

$$\text{col-per}\, A \;:=\; \sum_{\sigma \in \mathcal{S}_n} a_{\sigma(1)1}\, a_{\sigma(2)2} \,\cdots\, a_{\sigma(n)n} \tag{1.31}$$

and the *row-permanent*

$$\text{row-per}\, A \;:=\; \sum_{\sigma \in \mathcal{S}_n} a_{1\sigma(1)}\, a_{2\sigma(2)} \,\cdots\, a_{n\sigma(n)} \;. \tag{1.32}$$

(Note that $\text{col-per}\, A = \text{row-per}\, A^{\mathrm{T}}$.) We then prove the following slight generalization of Turnbull's formula:

Proposition 1.5 (Turnbull's antisymmetric analogue) *Let R be a (not-necessarily-commutative) ring, and let A and B be $n \times n$ matrices with elements in R. Suppose that*

$$[a_{ij}, b_{kl}] \;=\; -h\,(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) \tag{1.33}$$

where h is an element of R. Then, for any $I, J \subseteq [n]$ of cardinality $|I| = |J| = r$:

(a) If A is antisymmetric off-diagonal (i.e., $a_{ij} = -a_{ji}$ for $i \neq j$) and $[a_{ij}, h] = 0$ for all i, j, we have

$$\begin{aligned} \sum_{\sigma \in \mathcal{S}_r} \sum_{l_1,\ldots,l_r \in [n]} a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1} \cdots b_{l_r j_r} \;=\; & \\ =\;& \text{col-per}[(A^{\mathrm{T}}B)_{IJ} - Q_{\mathrm{col}}] && (1.34\mathrm{a}) \\ =\;& (-1)^r\; \text{col-per}[(AB)_{IJ} + Q_{\mathrm{col}}] && (1.34\mathrm{b}) \end{aligned}$$

where

$$(Q_{\mathrm{col}})_{\alpha\beta} \;=\; (r - \beta)\, h\delta_{i_\alpha j_\beta} \tag{1.35}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

[8] See also [29] for related work.

(b) If B is antisymmetric off-diagonal (i.e., $b_{ij} = -b_{ji}$ for $i \neq j$) and $[b_{ij}, h] = 0$ for all i, j, we have

$$\sum_{\sigma \in \mathcal{S}_r} \sum_{l_1, \ldots, l_r \in [n]} a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1} \cdots b_{l_r j_r} \;=\; \text{row-per}[(A^{\mathrm{T}} B)_{IJ} - Q_{\text{row}}] \tag{1.36}$$

where

$$(Q_{\text{row}})_{\alpha\beta} \;=\; (\alpha - 1)\, h \delta_{i_\alpha j_\beta} \tag{1.37}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

Note that no requirements are imposed on the $[a, a]$ and $[b, b]$ commutators (but see the Remark at the end of Section 4).

Let us remark that if $[a_{ij}, b_{kl}] = 0$, then the left-hand side of (1.34)/(1.36) is simply

$$\sum_{\sigma \in \mathcal{S}_r} \sum_{l_1, \ldots, l_r \in [n]} a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1} \cdots b_{l_r j_r} \;=\; \text{per}(A^{\mathrm{T}} B)_{IJ}\ , \tag{1.38}$$

so that Proposition 1.5 becomes the *trivial* statement $\text{per}(A^{\mathrm{T}} B)_{IJ} = \text{per}(A^{\mathrm{T}} B)_{IJ}$. So Turnbull's identity does *not* reduce in the commutative case to a formula of Cauchy–Binet type — indeed, no such formula exists for permanents[9] — which is why it is considerably less interesting than the formulae of Cauchy–Binet–Capelli type for determinants.

Turnbull [53] and Foata–Zeilberger [20] proved their identity for a specific choice of matrices $A = X^{\text{antisym}}$ and $B = \partial^{\text{antisym}}$ in a Weyl algebra, but their proof again depends only on the commutation properties and symmetry properties of A and B. Proposition 1.5 therefore generalizes their work in four principal ways: they consider only the case $r = n$, while we prove a general identity for minors; they assume that *both* A and B are antisymmetric, while we show that it suffices for *one* of the two to be antisymmetric *plus an arbitrary diagonal matrix*; and they assume that $[a_{ij}, a_{kl}] = 0$ and $[b_{ij}, b_{kl}] = 0$, while we show that these commutators play no role. We warn the reader that Foata–Zeilberger's [20] statement of this theorem contains a typographical error, inserting a factor $\text{sgn}(\sigma)$ that ought to be absent (and hence inadvertently converting col-per to col-det).[10] We prove Proposition 1.5 in Section 4.[11]

Finally, let us briefly mention some other generalizations of the Capelli identity that have appeared in the literature. One class of generalizations [41, 45, 46, 48] gives formulae for further elements in the (center of the) universal enveloping algebra $\mathcal{U}(\mathfrak{gl}(n))$, such as the so-called quantum immanants. Another class of generalizations extends these

---

[9]But see the Note Added at the end of this introduction.

[10]Also, their verbal description of the other side of the identity — "the matrix product $X^{\mathrm{T}} P$ that appears on the right side of TUR′ is taken with the assumption that the $x_{i,j}$ and $p_{i,j}$ commute" — is ambiguous, but we interpret it as meaning that all the factors $x_{i,j}$ should be moved to the left, as is done on the left-hand side of (1.34)/(1.36).

[11]In the first preprint version of this paper we mistakenly failed to include the hypotheses that $[a_{ij}, h] = 0$ or $[b_{ij}, h] = 0$. See the Remark at the end of Section 4.

formulae to Lie algebras other than $\mathfrak{gl}(n)$ [23–28, 33, 34, 41, 42, 56]. Finally, a third class of generalizations finds analogous formulae in more general structures such as quantum groups [49,50] and Lie superalgebras [44]. Our approach is rather more elementary than all of these works: we ignore the representation-theory context and simply treat the Capelli identity as a noncommutative generalization of the Cauchy–Binet formula. A different generalization along vaguely similar lines can be found in [43].

The plan of this paper is as follows: In Section 2 we make some preliminary comments about the properties of column- and row-determinants. In Section 3 we prove Propositions 1.2 and 1.2′ and Corollary 1.3. We also prove a variant of Proposition 1.2 in which the hypothesis on the commutators $[a_{ij}, a_{kl}]$ is weakened, at the price of a slightly weaker conclusion (see Proposition 3.8). In Section 4 we prove Propositions 1.4 and 1.5. Finally, in Section 5 we discuss whether these results are susceptible of further generalization. In the Appendix we prove a generalization of the "Cayley" identity (1.24).

In a companion paper [10] we shall extend these identities to the (considerably more difficult) case in which $[a_{ij}, b_{kl}] = -g_{ik}h_{jl}$ for general matrices (g_{ik}) and (h_{jl}), whose elements do not necessarily commute.

Note added. Subsequent to the posting of the present paper in preprint form, Chervov, Falqui and Rubtsov [16] posted an extremely interesting survey of the algebraic properties of row-pseudo-commutative matrices (which they call "Manin matrices") when the ring R is an associative algebra over a field of characteristic $\neq 2$. In particular, Section 6 of [16] contains an interesting generalization of the results of the present paper.[12] To state this generalization, note first that the hypotheses of our Proposition 1.2(a) are

(i) A is column-pseudo-commutative, and

(ii) $[a_{ij}, b_{kl}] \;=\; -\delta_{ik}h_{jl}\;.$

Left-multiplying (ii) by a_{km} and summing over k, we obtain

(ii′) $\sum\limits_{k} a_{km}\,[a_{ij}, b_{kl}] \;+\; a_{im}h_{jl} \;=\; 0\;;$

moreover, the converse is true if A is invertible. Furthermore, (i) and (ii) imply

(iii) $[a_{ij}, h_{ls}] \;=\; [a_{il}, h_{js}]$

as shown in Lemma 3.4 below. Then, Chervov *et al.* [16, Theorem 6] observed in essence (translated back to our own language) that our proof of Proposition 1.2(a) used only (i), (ii′) and (iii), and morover that (ii′) can be weakened to[13]

[12]Chervov *et al.* [16] also reformulated the hypotheses and proofs by using Grassmann variables (= exterior algebra) along the lines of [25, 28]. This renders the proofs slightly more compact, and some readers may find that it renders the proofs more transparent as well (this is largely a question of taste). But we do think that the *hypotheses* of the theorems are best stated without reference to Grassmann variables.

[13]Here we have made the translations from their notation to ours ($M \to A^{\mathrm{T}}$, $Y \to B$, $Q \to H$) and written their hypotheses without reference to Grassmann variables. Their Conditions 1 and 2 then correspond to (ii″) and (iii), respectively.

(ii$''$) $\sum_k a_{km}\,[a_{ij}, b_{kl}] \;+\; a_{im}h_{jl} \;=\; [j \leftrightarrow m]$

— that is, we need not demand the vanishing of the left-hand side of (ii$'$), but merely of its antisymmetric part under $j \leftrightarrow m$, *provided that we also assume (iii).* Their Theorem 6 also has the merit of including as a special case not only Proposition 1.2(a) but also Proposition 1.4.

Chervov *et al.* [16, Section 6.5] also provide an interesting rejoinder to our assertion above that no formula of Cauchy–Binet type exists for permanents. They show that if one defines a modified permanent for submatrices involving *possibly repeated indices*, which includes a factor $1/\nu!$ for each index that is repeated ν times, then one obtains a formula of Cauchy–Binet type in which the intermediate sum is over r-tuples of *not necessarily distinct* indices $l_1 \leqslant l_2 \leqslant \ldots \leqslant l_r$. Moreover, this formula of Cauchy–Binet type extends to a Capelli-type formula involving a "quantum correction" [16, Theorems 11–13]. In our opinion this is a very interesting observation, which goes a long way to restore the analogy between determinants and permanents (and which in their formalism reflects the analogy between Grassmann algebra and the algebra of polynomials).

2 Properties of column- and row-determinants

In this section we shall make some preliminary observations about the properties of column- and row-determinants, stressing the following question: Which commutation properties among the elements of the matrix imply which of the standard properties of the determinant? Readers who are impatient to get to the proof of our main results can skim this section lightly. We also call the reader's attention to the historical remarks appended at the end of this section, concerning the commutation hypotheses on matrix elements that have been employed for theorems in noncommutative linear algebra.

Let us begin by recalling two elementary facts that we shall use repeatedly in the proofs throughout this paper:

Lemma 2.1 (Translation Lemma) *Let $\mathcal{A}$ be an abelian group, and let $f\colon \mathcal{S}_n \to \mathcal{A}$. Then, for any $\tau \in \mathcal{S}_n$, we have*

$$\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, f(\sigma) \;=\; \operatorname{sgn}(\tau) \sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma)\, f(\sigma\circ\tau)\;. \tag{2.1}$$

Proof. Just note that both sides equal $\sum_{\sigma\in\mathcal{S}_n} \operatorname{sgn}(\sigma\circ\tau)\, f(\sigma\circ\tau)$. $\square$

Lemma 2.2 (Involution Lemma) *Let $\mathcal{A}$ be an abelian group, and let $f\colon \mathcal{S}_n \to \mathcal{A}$. Suppose that there exists a pair of distinct elements $i, j \in [n]$ such that*

$$f(\sigma) \;=\; f(\sigma\circ(ij)) \tag{2.2}$$

for all $\sigma \in \mathcal{S}_n$ [where (ij) denotes the transposition interchanging i with j]. Then

$$\sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma)\, f(\sigma) \;=\; 0\,. \tag{2.3}$$

Proof. We have

$$\begin{aligned}
\sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma)\, f(\sigma) \;&=\; \sum_{\sigma:\; \sigma(i)<\sigma(j)} \operatorname{sgn}(\sigma)\, f(\sigma) \;+\; \sum_{\sigma:\; \sigma(i)>\sigma(j)} \operatorname{sgn}(\sigma)\, f(\sigma) && \text{(2.4a)} \\
&=\; \sum_{\sigma:\; \sigma(i)<\sigma(j)} \operatorname{sgn}(\sigma)\, f(\sigma) \;-\; \sum_{\sigma':\; \sigma'(i)<\sigma'(j)} \operatorname{sgn}(\sigma')\, f(\sigma' \circ (ij)) && \text{(2.4b)} \\
&=\; 0\,, && \text{(2.4c)}
\end{aligned}$$

where in the second line we made the change of variables $\sigma' = \sigma \circ (ij)$ and used $\operatorname{sgn}(\sigma') = -\operatorname{sgn}(\sigma)$ [or equivalently used the Translation Lemma]. $\square$

With these trivial preliminaries in hand, let us consider noncommutative determinants. Let $M = (M_{ij})$ be a matrix (not necessarily square) with entries in a ring R. Let us call M

- *commutative* if $[M_{ij}, M_{kl}] = 0$ for all i, j, k, l;
- *row-commutative* if $[M_{ij}, M_{kl}] = 0$ whenever $i \neq k$ [i.e., all pairs of elements not in the same row commute];
- *column-commutative* if $[M_{ij}, M_{kl}] = 0$ whenever $j \neq l$ [i.e., all pairs of elements not in the same column commute];
- *weakly commutative* if $[M_{ij}, M_{kl}] = 0$ whenever $i \neq k$ and $j \neq l$ [i.e., all pairs of elements not in the same row or column commute].

Clearly, if M has one of these properties, then so do all its submatrices M_{IJ}. Also, M is commutative if and only if it is both row- and column-commutative.

Weak commutativity is a sufficient condition for the determinant to be defined unambiguously without any ordering prescription, since all the matrix elements in the product (1.1) differ in *both* indices. Furthermore, weak commutativity is sufficient for *single* determinants to have most of their basic properties:

Lemma 2.3 *For weakly commutative square matrices:*

(a) The determinant is antisymmetric under permutation of rows or columns.

(b) The determinant of a matrix with two equal rows or columns is zero.

(c) The determinant of a matrix equals the determinant of its transpose.

The easy proof, which uses the Translation and Involution Lemmas, is left to the reader (it is identical to the usual proof in the commutative case). We simply remark that *if the ring R has the property that $2x = 0$ implies $x = 0$*, then antisymmetry under permutation of rows (or columns) *implies* the vanishing with equal rows (or columns). But if the ring has elements $x \neq 0$ satisfying $2x = 0$ (for instance, if $R = \mathbb{Z}_n$ for n even), then a slightly more careful argument, using the Involution Lemma, is needed to establish the vanishing with equal rows (or columns).

The situation changes, however, when we try to prove a formula for the determinant of a product of *two* matrices, or more generally a formula of Cauchy–Binet type. We are then inevitably led to consider products of matrix elements in which some of the indices may be repeated — but only in *one* of the two positions. It therefore turns out (see Proposition 3.1 below) that we need something like *row*- or *column*-commutativity; indeed, the result can be false without it (see Example 3.2).

Some analogues of Lemma 2.3(a,b) can nevertheless be obtained for the column- and row-determinants under hypotheses *weaker* than weak commutativity. For brevity let us restrict attention to column-determinants; the corresponding results for row-determinants can be obtained by exchanging everywhere "row" with "column".

If $M = (M_{ij})_{i,j=1}^n$ is an $n \times n$ matrix and $\tau \in \mathcal{S}_n$ is a permutation, let us define the matrices obtained from M by permutation of rows or columns:

$$({}^{\tau}M)_{ij} \;:=\; M_{\tau(i)\, j} \tag{2.5a}$$

$$(M^{\tau})_{ij} \;:=\; M_{i\, \tau(j)} \tag{2.5b}$$

We then have the following trivial result:

Lemma 2.4 *For arbitrary square matrices:*

(a) The column-determinant is antisymmetric under permutation of rows*:*

$$\operatorname{col-det} {}^{\tau}M \;=\; \operatorname{sgn}(\tau)\, \operatorname{col-det} M \tag{2.6}$$

for any permutation τ.

(b) The column-determinant of a matrix with two equal rows *is zero.*

Indeed, statements (a) and (b) follow immediately from the Translation Lemma and the Involution Lemma, respectively.

On the other hand, the column-determinant is *not* in general antisymmetric under permutation of *columns*, nor is the column-determinant of a matrix with two equal columns necessarily zero. [For instance, in the Weyl algebra in one variable over a field of characteristic $\neq 2$, we have $\operatorname{col-det} \begin{pmatrix} d & d \\ x & x \end{pmatrix} = dx - xd = 1$, which is neither equal to -1 nor to 0.] It is therefore natural to seek sufficient conditions for these two properties to hold. We now proceed to give a condition, weaker than weak commutativity, that entails the first property and *almost* entails the second property.

Let us begin by observing that $\mu_{ijkl} := [M_{ij}, M_{kl}]$ is manifestly antisymmetric under the simultaneous interchange $i \leftrightarrow k$, $j \leftrightarrow l$. So symmetry under one of these interchanges is equivalent to antisymmetry under the other. Let us therefore say that a matrix M has

- *row-symmetric* (and *column-antisymmetric*) *commutators* if $[M_{ij}, M_{kl}] = [M_{kj}, M_{il}]$ for all i, j, k, l;
- *column-symmetric* (and *row-antisymmetric*) *commutators* if $[M_{ij}, M_{kl}] = [M_{il}, M_{kj}]$ for all i, j, k, l.

Let us further introduce the same types of weakening that we did for commutativity, saying that a matrix M has

- *weakly row-symmetric* (and *column-antisymmetric*) *commutators* if $[M_{ij}, M_{kl}] = [M_{kj}, M_{il}]$ whenever $i \neq k$ and $j \neq l$;
- *weakly column-symmetric* (and *row-antisymmetric*) *commutators* if $[M_{ij}, M_{kl}] = [M_{il}, M_{kj}]$ whenever $i \neq k$ and $j \neq l$.

(Note that row-symmetry is trivial when $i = k$, and column-symmetry is trivial when $j = l$.) Obviously, each of these properties is inherited by all the submatrices M_{IJ} of M. Also, each of these properties is manifestly weaker than the corresponding type of commutativity.

The following fact is sometimes useful:

Lemma 2.5 *Suppose that the square matrix M has either row-symmetric or column-symmetric commutators* and *is either symmetric or antisymmetric. Then $2[M_{ij}, M_{kl}] = 0$ for all i, j, k, l. In particular, if the ring R has the property that $2x = 0$ implies $x = 0$, then M is commutative.*

Proof. Suppose that M has row-symmetric commutators (the column-symmetric case is analogous) and that $M^{\mathrm{T}} = \pm M$. Then $[M_{ij}, M_{kl}] = [M_{kj}, M_{il}] = [M_{jk}, M_{li}] = [M_{lk}, M_{ji}] = [M_{kl}, M_{ij}]$, where the first and third equalities use the row-symmetric commutators, and the second and fourth equalities use symmetry or antisymmetry. □

Returning to the properties of column-determinants, we have:

Lemma 2.6 *If the square matrix M has weakly row-symmetric commutators:*

(a) The column-determinant is antisymmetric under permutation of columns, i.e.

$$\text{col-det}\, M^{\tau} \;=\; \operatorname{sgn}(\tau)\ \text{col-det}\, M \tag{2.7}$$

for any permutation τ.

(b) If M has two equal columns, then $2\,\text{col-det}\,M = 0$. *(In particular, if R is a ring in which $2x = 0$ implies $x = 0$, then* $\text{col-det}\,M = 0$.*)*

(c) If M has two equal columns and the elements in those columns commute among themselves, *then* $\text{col-det}\,M = 0$.

Proof. (a) It suffices to prove the claim when τ is the transposition exchanging i with $i+1$ (for arbitrary i). We have

$$\text{col-det}\,M \;=\; \sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma)\, M_{\sigma(1),1} \,\cdots\, M_{\sigma(i),i}\, M_{\sigma(i+1),i+1} \,\cdots\, M_{\sigma(n),n} \tag{2.8a}$$

$$=\; -\sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma)\, M_{\sigma(1),1} \,\cdots\, M_{\sigma(i+1),i}\, M_{\sigma(i),i+1} \,\cdots\, M_{\sigma(n),n} \tag{2.8b}$$

where the last equality uses the change of variables $\sigma' = \sigma \circ (i, i+1)$ and the fact that $\text{sgn}(\sigma') = -\,\text{sgn}(\sigma)$. Similarly,

$$\text{col-det}\,M^\tau \;=\; \sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma)\, M_{\sigma(1),1} \,\cdots\, M_{\sigma(i),i+1}\, M_{\sigma(i+1),i} \,\cdots\, M_{\sigma(n),n} \tag{2.9a}$$

$$=\; -\sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma)\, M_{\sigma(1),1} \,\cdots\, M_{\sigma(i+1),i+1}\, M_{\sigma(i),i} \,\cdots\, M_{\sigma(n),n}\,. \tag{2.9b}$$

It follows from (2.8a) and (2.9b) that

$$\text{col-det}\,M \;+\; \text{col-det}\,M^\tau \;=\; \sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma)\, M_{\sigma(1),1} \,\cdots\, [M_{\sigma(i),i}, M_{\sigma(i+1),i+1}] \,\cdots\, M_{\sigma(n),n}\,. \tag{2.10}$$

Under the hypothesis that M has weakly row-symmetric commutators [which applies here since $i \neq i+1$ and $\sigma(i) \neq \sigma(i+1)$], the summand [excluding $\text{sgn}(\sigma)$] is invariant under $\sigma \mapsto \sigma \circ (i, i+1)$, so the Involution Lemma implies that the sum is zero.

(b) is an immediate consequence of (a).

(c) Using (a), we may assume without loss of generality that the two equal columns are adjacent (say, in positions 1 and 2). Then, in

$$\text{col-det}\,M \;=\; \sum_{\sigma\in\mathcal{S}_n} \text{sgn}(\sigma)\, M_{\sigma(1)1}\, M_{\sigma(2)2} \,\cdots\, M_{\sigma(n)n}\,, \tag{2.11}$$

we have by hypothesis $M_{i1} = M_{i2}$ and

$$M_{\sigma(1)1}\, M_{\sigma(2)1} \;=\; M_{\sigma(2)1}\, M_{\sigma(1)1}\,, \tag{2.12}$$

so that the summand in (2.11) [excluding $\text{sgn}(\sigma)$] is invariant under $\sigma \mapsto \sigma \circ (12)$; the Involution Lemma then implies that the sum is zero. $\square$

The embarrassing factor of 2 in Lemma 2.6(b) is not simply an artifact of the proof; it is a fact of life when the ring R has elements $x \neq 0$ satisfying $2x = 0$:

Example 2.7 Let R be the ring of 2×2 matrices with elements in the field $GF(2)$, and let α and β be any two noncommuting elements of R [for instance, $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$]. Then the matrix $M = \begin{pmatrix} \alpha & \alpha \\ \beta & \beta \end{pmatrix}$ has both row-symmetric and column-symmetric commutators (and hence also row-*anti*symmetric and column-*anti*symmetric commutators! — note that symmetry is *equivalent* to antisymmetry in a ring of characteristic 2). But $\operatorname{col-det} M = \alpha\beta - \beta\alpha \neq 0$. $\square$

In Proposition 3.8 below, we shall prove a variant of Proposition 1.2 that requires the matrix A^{T} only to have row-symmetric commutators, but at the price of multiplying everything by this embarrassing factor of 2.

If we want to avoid this factor of 2 by invoking Lemma 2.6(c), then (as will be seen in Section 3) we shall need to impose a condition that is intermediate between row-commutativity and row-symmetry: namely, we say (as in Definition 1.1) that M is

- *row-pseudo-commutative* if $[M_{ij}, M_{kl}] = [M_{kj}, M_{il}]$ for all i,j,k,l and $[M_{ij}, M_{kj}] = 0$ for all i,j,k;
- *column-pseudo-commutative* if $[M_{ij}, M_{kl}] = [M_{il}, M_{jk}]$ for all i,j,k,l and $[M_{ij}, M_{il}] = 0$ for all i,j,l.

(Of course, the $[M,M] = [M,M]$ condition need be imposed only when $i \neq k$ and $j \neq l$, since in all other cases it is either trivial or else a consequence of the $[M,M] = 0$ condition.) We thus have M row-commutative $\Longrightarrow$ M row-pseudo-commutative $\Longrightarrow$ M has row-symmetric commutators; furthermore, the converse to the second implication holds whenever R is a ring in which $2x = 0$ implies $x = 0$. Row-pseudo-commutativity thus turns out to be exactly the strengthening of row-symmetry that we need in order to apply Lemma 2.6(c) and thus avoid the factor of 2 in Proposition 3.8, i.e. to prove the full Proposition 1.2.

The following intrinsic characterizations of row-pseudo-commutativity and row-symmetry are perhaps of some interest[14]:

Proposition 2.8 *Let $M = (M_{ij})$ be an $m \times n$ matrix with entries in a (not-necessarily-commutative) ring R.*

(a) Let $x_1, \ldots, x_n$ be commuting indeterminates, and define for $1 \leqslant i \leqslant m$ the elements $\widetilde{x}_i = \sum_{j=1}^{n} M_{ij} x_j$ in the polynomial ring $R[x_1, \ldots, x_n]$. Then the matrix M

[14] Proposition 2.8 is almost identical to a result of Chervov and Falqui [15, Proposition 1], from whom we got the idea; but since they work in an associative algebra over a field of characteristic $\neq 2$, they don't need to distinguish between row-pseudo-commutativity and row-symmetry. They attribute this result to Manin [38, top p. 199] [39, 40], but we are unable to find it there (or perhaps we have simply failed to understand what we have read). However, a result of similar flavor can be found in [38, p. 193, Proposition] [39, pp. 7–8, Theorem 4], and it is probably this to which the authors are referring.

is row-pseudo-commutative if and only if the elements $\widetilde{x}_1, \ldots, \widetilde{x}_m$ commute among themselves.

(b) Let $\eta_1, \ldots, \eta_m$ be Grassmann indeterminates (i.e. $\eta_i^2 = 0$ and $\eta_i \eta_j = -\eta_j \eta_i$), and define for $1 \leqslant j \leqslant n$ the elements $\widetilde{\eta}_j = \sum\limits_{i=1}^{m} \eta_i M_{ij}$ in the Grassmann ring $R[\eta_1, \ldots, \eta_m]_{\mathrm{Gr}}$. Then:

(i) The matrix M has row-symmetric commutators if and only if the elements $\widetilde{\eta}_1, \ldots, \widetilde{\eta}_n$ anticommute among themselves (i.e. $\widetilde{\eta}_i \widetilde{\eta}_j = -\widetilde{\eta}_j \widetilde{\eta}_i$).

(ii) The matrix M is row-pseudo-commutative if and only if the elements $\widetilde{\eta}_1, \ldots, \widetilde{\eta}_n$ satisfy all the Grassmann relations $\widetilde{\eta}_i \widetilde{\eta}_j = -\widetilde{\eta}_j \widetilde{\eta}_i$ and $\widetilde{\eta}_i^2 = 0$.

Proof. (a) We have

$$[\widetilde{x}_i, \widetilde{x}_k] \;=\; \left[\sum_j M_{ij} x_j, \, \sum_l M_{kl} x_l \right] \;=\; \sum_{j,l} [M_{ij}, M_{kl}] \, x_j x_l \, . \tag{2.13}$$

For $j \neq l$, the two terms in $x_j x_l = x_l x_j$ cancel if and only if $[M_{ij}, M_{kl}] = -[M_{il}, M_{kj}]$; and the latter equals $[M_{kj}, M_{il}]$. For $j = l$, there is only one term, and it vanishes if and only if $[M_{ij}, M_{kj}] = 0$.

(b) We have

$$\widetilde{\eta}_j \widetilde{\eta}_l + \widetilde{\eta}_l \widetilde{\eta}_j \;=\; \sum_{i,k} (\eta_i M_{ij} \eta_k M_{kl} + \eta_k M_{kl} \eta_i M_{ij}) \;=\; \sum_{i,k} \eta_i \eta_k \, [M_{ij}, M_{kl}] \tag{2.14}$$

since $\eta_k \eta_i = -\eta_i \eta_k$. For $i \neq k$, the two terms in $\eta_i \eta_k = -\eta_k \eta_i$ cancel if and only if $[M_{ij}, M_{kl}] = [M_{kj}, M_{il}]$. (Note that there is no term with $i = k$, so no further condition is imposed on the commutators $[M, M]$.) On the other hand,

$$\widetilde{\eta}_j^2 \;=\; \sum_{i,k} \eta_i M_{ij} \eta_k M_{kj} \;=\; \sum_{i<k} \eta_i \eta_k \, [M_{ij}, M_{kj}] \, , \tag{2.15}$$

which vanishes if and only if $[M_{ij}, M_{kj}] = 0$ for all i, k. $\square$

Some historical remarks. 1. Row-commutativity has arisen in some previous work on noncommutative linear algebra, beginning with the work of Cartier and Foata on noncommutative extensions of the MacMahon master theorem [12, Théorème 5.1]. For this reason, many authors [15, 30–32] call a row-commutative matrix a *Cartier–Foata matrix*. See e.g. [12, 19, 30, 32, 37] for theorems of noncommutative linear algebra for row-commutative matrices; and see also [32, secs. 5 and 7] for some beautiful q- and $\mathbf{q}$-generalizations.

2. Row-pseudo-commutativity has also arisen previously, beginning (indirectly) with Manin's early work on quantum groups [38–40]. Thus, some authors [15] call a row-pseudo-commutative matrix a *Manin matrix*; others [30–32] call it a *right-quantum matrix*. Results of noncommutative linear algebra for row-pseudo-commutative matrices include Cramer's rule for the inverse matrix [15, 31, 39] and the Jacobi identity for cofactors [31], the formula for the determinant of block matrices [15], Sylvester's determinantal identity [30], the Cauchy–Binet formula (Section 3 below), the Cayley–Hamilton theorem [15], the Newton identities between $\operatorname{tr} M^k$ and coefficients of $\det(tI + M)$ [15], and the MacMahon master theorem [31, 32]; see also [32, secs. 6 and 8] [30, 31] for some beautiful q- and $\mathbf{q}$-generalizations. See in particular [32, Lemma 12.2] for Lemma 2.6 specialized to row-pseudo-commutative matrices.

The aforementioned results suggest that row-pseudo-commutativity is the natural hypothesis for (most? all?) theorems of noncommutative linear algebra involving the column-determinant. Some of these results were derived earlier and/or have simpler proofs under the stronger hypothesis of row-commutativity.

We thank Luigi Cantini for drawing our attention to the paper [15], from which we traced the other works cited here.

3. Subsequent to the posting of the present paper in preprint form, Chervov, Falqui and Rubtsov [16] posted an extremely interesting survey of the algebraic properties of row-pseudo-commutative matrices (which they call "Manin matrices") when the ring R is an associative algebra over a field of characteristic $\neq 2$. This survey discusses the results cited in #2 above, plus many more; in particular, Section 6 of [16] contains an interesting generalization of the results of the present paper on Cauchy–Binet formulae and Capelli-type identities. These authors state explicitly that "the main aim of [their] paper is to argue the following claim: linear algebra statements hold true for Manin matrices in a form identical to the commutative case" [16, first sentence of Section 1.1].

4. The reader may well wonder (as one referee of the present paper did): Since the literature already contains two competing terminologies for the class of matrices in question ("Manin" and "right-quantum"), why muddy the waters by proposing yet another terminology ("row-pseudo-commutative") that is by no means guaranteed to catch on? We would reply by stating our belief that a "good" terminology ought to respect the symmetry $A \mapsto A^{\mathrm{T}}$; or in other words, rows and columns ought to be treated on the same footing, with neither one privileged over the other. (For the same reason, we endeavor to treat the row-determinant and the column-determinant on an equal basis.) We do not claim that our terminology is ideal — perhaps someone will find one that is more concise and easier to remember — but we do think that this symmetry property is important.

3 Proof of the ordinary Capelli-type identities

In this section we shall prove Proposition 1.2; then Proposition 1.2′ and Corollary 1.3 follow immediately. At the end we shall also prove a variant (Proposition 3.8) in which the hypotheses on the commutators are slightly weakened, with a corresponding slight weakening of the conclusion.

It is convenient to begin by reviewing the proof of the classical Cauchy–Binet formula (1.3) where the ring R is commutative. First fix $L = \{l_1, \ldots, l_r\}$ with $l_1 < \ldots < l_r$, and compute

$$\begin{aligned}(\det\, (A^{\mathrm{T}})_{IL})\, (\det B_{LJ}) \;&=\; \sum_{\tau,\pi \in \mathcal{S}_r} \mathrm{sgn}(\tau)\ \mathrm{sgn}(\pi)\, a_{l_1 i_{\tau(1)}} \cdots a_{l_r i_{\tau(r)}} b_{l_{\pi(1)} j_1} \cdots b_{l_{\pi(r)} j_r} &(3.1a)\\ &=\; \sum_{\sigma,\pi \in \mathcal{S}_r} \mathrm{sgn}(\sigma)\, a_{l_{\pi(1)} i_{\sigma(1)}} \cdots a_{l_{\pi(r)} i_{\sigma(r)}} b_{l_{\pi(1)} j_1} \cdots b_{l_{\pi(r)} j_r} \;, &(3.1b)\end{aligned}$$

where we have written $\sigma = \tau \circ \pi$ and exploited the commutativity of the elements of A (but not of B). Now the sum over L and π is equivalent to summing over all r-tuples of *distinct* elements $l_1, \ldots, l_r \in [m]$:

$$\sum_L (\det\, (A^{\mathrm{T}})_{IL})\, (\det B_{LJ}) \;=\; \sum_{l_1,\ldots,l_r \in [m] \,\mathrm{distinct}} f(l_1, \ldots, l_r)\, b_{l_1 j_1} \cdots b_{l_r j_r} \;, \qquad (3.2)$$

where we have defined

$$f(l_1, \ldots, l_r) \;:=\; \sum_{\sigma \in \mathcal{S}_r} \mathrm{sgn}(\sigma)\, a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} \qquad (3.3)$$

for arbitrary $l_1, \ldots, l_r \in [m]$. Note now that $f(l_1, \ldots, l_r) = 0$ whenever two or more arguments take the same value, because (3.3) is then the determinant of a matrix with two (or more) equal rows. We can therefore add such terms to the sum (3.2), yielding

$$\begin{aligned}\sum_L (\det\, (A^{\mathrm{T}})_{IL})\, (\det B_{LJ}) \;&=\; \sum_{l_1,\ldots,l_r \in [m]} \sum_{\sigma \in \mathcal{S}_r} \mathrm{sgn}(\sigma)\, a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1} \cdots b_{l_r j_r} &(3.4a)\\ &=\; \sum_{\sigma \in \mathcal{S}_r} \mathrm{sgn}(\sigma)\, (A^{\mathrm{T}} B)_{i_{\sigma(1)} j_1} \cdots (A^{\mathrm{T}} B)_{i_{\sigma(r)} j_r} &(3.4b)\\ &=\; \det\, (A^{\mathrm{T}} B)_{IJ} \;, &(3.4c)\end{aligned}$$

where we have here commuted the b's through the a's. Note that the order of the elements of B remains unchanged throughout these manipulations.

Let us also remark that this proof is valid even if $r > m$: the starting sum (3.2) is then empty, since there do not exist *distinct* elements $l_1, \ldots, l_r \in [m]$; but the sum (3.4a) is nonempty, since repetitions among the $l_1, \ldots, l_r$ are now allowed, and we prove the non-trivial result that $\det(A^{\mathrm{T}} B)_{IJ} = 0$. (Of course, in the commutative case this is no surprise, since the matrix $A^{\mathrm{T}} B$ has rank at most m; but the corresponding noncommutative result will be less trivial.)

Now let us examine this proof more closely, in order to see what commutation properties of the matrix elements were really needed to make it work. In the passage from

(3.1a) to (3.1b), the essence of the argument was that

$$(\text{col-det}\,(A^{\mathrm{T}})_{IL})\,(\text{col-det}\,B_{LJ}) \;=\; \sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)\,(\text{col-det}\,(A^{\mathrm{T}})_{IL})\, b_{l_{\pi(1)}j_1}\cdots b_{l_{\pi(r)}j_r} \tag{3.5a}$$

$$= \sum_{\pi\in\mathcal{S}_r} \operatorname{sgn}(\pi)^2\,[\text{col-det}\,((A^{\mathrm{T}})_{IL})^{\pi}]\, b_{l_{\pi(1)}j_1}\cdots b_{l_{\pi(r)}j_r}\;, \tag{3.5b}$$

where Lemma 2.6(a) justifies the passage from the first line to the second; so it suffices for A^{T} to have weakly row-symmetric commutators. In the argument that $f(l_1,\ldots,l_r)=0$ whenever two or more arguments take the same value, we need to apply Lemma 2.6(c) to a matrix that is a submatrix of A^{T} *with possibly repeated columns*; therefore we need, in addition to weak row-symmetry, the additional hypothesis that the matrix elements of A^{T} within each column commute among themselves — or in other words, we need A^{T} to be row-pseudo-commutative (Definition 1.1). Finally, in the step from (3.4a) to (3.4b), we commuted the b's through the a's. We have therefore proven:

Proposition 3.1 (easy noncommutative Cauchy–Binet) *Let R be a (not-necessarily-commutative) ring, and let A and B be $m\times n$ matrices with elements in R. Suppose that*

(a) A^{T} is row-pseudo-commutative, i.e. A is column-pseudo-commutative, i.e. $[a_{ij},a_{kl}] = [a_{il},a_{kj}]$ whenever $i\neq k$ and $j\neq l$ and $[a_{ij},a_{il}]=0$ whenever $j\neq l$;

(b) the matrix elements of A commute with those of B, i.e. $[a_{ij},b_{kl}]=0$ for all i,j,k,l.

Then, for any $I,J\subseteq[n]$ of cardinality $|I|=|J|=r$, we have

$$\sum_{\substack{L\subseteq[m]\\|L|=r}} (\text{col-det}\,(A^{\mathrm{T}})_{IL})(\text{col-det}\,B_{LJ}) \;=\; \text{col-det}\,(A^{\mathrm{T}}B)_{IJ}\;. \tag{3.6}$$

Note that no hypothesis whatsoever is needed concerning the commutators $[b_{ij},b_{kl}]$.

There is also a dual result using row-det, in which B is required to be column-pseudo-commutative and no hypothesis is needed on the $[a,a]$ commutators.

The hypothesis in Proposition 3.1 that A be column-pseudo-commutative really is necessary:

Example 3.2 Let α and β be any noncommuting elements of the ring R, and let $A=\begin{pmatrix}\alpha & \beta\\ 0 & 0\end{pmatrix}$ and $B=\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}$ [let us assume for simplicity that the ring R has an identity element], so that $A^{\mathrm{T}}B=\begin{pmatrix}\alpha & \alpha\\ \beta & \beta\end{pmatrix}$. Then A is row-commutative but not column-pseudo-commutative, while the elements of B commute with everything. We have $\det A^{\mathrm{T}} = \det B = 0$ but $\text{col-det}(A^{\mathrm{T}}B)=\alpha\beta-\beta\alpha\neq 0$.

This example can be streamlined by dropping the second row of the matrices A and B, i.e. considering it as an example with $m = 1$, $n = 2$ and $r = 2$. Then the left-hand side of (3.6) is an empty sum (since $r > m$), but the right-hand side does not vanish. $\square$

Example 3.3 It is instructive to consider the general case of 2×2 matrices (i.e. $m = n = 2$) under the sole hypothesis that $[a_{ij}, b_{kl}] = 0$ for all i, j, k, l. We have

$$\begin{aligned} \text{col-det}(A^{\mathrm{T}}B) \;-\; (\text{col-det}\, A^{\mathrm{T}})(\text{col-det}\, B) \;&=\; \big([a_{21}, a_{12}] \,+\, [a_{11}, a_{22}]\big)\, b_{21}b_{12} \\ &\qquad +\, [a_{11}, a_{12}]\, b_{11}b_{12} \;+\; [a_{21}, a_{22}]\, b_{21}b_{22} \,, \end{aligned} \tag{3.7}$$

where the terms on the first (resp. second) line of the right-hand side come from the first (resp. second) step of the proof. We see that column-pseudo-commutativity of A is *precisely* what we need in order to guarantee that (3.7) vanishes for arbitrary matrices B. $\square$

We are now ready to consider Proposition 1.2, which generalizes Proposition 3.1 by allowing nonzero commutators $[a_{ij}, b_{kl}] = -\delta_{ik} h_{jl}$, thereby producing a "quantum correction" on the right-hand side of the identity. In the proof of Proposition 1.2(a) it will be necessary (as we shall see) to commute the h's through the a's. We therefore begin with a lemma giving an important property of such commutators:

Lemma 3.4 *Let R be a (not-necessarily-commutative) ring, and let A and B be $m \times n$ matrices with elements in R. Suppose that for all $i, k \in [m]$ and $j, l \in [n]$ we have*

$$[a_{ij}, a_{il}] \;=\; 0 \tag{3.8a}$$

$$[a_{ij}, b_{kl}] \;=\; -\delta_{ik} h_{jl} \tag{3.8b}$$

where $(h_{jl})_{j,l=1}^{n}$ are elements of R. Then, for all $i \in [m]$ and $j, l, s \in [n]$ we have

$$[a_{ij}, h_{ls}] \;=\; [a_{il}, h_{js}] \,. \tag{3.9}$$

Note the very weak hypothesis here on the $[a, a]$ commutators: we require $[a_{ij}, a_{kl}] = 0$ only when $i = k$, i.e. between different columns *within the same row*. This is much weaker than the column-pseudo-commutativity assumed in Proposition 1.2(a), as it imposes (1.7) but omits (1.6).

Proof. For any indices $i, k, r \in [m]$ and $j, l, s \in [n]$, we have the Jacobi identity

$$[a_{ij}, [a_{kl}, b_{rs}]] \;+\; [a_{kl}, [b_{rs}, a_{ij}]] \;+\; [b_{rs}, [a_{ij}, a_{kl}]] \;=\; 0 \,. \tag{3.10}$$

By taking $k = r = i$ and using the hypotheses (3.8), we obtain the conclusion (3.9). $\square$

Remark. Since A and B play symmetrical roles in this problem (modulo the substitution $h \mapsto -h^{\mathrm{T}}$), a similar argument shows that if $[b_{ij}, b_{il}] = 0$, then $[b_{ij}, h_{ls}] = [b_{is}, h_{lj}]$. This will be relevant for Proposition 1.2(b). $\square$

One consequence of Lemma 3.4 is that h can be commuted through a when it arises inside a sum over permutations with the factor $\operatorname{sgn}(\sigma)$:

Corollary 3.5 *Fix distinct elements $\alpha, \beta \in [r]$ and fix a set $I \subseteq [n]$ of cardinality $|I| = r$. Then, under the hypotheses of Lemma 3.4, we have*

$$\sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma)\, F\big(\{\sigma(j)\}_{j \neq \alpha,\beta}\big)\, [a_{l i_{\sigma(\alpha)}}, h_{i_{\sigma(\beta)} k}]\, G\big(\{\sigma(j)\}_{j \neq \alpha,\beta}\big) \;=\; 0 \tag{3.11}$$

for arbitrary functions $F, G\colon\, [r]^{r-2} \to R$ and arbitrary indices $l \in [m]$ and $k \in [n]$.

Proof. By Lemma 3.4 we have

$$[a_{l i_{\sigma(\alpha)}}, h_{i_{\sigma(\beta)} j}] \;=\; [a_{l i_{\sigma(\beta)}}, h_{i_{\sigma(\alpha)} j}] \,. \tag{3.12}$$

This means that the summand in (3.11) [excluding the factor $\operatorname{sgn}(\sigma)$] is invariant under $\sigma \mapsto \sigma \circ (\alpha\beta)$. The claim then follows immediately from the Involution Lemma. □

We also have a dual version of Corollary 3.5, along the lines of the Remark above, stating that if $[b_{ij}, b_{il}] = 0$, then sums involving $[b_{l j_{\sigma(\alpha)}}, h_{k j_{\sigma(\beta)}}]$ vanish. Let us call this Corollary 3.5′.

We are now ready to prove Proposition 1.2:

Proof of Proposition 1.2. We begin with part (a). The first two steps in the proof are identical to those in Proposition 3.1: we therefore have

$$\sum_L (\text{col-det}\,(A^{\mathrm{T}})_{IL})\,(\text{col-det}\,B_{LJ}) \;=\; \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1, \cdots, l_r \in [m]} a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1} \cdots b_{l_r j_r} \,. \tag{3.13}$$

It is only now that we have to work harder, because of the noncommutativity of the b's with the a's. Let us begin by moving the factor $b_{l_1 j_1}$ to the left until it lies just to the right of $a_{l_1 i_{\sigma(1)}}$, using the general formula

$$x_1\, [x_2 \cdots x_r, y] \;=\; x_1 \sum_{s=2}^{r} x_2 \cdots x_{s-1}\, [x_s, y]\, x_{s+1} \cdots x_r \tag{3.14}$$

with $x_\alpha = a_{l_\alpha i_{\sigma(\alpha)}}$ and $y = b_{l_1 j_1}$. This gives

$$\begin{aligned} &\sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1, \ldots, l_r \in [m]} a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1} \cdots b_{l_r j_r} \\ &\quad = \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1, \ldots, l_r \in [m]} a_{l_1 i_{\sigma(1)}} \Bigg[b_{l_1 j_1} a_{l_2 i_{\sigma(2)}} \cdots a_{l_r i_{\sigma(r)}} \\ &\qquad - \sum_{s=2}^{r} \delta_{l_1 l_s}\, a_{l_2 i_{\sigma(2)}} \cdots a_{l_{s-1} i_{\sigma(s-1)}} h_{i_{\sigma(s)} j_1} a_{l_{s+1} i_{\sigma(s+1)}} \cdots a_{l_r i_{\sigma(r)}} \Bigg]\, b_{l_2 j_2} \cdots b_{l_r j_r} \,. \end{aligned} \tag{3.15}$$

Now we repeatedly use Corollary 3.5 to push the factor $h_{i_{\sigma(s)}j_1}$ to the left: we obtain

$$\sum_{\sigma\in\mathcal{S}_r}\operatorname{sgn}(\sigma)\sum_{l_1,\ldots,l_r\in[m]} a_{l_1 i_{\sigma(1)}}\Bigg[b_{l_1j_1}a_{l_2i_{\sigma(2)}}\cdots a_{l_ri_{\sigma(r)}} - \sum_{s=2}^{r} h_{i_{\sigma(s)}j_1}\delta_{l_1l_s}a_{l_2i_{\sigma(2)}}\cdots a_{l_{s-1}i_{\sigma(s-1)}}a_{l_{s+1}i_{\sigma(s+1)}}\cdots a_{l_ri_{\sigma(r)}}\Bigg] b_{l_2j_2}\cdots b_{l_rj_r} \tag{3.16a}$$

$$= \sum_{\sigma\in\mathcal{S}_r}\operatorname{sgn}(\sigma)\sum_{l_2,\ldots,l_r\in[m]}\Bigg[(A^{\mathrm{T}}B)_{i_{\sigma(1)}j_1}a_{l_2i_{\sigma(2)}}\cdots a_{l_ri_{\sigma(r)}} - \sum_{s=2}^{r} h_{i_{\sigma(s)}j_1}a_{l_2i_{\sigma(2)}}\cdots a_{l_{s-1}i_{\sigma(s-1)}}a_{l_si_{\sigma(1)}}a_{l_{s+1}i_{\sigma(s+1)}}\cdots a_{l_ri_{\sigma(r)}}\Bigg] b_{l_2j_2}\cdots b_{l_rj_r} \tag{3.16b}$$

$$= \sum_{\sigma\in\mathcal{S}_r}\operatorname{sgn}(\sigma)\Bigg[(A^{\mathrm{T}}B)_{i_{\sigma(1)}j_1} + \sum_{s=2}^{r} h_{i_{\sigma(1)}j_1}\Bigg]\sum_{l_2,\ldots,l_r\in[m]} a_{l_2i_{\sigma(2)}}\cdots a_{l_ri_{\sigma(r)}}b_{l_2j_2}\cdots b_{l_rj_r} \tag{3.16c}$$

$$= \sum_{\sigma\in\mathcal{S}_r}\operatorname{sgn}(\sigma)\left[A^{\mathrm{T}}B + (r-1)\,h\right]_{i_{\sigma(1)}j_1}\sum_{l_2,\ldots,l_r\in[m]} a_{l_2i_{\sigma(2)}}\cdots a_{l_ri_{\sigma(r)}}b_{l_2j_2}\cdots b_{l_rj_r}\,, \tag{3.16d}$$

where we have simply executed the sum over l_1 and, in the second summand, interchanged $\sigma(1)$ with $\sigma(s)$ [which multiplies $\operatorname{sgn}(\sigma)$ by -1]. This procedure can be now iterated to obtain

$$\sum_L (\det\,(A^{\mathrm{T}})_{IL})(\text{col-det}\,B_{LJ}) \tag{3.17a}$$

$$= \sum_{\sigma\in\mathcal{S}_r}\operatorname{sgn}(\sigma)\left[A^{\mathrm{T}}B + (r-1)\,h\right]_{i_{\sigma(1)}j_1}\left[A^{\mathrm{T}}B + (r-2)\,h\right]_{i_{\sigma(2)}j_2}\cdots\left[A^{\mathrm{T}}B\right]_{i_{\sigma(r)}j_r} \tag{3.17b}$$

$$= \text{col-det}\left[(A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}}\right]\,, \tag{3.17c}$$

which is the desired result of part (a).

For part (b), let us start as before:

$$\sum_L (\text{row-det}(A^{\mathrm{T}})_{IL})\;(\text{row-det}\,B_{LJ}) \tag{3.18a}$$

$$= \sum_L\sum_{\tau,\pi\in\mathcal{S}_r}\operatorname{sgn}(\tau)\operatorname{sgn}(\pi)\,a_{l_{\tau(1)}i_1}\cdots a_{l_{\tau(r)}i_r}b_{l_1j_{\pi(1)}}\cdots b_{l_rj_{\pi(r)}} \tag{3.18b}$$

$$= \sum_L\sum_{\tau,\sigma\in\mathcal{S}_r}\operatorname{sgn}(\sigma)\,a_{l_{\tau(1)}i_1}\cdots a_{l_{\tau(r)}i_r}b_{l_{\tau(1)}j_{\sigma(1)}}\cdots b_{l_{\tau(r)}j_{\sigma(r)}} \tag{3.18c}$$

where we have written $\sigma = \pi\circ\tau$ and exploited the commutativity of the elements of B (but not of A). An argument as in Proposition 3.1 allows us to rewrite this as

$$\sum_{\sigma\in\mathcal{S}_r}\operatorname{sgn}(\sigma)\sum_{l_1,\cdots,l_r\in[m]} a_{l_1i_1}\cdots a_{l_ri_r}b_{l_1j_{\sigma(1)}}\cdots b_{l_rj_{\sigma(r)}}\,. \tag{3.19}$$

We first move the factor $a_{l_r i_r}$ to the right, giving

$$\sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1,\ldots,l_r \in [m]} a_{l_1 i_1} \cdots a_{l_{r-1} i_{r-1}} \Bigg[b_{l_1 j_{\sigma(1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} a_{l_r i_r} - \sum_{s=1}^{r-1} \delta_{l_r l_s} b_{l_1 j_{\sigma(1)}} \cdots b_{l_{s-1} j_{\sigma(s-1)}} h_{i_r j_{\sigma(s)}} b_{l_{s+1} j_{\sigma(s+1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} \Bigg] b_{l_r j_{\sigma(r)}} . \tag{3.20}$$

Now we repeatedly use Corollary 3.5′ to push the factor $h_{i_r j_{\sigma(s)}}$ to the right: we obtain

$$\sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1,\ldots,l_r \in [m]} a_{l_1 i_1} \cdots a_{l_{r-1} i_{r-1}} \Bigg[b_{l_1 j_{\sigma(1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} a_{l_r i_r} - \sum_{s=1}^{r-1} \delta_{l_r l_s} b_{l_1 j_{\sigma(1)}} \cdots b_{l_{s-1} j_{\sigma(s-1)}} b_{l_{s+1} j_{\sigma(s+1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} h_{i_r j_{\sigma(s)}} \Bigg] b_{l_r j_{\sigma(r)}} \tag{3.21a}$$

$$= \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1,\ldots,l_{r-1} \in [m]} a_{l_1 i_1} \cdots a_{l_{r-1} i_{r-1}} \Bigg[b_{l_1 j_{\sigma(1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} (A^{\mathrm{T}} B)_{i_r j_{\sigma(r)}} - \sum_{s=1}^{r-1} b_{l_1 j_{\sigma(1)}} \cdots b_{l_{s-1} j_{\sigma(s-1)}} b_{l_s j_{\sigma(r)}} b_{l_{s+1} j_{\sigma(s+1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} h_{i_r j_{\sigma(s)}} \Bigg] \tag{3.21b}$$

$$= \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1,\ldots,l_{r-1} \in [m]} a_{l_1 i_1} \cdots a_{l_{r-1} i_{r-1}} b_{l_1 j_{\sigma(1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} \Bigg[(A^{\mathrm{T}} B)_{i_r j_{\sigma(r)}} + \sum_{s=1}^{r-1} h_{i_r j_{\sigma(r)}} \Bigg] \tag{3.21c}$$

$$= \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1,\ldots,l_{r-1} \in [m]} a_{l_1 i_1} \cdots a_{l_{r-1} i_{r-1}} b_{l_1 j_{\sigma(1)}} \cdots b_{l_{r-1} j_{\sigma(r-1)}} [A^{\mathrm{T}} B + (r-1)h]_{i_r j_{\sigma(r)}} , \tag{3.21d}$$

where we exchanged $\sigma(s)$ with $\sigma(r)$. This procedure can be iterated as before to obtain

$$\sum_L (\text{row-det}\, (A^{\mathrm{T}})_{IL})(\det B_{LJ}) \tag{3.22a}$$

$$= \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \left[A^{\mathrm{T}} B\right]_{i_1 j_{\sigma(1)}} \cdots \left[A^{\mathrm{T}} B + (r-2)\, h\right]_{i_{r-1} j_{\sigma(r-1)}} \left[A^{\mathrm{T}} B + (r-1)\, h\right]_{i_r j_{\sigma(r)}} \tag{3.22b}$$

$$= \text{row-det} \left[(A^{\mathrm{T}} B)_{IJ} + Q_{\mathrm{row}}\right] , \tag{3.22c}$$

which is the desired result of part (b). □

Let us remark that if we care only about the Capelli identity (i.e., Corollary 1.3 with $h = 1$), then the proof becomes even simpler: all the discussion about column-pseudo-commutativity is unnecessary because we have the stronger hypothesis $[a_{ij}, a_{kl}] = 0$, so the first steps in the proof proceed exactly as in the commutative case; and Lemma 3.4

is unnecessary because the quantities h_{jl} are by hypothesis central elements of the ring. The key nontrivial steps in the proof of the Capelli identity are thus the manipulations of commutators leading from (3.13) to (3.17).

Example 3.6 Itoh [25, Theorem A] proves a Capelli-type identity for some matrices arising from the action of $O(m) \times Sp(n)$ on $m \times n$ matrices (n even). His result can be obtained as a special case of Proposition 1.2′, as follows: We write $\nu = n/2$ and work in the Weyl algebra $A_{m\times\nu}(K)$ generated by an $m\times\nu$ matrix of indeterminates $X = (x_{ij})$ and the corresponding matrix $\partial = (\partial/\partial x_{ij})$ of differential operators. Now form the $m\times n$ matrices $A = (X, \partial)$ and $B = (-\partial, X)$, which satisfy the commutation relations $[a_{ij}, b_{kl}] = \delta_{ik}\delta_{jl}$. Note that A and B are row-commutative but *not* column-pseudo-commutative. We can therefore apply Proposition 1.2′(c) to obtain the identity

$$\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\det A_{IL})(\det\, (B^{\mathrm{T}})_{LJ}) \;=\; \text{col-det}[(AB^{\mathrm{T}})_{IJ} + Q_{\mathrm{col}}] \tag{3.23a}$$

$$\;=\; \text{row-det}[(AB^{\mathrm{T}})_{IJ} + Q_{\mathrm{row}}]\;, \tag{3.23b}$$

in which the matrix AB^{T} describes the left action of $\mathfrak{o}(m)$. But we *cannot* obtain a corresponding identity for the matrix $A^{\mathrm{T}}B$, which describes the right action of $\mathfrak{sp}(n)$, because A and B fail to be column-pseudo-commutative. Itoh [25] also notes this failure and interprets it from the point of view of his own proof (which uses exterior algebra); he goes on to provide a weaker alternative formula. $\square$

Example 3.7 (generalizing Example 3.2) It is instructive to consider the most general case with $m = 1$, $n = 2$ and $r = 2$: here the left-hand side of the identity is automatically zero (since $r > m$), but the right-hand side need not vanish unless we make suitable hypotheses on the matrices $A = (\alpha, \beta)$ and $B = (\gamma, \delta)$. We have $A^{\mathrm{T}}B = \begin{pmatrix} \alpha\gamma & \alpha\delta \\ \beta\gamma & \beta\delta \end{pmatrix}$, $H = \begin{pmatrix} \gamma\alpha - \alpha\gamma & \delta\alpha - \alpha\delta \\ \gamma\beta - \beta\gamma & \delta\beta - \beta\delta \end{pmatrix}$ and $Q_{\mathrm{col}} = \begin{pmatrix} \gamma\alpha - \alpha\gamma & 0 \\ \gamma\beta - \beta\gamma & 0 \end{pmatrix}$, so that $A^{\mathrm{T}}B + Q_{\mathrm{col}} = \begin{pmatrix} \gamma\alpha & \alpha\delta \\ \gamma\beta & \beta\delta \end{pmatrix}$ (a beautiful cancellation!) and hence

$$\text{col-det}(A^{\mathrm{T}}B + Q_{\mathrm{col}}) \;=\; \gamma\alpha\beta\delta - \gamma\beta\alpha\delta \;=\; \gamma\,[\alpha, \beta]\,\delta\;. \tag{3.24}$$

In Example 3.2 we took $\gamma = \delta = 1$ and found that the identity fails unless $[\alpha, \beta] = 0$, i.e. A is column-(pseudo-)commutative. In Itoh's [25] $\mathfrak{sp}(n)$ action (specialized to $m = 1$ and $n = 2$) we work in the Weyl algebra $A_1(K)$ and have $\alpha = x$, $\beta = d$, $\gamma = -d$ and $\delta = x$, so that $[\alpha, \beta] = -1 \neq 0$ and the identity again fails. $\square$

Remarks. 1. Part (b) of Proposition 1.2 is essentially equivalent to part (a) under a duality that reverses the order of products inside each monomial, when R is the algebra of noncommutative polynomials (over $\mathbb{Z}$) in indeterminates (a_{ij}) and (b_{ij}) with the appropriate relations. For brevity we omit the details.

2. Suppose that $[a_{ij}, b_{kl}] = -g_{ik} h_{jl}$ where $G = (g_{ik})$ and $H = (h_{jl})$ are two matrices. Now make the replacements $A \to PAQ$ and $B \to RAS$, where P, Q, R, S are matrices whose elements commute with each other and with those of A, B, G, H; then $G \to PGR^{\mathrm{T}}$ and $H \to Q^{\mathrm{T}} HS$. It follows that Proposition 1.2 with general h — or even the extension to general g and h, *provided that* $[g_{ik}, h_{jl}] = 0$ — is not terribly much more general than the traditional case $g = h = \delta$. However, the case in which $[g_{ik}, h_{jl}] \neq 0$ is much more difficult [10]. □

Let us now state and prove a variant of Proposition 1.2 in which the hypotheses on the $[a, a]$ or $[b, b]$ commutators are weakened: instead of assuming that A (or B) is column-pseudo-commutative, we shall assume only that it has column-symmetric commutators. The price that we have to pay (as in Lemma 2.6(b)) is that the final identity must be multiplied by a factor of 2 — or in other words, the identity holds only up to addition of an unknown element $x \in R$ satisfying $2x = 0$. Of course, if the ring R is such that $2x = 0$ implies $x = 0$ (as it is in most applications), then both the hypotheses and the conclusion are equivalent to those of Proposition 1.2; but in general rings, the hypotheses are slightly weaker, as is the conclusion.

Proposition 3.8 (noncommutative Cauchy–Binet, variant) *Let R be a (not-necessarily-commutative) ring, and let A and B be $m \times n$ matrices with elements in R. Suppose that*

$$[a_{ij}, b_{kl}] \;=\; -\delta_{ik} h_{jl} \tag{3.25}$$

where $(h_{jl})_{j,l=1}^{n}$ are elements of R. Then, for any $I, J \subseteq [n]$ of cardinality $|I| = |J| = r$:

(a) If A has column-symmetric commutators, then

$$2 \sum_{\substack{L \subseteq [m] \\ |L| = r}} (\text{col-det}\,(A^{\mathrm{T}})_{IL})(\text{col-det}\,B_{LJ}) \;=\; 2\,\text{col-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}}] \tag{3.26}$$

where

$$(Q_{\mathrm{col}})_{\alpha\beta} \;=\; (r - \beta)\, h_{i_\alpha j_\beta} \tag{3.27}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

(b) If B has column-symmetric commutators, then

$$2 \sum_{\substack{L \subseteq [m] \\ |L| = r}} (\text{row-det}\,(A^{\mathrm{T}})_{IL})(\text{row-det}\,B_{LJ}) \;=\; 2\,\text{row-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{row}}] \tag{3.28}$$

where

$$(Q_{\mathrm{row}})_{\alpha\beta} \;=\; (\alpha - 1)\, h_{i_\alpha j_\beta} \tag{3.29}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

In proving Proposition 3.8, we shall need a variant of Lemma 3.4:

Lemma 3.9 *Let R be a (not-necessarily-commutative) ring, and let A and B be $m \times n$ matrices with elements in R. Suppose that for all $i, k \in [m]$ and $j, l \in [n]$ we have*

$$[a_{ij}, a_{kl}] \;=\; [a_{il}, a_{kj}] \tag{3.30a}$$
$$[a_{ij}, b_{kl}] \;=\; -\delta_{ik} h_{jl} \tag{3.30b}$$

where $(h_{jl})_{j,l=1}^{n}$ are elements of R. Then, for all $i \in [m]$ and $j, l, s \in [n]$ we have

$$2\,[a_{ij}, h_{ls}] \;=\; 2\,[a_{il}, h_{js}]\;; \tag{3.31}$$

and if $m \geqslant 2$ we have

$$[a_{ij}, h_{ls}] \;=\; [a_{il}, h_{js}]\;. \tag{3.32}$$

Proof. We consider the Jacobi identity (3.10) and the corresponding identity with j and l interchanged:

$$[a_{ij}, [a_{kl}, b_{rs}]] \,+\, [a_{kl}, [b_{rs}, a_{ij}]] \,+\, [b_{rs}, [a_{ij}, a_{kl}]] \;=\; 0 \tag{3.33a}$$
$$[a_{il}, [a_{kj}, b_{rs}]] \,+\, [a_{kj}, [b_{rs}, a_{il}]] \,+\, [b_{rs}, [a_{il}, a_{kj}]] \;=\; 0 \tag{3.33b}$$

Now subtract the two equations and use the hypotheses (3.30): we obtain

$$\delta_{kr}\left([a_{ij}, h_{ls}] - [a_{il}, h_{js}]\right) \;=\; \delta_{ir}\left([a_{kl}, h_{js}] - [a_{kj}, h_{ls}]\right)\;. \tag{3.34}$$

Now fix arbitrarily the indices $i \in [m]$ and $j, l, s \in [n]$. Choosing $k = r = i$, we get

$$[a_{ij}, h_{ls}] - [a_{il}, h_{js}] \;=\; [a_{il}, h_{js}] - [a_{ij}, h_{ls}]\;, \tag{3.35}$$

which is (3.31). For $m \geqslant 2$ we can choose $k = r \in [m]$ different from i, which yields (3.32). $\square$

Proof of Proposition 3.8. The proof is almost identical to that of Proposition 1.2, with the following changes: in the argument that $f(l_1, \ldots, l_r) = 0$ whenever two or more arguments take the same value, we use Lemma 2.6(b) instead of Lemma 2.6(c); and in the commutation argument involving h, we use Lemma 3.9 [eq. (3.31)] instead of Lemma 3.4. These two weakenings are permissible since we are multiplying the identity by 2. $\square$

4 Proof of the Turnbull-type identities

In this section we shall prove Propositions 1.4 and 1.5, following the same strategy as in the previous proofs.

We begin with a variant of Lemma 3.4 that will be useful in both the symmetric and antisymmetric cases.

Lemma 4.1 *Let $n \geqslant 2$, let R be a (not-necessarily-commutative) ring, and let A and B be $n \times n$ matrices with elements in R. Suppose that for all $i, j, k, l \in [n]$ we have*

$$[a_{ij}, a_{il}] \;=\; 0 \tag{4.1a}$$

$$[a_{ij}, b_{kl}] \;=\; -h_1 \delta_{ik} h_{jl} - h_2 \delta_{il} \delta_{jk} \tag{4.1b}$$

where h_1 and h_2 are elements of R. Then:

(a) If either $i = j$ or $n \geqslant 3$ (or both), we have $[a_{ij}, h_1] = 0$.

(b) If $i \neq j$, we have $[a_{ij}, h_1 + h_2] = 0$.

Proof. For any indices $i, j, k, l, r, s \in [n]$ we have the Jacobi identity

$$[a_{ij}, [a_{kl}, b_{rs}]] \;+\; [a_{kl}, [b_{rs}, a_{ij}]] \;+\; [b_{rs}, [a_{ij}, a_{kl}]] \;=\; 0 \,. \tag{4.2}$$

Now fix indices i, j, l such that $j \neq l$ (here we use $n \geqslant 2$) and take $k = r = i$ and $s = l$. Using the hypotheses (4.1), we obtain

$$[a_{ij}, h_1 + h_2 \delta_{il}] \;=\; 0 \quad \text{whenever } j \neq l \,. \tag{4.3}$$

If $i = j$ or $n \geqslant 3$, then we can choose l to be different from both i and j, and conclude that $[a_{ij}, h_1] = 0$. If $i \neq j$, then we can choose $l = i$ and conclude that $[a_{ij}, h_1 + h_2] = 0$. □

Proof of Proposition 1.4. Let us first consider the case in which A is symmetric and we seek a result with col-det. The result is trivial when $n = 1$, so we can assume that $n \geqslant 2$ and apply Lemma 4.1 with $h_1 = h_2 = h$. The conclusions of Lemma 4.1, together with the supplementary hypotheses (i) or (ii) of Proposition 1.4(a) in case $n = 2$, ensure that $[a_{ij}, h] = 0$ for all i, j and hence that we can push all h factors freely to the left.

The first two steps in the proof are identical to those in Proposition 3.1: we therefore have

$$\sum_L (\text{col-det}\, A_{LI})\, (\text{col-det}\, B_{LJ}) \;=\; \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1, \cdots, l_r \in [m]} a_{l_1 i_{\sigma(1)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1} \cdots b_{l_r j_r} \,. \tag{4.4}$$

We now push the factor $b_{l_1 j_1}$ to the left:

$$\sum_L (\text{col-det}\, A_{LI})\, (\text{col-det}\, B_{LJ}) \tag{4.5a}$$

$$\begin{aligned}
= \quad & \sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_1,\ldots,l_r\in[n]} a_{l_1 i_{\sigma(1)}} \Bigg[b_{l_1 j_1} a_{l_2 i_{\sigma(2)}} \cdots a_{l_r i_{\sigma(r)}} \\
& -h \sum_{s=2}^{r} (\delta_{i_{\sigma(s)} j_1} \delta_{l_1 l_s} + \delta_{i_{\sigma(s)} l_1} \delta_{l_s j_1})\, a_{l_2 i_{\sigma(2)}} \cdots a_{l_{s-1} i_{\sigma(s-1)}} a_{l_{s+1} i_{\sigma(s+1)}} \cdots a_{l_r i_{\sigma(r)}} \Bigg] \times \\
& \qquad b_{l_2 j_2} \cdots b_{l_r j_r}
\end{aligned} \tag{4.5b}$$

$$\begin{aligned}
= \quad & \sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}(\sigma) \sum_{l_2,\ldots,l_r\in[n]} \Bigg[(A^{\mathrm{T}} B)_{i_{\sigma(1)} j_1} a_{l_2 i_{\sigma(2)}} \cdots a_{l_r i_{\sigma(r)}} \\
& -h \sum_{s=2}^{r} \delta_{i_{\sigma(s)} j_1} a_{l_2 i_{\sigma(2)}} \cdots a_{l_{s-1} i_{\sigma(s-1)}} a_{l_s i_{\sigma(1)}} a_{l_{s+1} i_{\sigma(s+1)}} \cdots a_{l_r i_{\sigma(r)}} \\
& -h \sum_{s=2}^{r} \delta_{l_s j_1} a_{l_2 i_{\sigma(2)}} \cdots a_{l_{s-1} i_{\sigma(s-1)}} a_{i_{\sigma(s)} i_{\sigma(1)}} a_{l_{s+1} i_{\sigma(s+1)}} \cdots a_{l_r i_{\sigma(r)}} \Bigg] b_{l_2 j_2} \cdots b_{l_r j_r}
\end{aligned} \tag{4.5c}$$

$$= \quad \sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}(\sigma) \left[(A^{\mathrm{T}} B)_{i_{\sigma(1)} j_1} + h \sum_{s=2}^{r} \delta_{i_{\sigma(1)} j_1} \right] \sum_{l_2,\ldots,l_r\in[n]} a_{l_2 i_{\sigma(2)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_2 j_2} \cdots b_{l_r j_r} \tag{4.5d}$$

$$= \quad \sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}(\sigma) \left[A^{\mathrm{T}} B + h(r-1) \right]_{i_{\sigma(1)} j_1} \sum_{l_2,\ldots,l_r\in[n]} a_{l_2 i_{\sigma(2)}} \cdots a_{l_r i_{\sigma(r)}} b_{l_2 j_2} \cdots b_{l_r j_r}\ , \tag{4.5e}$$

where in the interchange of $\sigma(1)$ with $\sigma(s)$ we dropped out the third summand within square brackets because it is symmetric, hence vanishing when summed over σ with a factor $\operatorname{sgn}(\sigma)$. Iterating this procedure, we obtain the desired final result

$$\sum_L (\det A_{LI})(\det B_{LJ}) \tag{4.6a}$$

$$= \quad \sum_{\sigma\in\mathcal{S}_r} \operatorname{sgn}(\sigma) \left[A^{\mathrm{T}} B + (r-1)\, h \right]_{i_{\sigma(1)} j_1} \left[A^{\mathrm{T}} B + (r-2)\, h \right]_{i_{\sigma(2)} j_2} \cdots \left[A^{\mathrm{T}} B \right]_{i_{\sigma(r)} j_r} \tag{4.6b}$$

$$= \quad \text{col-det} \left[(A^{\mathrm{T}} B)_{IJ} + Q_{\mathrm{col}} \right]\ . \tag{4.6c}$$

The proof of part (b) is similar. □

Remark. In the proof of Proposition 1.4 contained in the first preprint version of this paper, we inadvertently assumed without justification that $[a_{ij}, h] = 0$.[15] We have now

[15] We realized our error when we read [16] and in particular their Lemmas 21 and 25, which correspond to our Lemma 4.1 albeit restricted to the case where $2x = 0$ implies $x = 0$ in the ring R.

repaired the error by including Lemma 4.1 and adding the extra conditions (i) or (ii) in Proposition 1.4 when $n = 2$. The importance of these extra conditions is illustrated by the following example. $\square$

Example 4.2 Let us consider the general case of 2×2 matrices (i.e. $n = 2$) under the hypotheses that A is column-pseudo-commutative and that $[a_{ij}, b_{kl}] = -h(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$. With $Q = \begin{pmatrix} h & 0 \\ 0 & 0 \end{pmatrix}$, we have

$$\text{col-det}(A^{\mathrm{T}}B + Q) \;-\; (\text{col-det}\, A^{\mathrm{T}})(\text{col-det}\, B) \;=\; -[a_{12}, h]b_{12} \;+\; (a_{21} - a_{12})hb_{12}\;. \tag{4.7}$$

The second term illustrates the importance of assuming that A is symmetric. The first term illustrates the importance of assuming that $[a_{12}, h] = 0$ when $n = 2$ (i.e. when this is not guaranteed by Lemma 4.1).

To see that $[a_{12}, h] \neq 0$ and $(4.7) \neq 0$ can actually arise, let R be the ring of 2×2 matrices with elements in the field $GF(2)$, and consider the elements $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ in R. Note that $[\alpha, \beta] = \beta$. Now let $A = \begin{pmatrix} 0 & \alpha \\ \alpha & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & \beta \\ \beta & 0 \end{pmatrix}$. These matrices satisfy $[a_{ij}, a_{kl}] = [b_{ij}, b_{kl}] = 0$ and $[a_{ij}, b_{kl}] = -h(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$ with $h = \beta$ (this works when $ij = kl = 11$ or 22 because $2h = 0$). But $[a_{12}, h] = [\alpha, \beta] = \beta \neq 0$ and $(4.7) = \beta^2 \neq 0$. $\square$

Proof of Proposition 1.5. Let us first consider the case in which A is antisymmetric off-diagonal and we seek a result with col-per. We follow closely the proof of the previous Proposition 1.4, starting from the expression in which the a's and b's are ordered and first pushing the factor $b_{l_1 j_1}$ to the left:

$$\sum_{\sigma\in\mathcal{S}_r}\sum_{l_1,\cdots,l_r\in[n]} a_{l_1 i_{\sigma(1)}}\cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1}\cdots b_{l_r j_r} \tag{4.8a}$$

$$\begin{aligned} = & \sum_{\sigma\in\mathcal{S}_r}\sum_{l_1,\ldots,l_r\in[n]} a_{l_1 i_{\sigma(1)}} \Bigg[b_{l_1 j_1} a_{l_2 i_{\sigma(2)}}\cdots a_{l_r i_{\sigma(r)}} \\ & -h \sum_{s=2}^{r} (\delta_{i_{\sigma(s)} j_1}\delta_{l_1 l_s} - \delta_{i_{\sigma(s)} l_1}\delta_{l_s j_1})\, a_{l_2 i_{\sigma(2)}}\cdots a_{l_{s-1} i_{\sigma(s-1)}} a_{l_{s+1} i_{\sigma(s+1)}}\cdots a_{l_r i_{\sigma(r)}} \Bigg] \times \\ & \qquad b_{l_2 j_2}\cdots b_{l_r j_r} \end{aligned} \tag{4.8b}$$

$$\begin{aligned} = & \sum_{\sigma\in\mathcal{S}_r}\sum_{l_2,\ldots,l_r\in[n]} \Bigg[(A^{\mathrm{T}}B)_{i_{\sigma(1)} j_1} a_{l_2 i_{\sigma(2)}}\cdots a_{l_r i_{\sigma(r)}} \\ & -h \sum_{s=2}^{r} \delta_{i_{\sigma(s)} j_1} a_{l_2 i_{\sigma(2)}}\cdots a_{l_{s-1} i_{\sigma(s-1)}} a_{l_s i_{\sigma(1)}} a_{l_{s+1} i_{\sigma(s+1)}}\cdots a_{l_r i_{\sigma(r)}} \\ & +h \sum_{s=2}^{r} \delta_{l_s j_1} a_{l_2 i_{\sigma(2)}}\cdots a_{l_{s-1} i_{\sigma(s-1)}} a_{i_{\sigma(s)} i_{\sigma(1)}} a_{l_{s+1} i_{\sigma(s+1)}}\cdots a_{l_r i_{\sigma(r)}} \Bigg] b_{l_2 j_2}\cdots b_{l_r j_r} \end{aligned} \tag{4.8c}$$

$$= \sum_{\sigma\in\mathcal{S}_r} \Bigg[(A^{\mathrm{T}}B)_{i_{\sigma(1)} j_1} - h\sum_{s=2}^{r} \delta_{i_{\sigma(1)} j_1} \Bigg] \sum_{l_2,\ldots,l_r\in[n]} a_{l_2 i_{\sigma(2)}}\cdots a_{l_r i_{\sigma(r)}} b_{l_2 j_2}\cdots b_{l_r j_r} \tag{4.8d}$$

$$= \sum_{\sigma\in\mathcal{S}_r} \big[A^{\mathrm{T}}B - h(r-1)\big]_{i_{\sigma(1)} j_1} \sum_{l_2,\ldots,l_r\in[n]} a_{l_2 i_{\sigma(2)}}\cdots a_{l_r i_{\sigma(r)}} b_{l_2 j_2}\cdots b_{l_r j_r}\ , \tag{4.8e}$$

where now in the interchange of $\sigma(1)$ with $\sigma(s)$ there is *no* change of sign in the second summand because there is *no* $\mathrm{sgn}(\sigma)$ factor, and we dropped out the third summand within square brackets because it is *anti*symmetric and hence vanishing when summed over σ (note that we used here only the *off-diagonal* antisymmetry of A). Iterating this procedure, we obtain the desired final result

$$\sum_{\sigma\in\mathcal{S}_r}\sum_{l_1,\cdots,l_r\in[n]} a_{l_1 i_{\sigma(1)}}\cdots a_{l_r i_{\sigma(r)}} b_{l_1 j_1}\cdots b_{l_r j_r} \tag{4.9a}$$

$$= \sum_{\sigma\in\mathcal{S}_r} \big[A^{\mathrm{T}}B - (r-1)\,h\big]_{i_{\sigma(1)} j_1} \big[A^{\mathrm{T}}B - (r-2)\,h\big]_{i_{\sigma(2)} j_2}\cdots \big[A^{\mathrm{T}}B\big]_{i_{\sigma(r)} j_r} \tag{4.9b}$$

$$= \text{col-per}\,\big[(A^{\mathrm{T}}B)_{IJ} - Q_{\mathrm{col}}\big]\ . \tag{4.9c}$$

The proof of part (b) is similar. □

Remark. In the proof of Proposition 1.5 contained in the first preprint version of this paper, we inadvertently assumed without justification that $[a_{ij}, h] = 0$ or $[b_{ij}, h] = 0$. We have now repaired the error by including these hypotheses explicitly. Alternatively, if we assume that $[a_{ij}, a_{il}] = 0$ for all i, j, l, then Lemma 4.1 with $h_1 = -h_2 = h$ implies that $[a_{ij}, h] = 0$ for all i, j whenever $n \geqslant 3$ (and analogously for B). $\square$

5 Further generalizations?

In this section we investigate whether any of our results can be further generalized. We always assume, for simplicity, that $[a_{ij}, a_{kl}] = 0$ and $[b_{ij}, b_{kl}] = 0$ for all i, j, k, l, and we ask the following question:

In which cases does there exist a diagonal matrix Q such that

$$(\det A)(\det B) \;=\; \text{col-det}(A^{\mathrm{T}}B + Q) \; ? \tag{5.1}$$

And what are the elements of Q?

More specifically, we investigate two types of possible extensions:

- Can we allow commutators $[a_{ij}, b_{kl}] = -h_1\delta_{ik}\delta_{jl} - h_2\delta_{il}\delta_{jk}$ beyond the ordinary case ($h_2 = 0$) and the symmetric case ($h_2 = h_1$)?
- In the symmetric case $h_2 = h_1$, is it really necessary that the matrix A be symmetric?

Our approach is to perform exact calculations for small matrices, in order to conjecture possible identities and to rule out others. We shall find, not surprisingly, that there exists an antisymmetric case ($h_2 = -h_1$) corresponding to the Howe–Umeda–Kostant–Sahi identity [23, 33] with n even; but it appears that there exists also *another* antisymmetric case, with a *different* "quantum correction", when n is odd.

We shall perform our calculations in two alternative frameworks:

Abstract framework. We work in the ring $\mathbb{Z}\langle A, B\rangle[h_1, h_2]/\mathcal{R}$ generated by noncommuting indeterminates $A = (a_{ij})_{i,j=1}^n$ and $B = (b_{ij})_{i,j=1}^n$ and commuting indeterminates h_1 and h_2 modulo the two-sided ideal $\mathcal{R}$ generated by the relations $[a_{ij}, a_{kl}] = 0$, $[b_{ij}, b_{kl}] = 0$ and $[a_{ij}, b_{kl}] = -h_1\delta_{ik}\delta_{jl} - h_2\delta_{il}\delta_{jk}$. We then introduce a matrix $Q = \mathrm{diag}(q_1, \ldots, q_n)$ of central elements and expand out the polynomial

$$f(A, B, Q) \;:=\; (\det A)(\det B) \;-\; \text{col-det}(A^{\mathrm{T}}B + Q) \,. \tag{5.2}$$

We ask whether $q_1, \ldots, q_n$ can be chosen so that either

- $f(A, B, Q) \equiv 0$, or
- $f(A, B, Q) \equiv 0$ modulo the ideal corresponding to the symmetry or antisymmetry of A and/or B. More precisely, there are six nontrivial ideals to be considered:
    - A symmetric

– B symmetric

– A and B symmetric

– A antisymmetric

– B antisymmetric

– A and B antisymmetric

In the first three cases we have $h_2 = h_1$, while in the latter three cases we have $h_2 = -h_1$. (If A and B have opposite symmetries, then $h_1 = h_2 = 0$ and we are back in the commutative case.)

Concrete framework. We work in the Weyl algebra $A_{n\times n}(K)$ over a field K of characteristic 0 (e.g. $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$) generated by a matrix $X = (x_{ij})_{i,j=1}^n$ of indeterminates and the corresponding matrix $\partial = (\partial/\partial x_{ij})_{i,j=1}^n$ of partial derivatives. The matrices X and $Y = h\partial$ (where $h \in K$, $h \neq 0$) satisfy the commutation relations

$$[x_{ij}, x_{kl}] = 0\,; \qquad [y_{ij}, y_{kl}] = 0\,; \qquad [x_{ij}, y_{kl}] = -h\,\delta_{ik}\,\delta_{jl}\ . \tag{5.3}$$

Now consider the matrices $A = (a_{ij})_{i,j=1}^n$ and $B = (b_{ij})_{i,j=1}^n$ defined by

$$A = \alpha X + \beta X^{\mathrm{T}}\,; \qquad B = \gamma Y + \delta Y^{\mathrm{T}} \tag{5.4}$$

(where $\alpha, \beta, \gamma, \delta \in K$), so that

$$\begin{aligned} [a_{ij}, a_{kl}] &= 0 && (5.5a)\\ [b_{ij}, b_{kl}] &= 0 && (5.5b)\\ [a_{ij}, b_{kl}] &= -h\left[(\alpha\gamma + \beta\delta)\,\delta_{ik}\delta_{jl} \ + \ (\alpha\delta + \beta\gamma)\,\delta_{il}\delta_{jk}\right]\ . && (5.5c) \end{aligned}$$

Some cases of special interest are:

- $\alpha\delta + \beta\gamma = 0$, so that (5.5) reduces to the starting case (5.3) with the replacement $h \to h\,(\alpha\gamma + \beta\delta)$;
- $\beta = \alpha$, so that A is symmetric and $\alpha\gamma + \beta\delta = \alpha\delta + \beta\gamma = \alpha(\gamma + \delta)$;
- $\delta = \gamma$, so that B is symmetric and $\alpha\gamma + \beta\delta = \alpha\delta + \beta\gamma = (\alpha + \beta)\gamma$;
- $\beta = \alpha$ and $\delta = \gamma$, so that both A and B are symmetric and $\alpha\gamma+\beta\delta = \alpha\delta+\beta\gamma = 2\alpha\gamma$;
- $\beta = -\alpha$, so that A is antisymmetric and $\alpha\gamma + \beta\delta = -(\alpha\delta + \beta\gamma) = \alpha(\gamma - \delta)$;
- $\delta = -\gamma$, so that B is antisymmetric and $\alpha\gamma + \beta\delta = -(\alpha\delta + \beta\gamma) = (\alpha - \beta)\gamma$;
- $\beta = -\alpha$ and $\delta = -\gamma$, so that both A and B are antisymmetric and $\alpha\gamma + \beta\delta = -(\alpha\delta + \beta\gamma) = 2\alpha\gamma$.

We shall use the abstract framework to conjecture possible identities, and the concrete framework to find counterexamples.

In the abstract framework we know the following general solutions:

(a) Capelli: For any n, with $h_2 = 0$ and $q_i = (n-i)h_1$.

(b) Turnbull: For any n, with $h_2 = h_1$, A symmetric and $q_i = (n-i)h_1$. [Turnbull requires symmetry of both A and B, but we have shown in Proposition 1.4 that only the symmetry of A is needed.]

(c) Howe–Umeda–Konstant–Sahi: For any *even* n, with $h_2 = -h_1$, A and B antisymmetric and $q_i = (n-i-1)h_1$. [But we suspect that only B need be assumed antisymmetric: see below.]

Our results lead us to *conjecture* a fourth general solution:

(d) For any *odd* n, with $h_2 = -h_1$, A or B antisymmetric (or both) and $q_i = (n-i)h_1$.

We shall show — with the aid of the symbolic-algebra package MATHEMATICA — that for $n \leqslant 5$ there are no other solutions, except for a special antisymmetric solution at $n = 2$.

Case $n = 1$. This case is trivial and we have $Q = 0$.

Case $n = 2$. For $n = 2$ we find the following solutions:

- Capelli: $h_2 = 0$, $q_1 = h_1$, $q_2 = 0$.
- Turnbull: $h_2 = h_1$, A symmetric, $q_1 = h_1$, $q_2 = 0$.
- Howe–Umeda–Konstant–Sahi: $h_2 = -h_1$, B antisymmetric, $q_1 = 0$, $q_2 = -h_1$.
- New antisymmetric solution: $h_2 = -h_1$, A antisymmetric, $q_1 = h_1$, $q_2 = 0$.

We see, therefore, that for $n = 2$ the Howe–Umeda–Konstant–Sahi solution requires only the antisymmetry of B, not of A. Furthermore, there is an additional antisymmetric solution with a different quantum correction; this solution does not, however, seem to generalize to larger even n (see $n = 4$ below).

There are no solutions for $n = 2$ besided the ones listed above. Indeed, even in the more restrictive concrete framework we can show that unless we are in one of the foregoing cases, there is no choice of q_1 and q_2 for which (5.1) holds.

Case $n = 3$. For $n = 3$ we find the following solutions:

- Capelli: $h_2 = 0$, $q_i = (n-i)h_1$.
- Turnbull: $h_2 = h_1$, A symmetric, $q_i = (n-i)h_1$.
- New antisymmetric solution: $h_2 = -h_1$, either A or B (or both) antisymmetric, $q_i = (n-i)h_1$.

There are no solutions for $n = 3$ besided the ones listed above. Indeed, even in the more restrictive concrete framework we can show that unless we are in one of the foregoing cases, there is no choice of q_1, q_2 and q_3 for which (5.1) holds.

Case $n = 4$. For $n = 4$ we find the following solutions:

- Capelli: $h_2 = 0$, $q_i = (n-i)h_1$.
- Turnbull: $h_2 = h_1$, A symmetric, $q_i = (n-i)h_1$.
- Howe–Umeda–Konstant–Sahi: $h_2 = -h_1$, B antisymmetric, $q_i = (n-i-1)h_1$.

There are no solutions for $n = 4$ besided the ones listed above. (With our available computer facilities we were able to perform the computation for $n = 4, 5$ only in the abstract framework, not in the concrete framework.)

Case $n = 5$. For $n = 5$ we find the following solutions:

- Capelli: $h_2 = 0$, $q_i = (n-i)h_1$.
- Turnbull: $h_2 = h_1$, A symmetric, $q_i = (n-i)h_1$.
- New antisymmetric solution: $h_2 = -h_1$, either A or B (or both) antisymmetric, $q_i = (n-i)h_1$.

There are no solutions for $n = 5$ besided the ones listed above.

These results for $n \leqslant 5$ lead us to make the following conjectures:

Conjecture 5.1 (generalized Howe–Umeda–Konstant–Sahi) *If n is* even *and B is antisymmetric (but A can be arbitrary), then (5.1) holds with*

$$q_i \;=\; (n-i-1)h_1 \; . \tag{5.6}$$

Conjecture 5.2 (new antisymmetric solution) *If n is* odd *and at least one of the matrices A and B is antisymmetric, then (5.1) holds with*

$$q_i \;=\; (n-i)h_1 \; . \tag{5.7}$$

A A generalized "Cayley" identity

In this appendix we use Proposition 1.2 to prove a generalization of the "Cayley" identity (1.24).

Proposition A.1 (generalized Cayley identity) *Let R be a (not-necessarily-commutative) ring, and let A and B be $n \times n$ matrices with elements in R. Suppose that*

$$[a_{ij}, a_{kl}] \;=\; 0 \tag{A.1a}$$

$$[a_{ij}, b_{kl}] \;=\; -\delta_{ik} h_{jl} \tag{A.1b}$$

$$[a_{ij}, h_{kl}] \;=\; 0 \tag{A.1c}$$

where $H = (h_{jl})_{j,l=1}^{n}$ is a matrix with elements in R. Then, for any $I, J \subseteq [n]$ of cardinality $|I| = |J| = r$ and any nonnegative integer s, we have

$$\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\det\, (A^{\mathrm{T}})_{IL})(\text{col-det}\, B_{LJ})(\det A)^s \;=\; (\det A)^s \;\text{col-det}[(A^{\mathrm{T}}B + sH)_{IJ} + Q_{\text{col}}] \;, \tag{A.2}$$

where

$$(Q_{\text{col}})_{\alpha\beta} \;=\; (r - \beta)\, h_{i_\alpha j_\beta} \tag{A.3}$$

for $1 \leqslant \alpha, \beta \leqslant r$.

The identity (A.2) clearly generalizes the identity (1.9) from Proposition 1.2(a), to which it reduces when $s = 0$ (and A is commutative: see also Remarks 2 and 3 below).

The key fact needed in the proof of Proposition A.1 is the following identity:

Lemma A.2 *Let A and B be as in Proposition A.1. Then, for all $i, j \in [n]$ and all nonnegative integers s, we have*

$$[(A^{\mathrm{T}}B)_{ij}, (\det A)^s] \;=\; s\, h_{ij}\, (\det A)^s \;=\; s\, (\det A)^s\, h_{ij} \;. \tag{A.4}$$

Proof. A simple computation using the hypotheses $[a, a] = 0$ shows that

$$[(A^{\mathrm{T}}B)_{ij}, a_{kl}] \;=\; a_{ki} h_{lj} \;. \tag{A.5}$$

We therefore have (using the hypotheses $[a, h] = 0$)

$$[(A^{\mathrm{T}}B)_{ij}, \det A]$$

$$= \sum_{\sigma \in \mathcal{S}_n} \sum_{r=1}^{n} \text{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(r-1),r-1}\; [(A^{\mathrm{T}}B)_{ij}, a_{\sigma(r)r}]\, a_{\sigma(r+1),r+1} \cdots a_{\sigma(n)n} \tag{A.6a}$$

$$= \sum_{r=1}^{n} h_{rj} \sum_{\sigma \in \mathcal{S}_n} \text{sgn}(\sigma)\, a_{\sigma(1)1} \cdots a_{\sigma(r-1),r-1}\, a_{\sigma(r)i}\, a_{\sigma(r+1),r+1} \cdots a_{\sigma(n)n} \tag{A.6b}$$

$$= \sum_{r=1}^{n} h_{rj}\, \delta_{ir}\, (\det A) \tag{A.6c}$$

$$= h_{ij}\, (\det A) \;, \tag{A.6d}$$

where the third equality used the fact that the determinant of a (commutative) matrix with two equal columns is zero (so that the terms with $i \neq r$ vanish). This proves (A.4) for the case $s = 1$. The general case easily follows by induction using $[x, yz] = [x, y]z + y[x, z]$ along with the hypotheses $[a, h] = 0$. $\square$

Proof of Proposition A.1. By Proposition 1.2 we have

$$\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\det (A^{\mathrm{T}})_{IL})(\text{col-det}\, B_{LJ})(\det A)^s \;=\; \text{col-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}}]\, (\det A)^s \,. \tag{A.7}$$

Now let us work on the right-hand side. By Lemma A.2 we have

$$[((A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}})_{\alpha\beta}, (\det A)^s] \;=\; s\, (\det A)^s\, h_{i_\alpha j_\beta} \tag{A.8}$$

(since the matrix elements of Q_{col} commute with those of A), or in other words

$$((A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}})_{\alpha\beta}\, (\det A)^s \;=\; (\det A)^s\, ((A^{\mathrm{T}}B + sH)_{IJ} + Q_{\mathrm{col}})_{\alpha\beta} \,. \tag{A.9}$$

Now expand out $\text{col-det}[(A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}}]$ and right-multiply it by $(\det A)^s$; repeatedly using (A.9) to push $(\det A)^s$ to the left, we obtain [using the shorthands $M = (A^{\mathrm{T}}B)_{IJ} + Q_{\mathrm{col}}$ and $M' = (A^{\mathrm{T}}B + sH)_{IJ} + Q_{\mathrm{col}}$]

$$\begin{aligned}
(\text{col-det}\, M)\, (\det A)^s \;&=\; \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma)\, M_{\sigma(1)1} \cdots M_{\sigma(r)r}\, (\det A)^s && \text{(A.10a)} \\
&=\; \sum_{\sigma \in \mathcal{S}_r} \operatorname{sgn}(\sigma)\, (\det A)^s\, M'_{\sigma(1)1} \cdots M'_{\sigma(r)r} && \text{(A.10b)} \\
&=\; (\det A)^s\, (\text{col-det}\, M') \,, && \text{(A.10c)}
\end{aligned}$$

as was to be proved. $\square$

Specializing Proposition A.1 to the Weyl algebra $R = A_{n\times n}(K)$ over a field K of characteristic 0, we obtain:

Corollary A.3 (Cayley identity, Weyl algebra version) *Let $X = (x_{ij})_{i,j=1}^n$ be a square matrix of commuting indeterminates, and let $\partial = (\partial/\partial x_{ij})_{i,j=1}^n$ be the corresponding matrix of partial derivatives. Then, for any $I, J \subseteq [n]$ of cardinality $|I| = |J| = r$ and any nonnegative integer s, we have*

$$\sum_{\substack{L \subseteq [n] \\ |L| = r}} (\det (X^{\mathrm{T}})_{IL})(\det \partial_{LJ})(\det X)^s \;=\; (\det X)^s\, \text{col-det}[(X^{\mathrm{T}}\partial)_{IJ} + Q_{\mathrm{col}}(s)] \,, \tag{A.11}$$

where

$$Q_{\mathrm{col}}(s)_{\alpha\beta} \;=\; (s + r - \beta)\, \delta_{i_\alpha j_\beta} \tag{A.12}$$

for $1 \leqslant \alpha, \beta \leqslant r$. In particular, for $I = J = [n]$ we have

$$(\det X)(\det \partial)(\det X)^s \;=\; (\det X)^s \text{ col-det}(X^{\mathrm{T}}\partial + \Delta) \tag{A.13}$$

where

$$\Delta_{ij} = (s + n - j)\, \delta_{ij}\,. \tag{A.14}$$

Finally, let us take the identity (A.13) in the Weyl algebra and apply it to the constant polynomial $1 \in K[X]$; then all the derivatives ∂_{ij} annihilate 1, and by removing an overall factor $(\det X)$ from the left we obtain the usual Cayley identity:

Corollary A.4 (Cayley identity, traditional version) *Let $X = (x_{ij})_{i,j=1}^n$ be a square matrix of commuting indeterminates, and let $\partial = (\partial/\partial x_{ij})_{i,j=1}^n$ be the corresponding matrix of partial derivatives. Then, for any nonnegative integer s, we have*

$$(\det \partial)(\det X)^s \;=\; s(s+1)\ldots(s+n-1)\,(\det X)^{s-1}\,. \tag{A.15}$$

Remarks. 1. Let us stress that (A.11)/(A.13) are identities in the Weyl algebra; they can be applied to *any* polynomial $P \in K[X]$ to obtain an identity in the polynomial algebra. They are therefore stronger than the traditional Cayley identity (A.15), which corresponds to taking $P = 1$ only. This example suggests that it might be fruitful to investigate more generally whether identities of Bernstein–Sato type [3, 4, 17, 36] can be extended in a useful way to identities in the Weyl algebra.

2. The hypothesis in Proposition A.1 that $[a_{ij}, h_{kl}] = 0$ can be avoided when $n \geqslant 2$, since by a simple modification of Lemma 3.4 it can be shown that $[a_{ij}, a_{kl}] = 0$ and $[a_{ij}, b_{kl}] = -\delta_{ik} h_{jl}$ for all i, j, k, l *implies* $[a_{ij}, h_{kl}] = 0$ for all i, j, k, l, *provided that* $n \geqslant 2$. For $n = 1$, by contrast, this implication is in general false, and neither Proposition A.1 nor Lemma A.2 holds when $[a, h] \neq 0$.

3. It would be interesting to know whether the hypothesis on $[a, a]$ commutators in Proposition A.1 can be weakened from commutativity to column-pseudo-commutativity (of course replacing all occurrences of det by col-det) or, more modestly, to column-commutativity. When $[a, a] \neq 0$, (A.5) must be replaced by

$$[(A^{\mathrm{T}}B)_{ij}, a_{kl}] \;=\; a_{ki} h_{lj} \;+\; \sum_m [a_{mi}, a_{kl}]\, b_{mj}\;, \tag{A.16}$$

and we do not know how to handle the extra terms in (A.6).

4. For simplicity we have assumed that s is a nonnegative integer, so that the formulae make sense in arbitrary rings R. But in specific applications we can often allow s to be an arbitrary real or complex number, or an indeterminate. For instance, (A.11)/(A.13) make sense in a ring of differential operators (with C^∞ or analytic coefficients) over a connected open set $D \subset \mathbb{R}^{n\times n}$ or $\mathbb{C}^{n\times n}$ where $\det X$ is nonvanishing; here s is an arbitrary real or complex number, and $(\det X)^s$ denotes any fixed branch of the corresponding analytic function. Alternatively, these formulae can be interpreted algebraically, with s treated as an indeterminate, in a well-known way [17, pp. 93–94] [36, pp. 96 ff.].

5. The Cayley identity (A.15) has an extension to arbitrary minors,

$$\det(\partial_{IJ})\,(\det X)^s \;=\; s(s+1)\cdots(s+k-1)\,(\det X)^{s-1}\,\epsilon(I,J)\,(\det X_{I^c J^c}) \qquad \text{(A.17)}$$

where $|I| = |J| = k$ and $\epsilon(I,J) = (-1)^{\sum_{i\in I} i + \sum_{j\in J} j}$. Unfortunately we do *not* know how to derive (A.17) from the Capelli identity (we always seem to get *sums* over minors rather than one particular minor). See [11] for alternate combinatorial proofs of the Cayley identity, which *do* include the all-minors version (A.17). $\square$

Acknowledgments

We wish to thank Luigi Cantini for drawing our attention to references [15, 43]. We also thank an anonymous referee for drawing our attention to [16] and for other helpful comments.

We thank the Isaac Newton Institute for Mathematical Sciences, University of Cambridge, for support during the programme on Combinatorics and Statistical Mechanics (January–June 2008), where this work was (almost) completed.

This research was supported in part by U.S. National Science Foundation grant PHY–0424082.

References

[1] A. Abdesselam, T. Crilly and 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.

[2] M. Atiyah, R. Bott and V.K. Patodi, On the heat equation and the index theorem, *Invent. Math.* **19**, 279–330 (1973) and erratum **28**, 277–280 (1975).

[3] I.N. Bernšteĭn, The analytic continuation of generalized functions with respect to a parameter, *Funkcional. Anal. i Priložen.* **6**(4), 26–40 (1972) [= *Funct. Anal. Appl.* **6**, 273–285 (1972)].

[4] J.-E. Björk, *Rings of Differential Operators* (North-Holland, Amsterdam–Oxford–New York, 1979).

[5] A. Capelli, Fondamenti di una teoria generale delle forme algebriche, *Atti Reale Accad. Lincei, Mem. Classe Sci. Fis. Mat. Nat.* (serie 3) **12**, 529–598 (1882).

[6] A. Capelli, Ueber die Zurückführung der Cayley'schen Operation Ω auf gewöhnliche Polar-Operationen, *Math. Annalen* **29**, 331–338 (1887).

[7] 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**, 1–17 (1888).

[8] A. Capelli, Sur les opérations dans la théorie des formes algébriques, *Math. Annalen* **37**, 1–37 (1890).

[9] A. Capelli, *Lezioni sulla Teoria delle Forme Algebriche* (Pellerano, Napoli, 1902).

[10] S. Caracciolo and A. Sportiello, Noncommutative determinants, Cauchy–Binet formulae, and Capelli-type identities. II. Grassmann and quantum oscillator algebra representation, in preparation.

[11] S. Caracciolo, A. Sportiello and A.D. Sokal, Combinatorial proofs of Cayley-type identities for derivatives of determinants and pfaffians, in preparation.

[12] P. Cartier and D. Foata, *Problèmes Combinatoires de Commutation et Réarrangements*, Lecture Notes in Mathematics #85 (Springer-Verlag, Berlin–Heidelberg–New York, 1969). Electronic reedition (2006), with three new appendices, available on-line at `http://www.emis.de/journals/SLC/books/cartfoa.html`

[13] A. Cayley, On linear transformations, *Cambridge and Dublin Math. J.* **1**, 104–122 (1846). [Also in *The Collected Mathematical Papers of Arthur Cayley* (Cambridge University Press, Cambridge, 1889–1897), vol. 1, pp. 95–112.]

[14] 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.]

[15] A. Chervov and G. Falqui, Manin matrices and Talalaev's formula, *J. Phys. A: Math. Theor.* **41**, 194006 (2008), arXiv:0711.2236 [math.QA] at arXiv.org.

[16] A. Chervov, G. Falqui and V. Rubtsov, Algebraic properties of Manin matrices 1, arXiv:0901.0235 [math.QA] at arXiv.org.

[17] S.C. Coutinho, *A Primer of Algebraic D-Modules*, London Mathematical Society Student Texts #33 (Cambridge University Press, Cambridge, 1995).

[18] I. Dolgachev, *Lectures on Invariant Theory*, London Mathematical Society Lecture Note Series #296 (Cambridge University Press, Cambridge, 2003).

[19] D. Foata, A noncommutative version of the matrix inversion formula, *Adv. Math.* **31**, 330–349 (1979).

[20] D. Foata and D. Zeilberger, Combinatorial proofs of Capelli's and Turnbull's identities from classical invariant theory, *Electron. J. Combin.* **1**, #R1 (1994).

[21] W. Fulton and J. Harris, *Representation Theory: A First Course* (Springer-Verlag, New York–Heidelberg, 1991), Appendix F.

[22] R. Howe, Remarks on classical invariant theory, *Trans. Amer. Math. Soc.* **313**, 539–570 (1989) and erratum **318**, 823 (1990).

[23] R. Howe and T. Umeda, The Capelli identity, the double commutant theorem, and multiplicity-free actions, *Math. Ann.* **290**, 565–619 (1991).

[24] M. Itoh, Capelli elements for the orthogonal Lie algebras, *J. Lie Theory* **10**, 463–489 (2000).

[25] M. Itoh, Capelli identities for the dual pair (O_M, Sp_N), *Math. Z.* **246**, 125–154 (2004).

[26] M. Itoh, Capelli identities for reductive dual pairs, *Adv. Math.* **194**, 345–397 (2005).

[27] M. Itoh, Two determinants in the universal enveloping algebras of the orthogonal Lie algebras, *J. Algebra* **314**, 479–506 (2007).

[28] M. Itoh and T. Umeda, On central elements in the universal enveloping algebras of the orthogonal Lie algebras, *Compositio Math.* **127**, 333–359 (2001).

[29] K. Kinoshita and M. Wakayama, Explicit Capelli identities for skew symmetric matrices, *Proc. Edinburgh Math. Soc.* **45**, 449–465 (2002).

[30] M. Konvalinka, Non-commutative Sylvester's determinantal identity, *Electron. J. Combin.* **14**, #R42 (2008), arxiv:math/0703213 [math.CO] at arXiv.org.

[31] M. Konvalinka, An inverse matrix formula in the right-quantum algebra, *Electron. J. Combin.* **15**, #R23 (2008), arxiv:math/0703203 [math.CO] at arXiv.org.

[32] M. Konvalinka and I. Pak, Non-commutative extensions of the MacMahon Master Theorem, *Adv. Math.* **216**, 29–61 (2007), arXiv:math/0607737 [math.CO] at arXiv.org.

[33] B. Kostant and S. Sahi, The Capelli identity, tube domains, and the generalized Laplace transform, *Adv. Math.* **87**, 71–92 (1991).

[34] B. Kostant and S. Sahi, Jordan algebras and Capelli identities, *Invent. Math.* **112**, 657–664 (1993).

[35] H. Kraft and C. Procesi, *Classical Invariant Theory: A Primer* (preliminary version, July 1996), available on-line at `http://www.math.unibas.ch/~kraft/Papers/KP-Primer.pdf`

[36] G.R. Krause and T.H. Lenagan, *Growth of Algebras and Gelfand–Kirillov Dimension*, rev. ed. (American Mathematical Society, Providence RI, 2000).

[37] P. Lalonde, A non-commutative version of Jacobi's equality on the cofactors of a matrix, *Discrete Math.* **158**, 161–172 (1996).

[38] Yu. I. Manin, Some remarks on Koszul algebras and quantum groups, *Ann. Inst. Fourier (Grenoble)* **37**, 191–205 (1987).

[39] Yu.I. Manin, *Quantum Groups and Non-Commutative Geometry* (Centre de Recherches Mathématiques, Université de Montréal, 1988).

[40] Y.I. Manin, *Topics in Noncommutative Geometry* (Princeton University Press, Princeton, NJ, 1991).

[41] A. Molev, *Yangians and Classical Lie Algebras*, Mathematical Surveys and Monographs #143 (American Mathematical Society, Providence, RI, 2007).

[42] A. Molev and M. Nazarov, Capelli identities for classical Lie algebras, *Math. Ann.* **313**, 315–357 (1999).

[43] E. Mukhin, V. Tarasov and A. Varchenko, A generalization of the Capelli identity, preprint (October 2006), arXiv:math/0610799 [math.QA] at arXiv.org.

[44] M. Nazarov, Capelli identities for Lie superalgebras, *Ann. Sci. École Norm. Sup.* **30**, 847–872 (1997).

[45] A. Okounkov, Quantum immanants and higher Capelli identities, *Transform. Groups* **1**, 99–126 (1996).

[46] A. Okounkov, Young basis, Wick formula, and higher Capelli identities, *Internat. Math. Res. Notices* **1996**, no. 17, 817–839.

[47] P.J. Olver, *Classical Invariant Theory*, London Mathematical Society Student Texts #44 (Cambridge University Press, Cambridge, 1999).

[48] M. Nazarov, Yangians and Capelli identities (Kirillov's seminar on representation theory), *Amer. Math. Soc. Transl. Ser. 2* **181**, 139–163 (1998).

[49] M. Noumi, T. Umeda and M. Wakayama, A quantum analogue of the Capelli identity and an elementary differential calculus on $\mathrm{GL}_q(n)$, *Duke Math. J.* **76**, 567–594 (1994).

[50] M. Noumi, T. Umeda and M. Wakayama, Dual pairs, spherical harmonics and a Capelli identity in quantum group theory, *Compositio Math.* **104**, 227–277 (1996).

[51] I. Schur, *Vorlesungen über Invariantentheorie* (Springer-Verlag, Berlin–Heidelberg–New York, 1968).

[52] P.R. Stein, On an identity from classical invariant theory, *Lin. Multilin. Alg.* **11**, 39–44 (1982).

[53] H.W. Turnbull, Symmetric determinants and the Cayley and Capelli operators, *Proc. Edinburgh Math. Soc.* (2)**8**, 76–86 (1948).

[54] T. Umeda, The Capelli identities, a century after, *Amer. Math. Soc. Transl. Ser. 2* **183**, 51–78 (1998).

[55] T. Umeda, On the proof of the Capelli identities, *Funkcial. Ekvac.* **51**, 1–15 (2008).

[56] A. Wachi, Central elements in the universal enveloping algebras for the split realization of the orthogonal Lie algebras, *Lett. Math. Phys.* **77**, 155–168 (2006).

[57] A.H. Wallace, A note on the Capelli operators associated with a symmetric matrix, *Proc. Edinburgh Math. Soc.* (2)**9**, 7–12 (1953).

[58] H. Weyl, *The Classical Groups, Their Invariants and Representations*, 2nd ed. (Princeton University Press, Princeton NJ, 1946).