Murray R. Bremner; Sara Madariaga; Luiz A. Peresi

Structure theory for the group algebra of the symmetric group, with applications to polynomial identities for the octonions

Commentationes Mathematicae Universitatis Carolinae, Vol. 57 (2016), No. 4, 413–452

Persistent URL: `http://dml.cz/dmlcz/145950`

Structure theory for the group algebra of the symmetric group, with applications to polynomial identities for the octonions

Murray R. Bremner, Sara Madariaga, Luiz A. Peresi

To our colleague Irvin Roy Hentzel on his retirement.

Abstract. This is a survey paper on applications of the representation theory of the symmetric group to the theory of polynomial identities for associative and nonassociative algebras. In §1, we present a detailed review (with complete proofs) of the classical structure theory of the group algebra $\mathbb{F}S_n$ of the symmetric group S_n over a field $\mathbb{F}$ of characteristic 0 (or $p > n$). The goal is to obtain a constructive version of the isomorphism $\psi\colon \bigoplus_\lambda M_{d_\lambda}(\mathbb{F}) \longrightarrow \mathbb{F}S_n$ where λ is a partition of n and d_λ counts the standard tableaux of shape λ. Young showed how to compute ψ; to compute its inverse, we use an efficient algorithm for representation matrices discovered by Clifton. In §2, we discuss constructive methods based on §1 which allow us to analyze the polynomial identities satisfied by a specific (non)associative algebra: fill and reduce algorithm, module generators algorithm, Bondari's algorithm for finite dimensional algebras. In §3, we study the multilinear identities satisfied by the octonion algebra $\mathbb{O}$ over a field of characteristic 0. For $n \leq 6$ we compare our computational results with earlier work of Racine, Hentzel & Peresi, Shestakov & Zhukavets. Going one step further, we verify computationally that every identity in degree 7 is a consequence of known identities of lower degree; this result is our main original contribution. This gap (no new identities in degree 7) motivates our concluding conjecture: the known identities for $n \leq 6$ generate all of the octonion identities in characteristic 0.

Keywords: symmetric group; group algebra; Young diagrams; standard tableaux; idempotents; matrix units; two-sided ideals; Wedderburn decomposition; representation theory; Clifton's algorithm; computer algebra; polynomial identities; nonassociative algebra; octonions

Classification: Primary 20C30; Secondary 16R10, 16S34, 16Z05, 17-04, 17-08, 17A50, 17A75, 17B01, 17C05, 17D05, 18D50, 20B30, 20B40, 20C40, 68W30

1. Structure theory for the group algebra of the symmetric group

In this first part, we study the structure of the group algebra $\mathbb{F}S_n$ of the symmetric group S_n on n letters. As a vector space over $\mathbb{F}$, $\mathbb{F}S_n$ has basis S_n, and the associative multiplication is defined on basis elements by the product in S_n and extended bilinearly. Our goal is to provide a useful and readable survey

DOI 10.14712/1213-7243.2015.188

of the structure of $\mathbb{F}S_n$ from a modern point of view. We assume throughout that $\mathbb{F}$ is a field of characteristic 0, but most of the results hold also for characteristic $p > n$.

By the classical structure theory of associative algebras, we know that $\mathbb{F}S_n$ is semisimple, and hence isomorphic to the direct sum of full matrix algebras with entries in division algebras over $\mathbb{F}$. In fact, each of these division algebras is isomorphic to $\mathbb{F}$, and the Wedderburn decomposition is given by two isomorphisms,

$$(1) \qquad \phi\colon \mathbb{F}S_n \longrightarrow \bigoplus_\lambda M_{d_\lambda}(\mathbb{F}), \qquad \psi\colon \bigoplus_\lambda M_{d_\lambda}(\mathbb{F}) \longrightarrow \mathbb{F}S_n,$$

where the sum is over all partitions λ of n, and d_λ is the dimension of the irreducible representation of S_n corresponding to λ.

The matrices obtained by restricting ϕ to S_n, and taking the component of ϕ for partition λ, have entries in $\{0, \pm 1\}$ and form the natural representation of S_n. We will show how to efficiently compute these matrices for all λ and all $p \in S_n$.

Each matrix algebra $M_{d_\lambda}(\mathbb{F})$ has a basis of matrix units E^λ_{ij} for $i, j = 1, \ldots, d_\lambda$ which multiply according to the standard relations,

$$(2) \qquad E^\lambda_{ij} E^\mu_{k\ell} = \delta_{\lambda\mu} \delta_{jk} E^\lambda_{i\ell}.$$

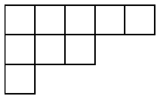
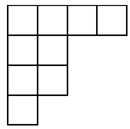
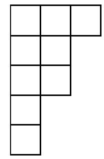
The isomorphism ψ produces elements $\psi(E^\lambda_{ij})$ in $\mathbb{F}S_n$ which obey the same equations. We show how to calculate these elements of $\mathbb{F}S_n$.

None of the material in this first part is original. We compiled the results from many sources, and attempted to make the terminology more contemporary and the notation simpler and more consistent. The structure theory of $\mathbb{F}S_n$ was original worked out by Young [52]. The proofs in Young's papers were simplified by Rutherford [45], and the theory was reformulated in more modern terminology and notation by Boerner [4], following suggestions by von Neumann and van der Waerden [50]. A substantial simplification of the algorithms for computing the matrices in the natural representation (the isomorphism ϕ) was introduced by Clifton [14], [15]. Our exposition is based on the Ph.D. thesis of Bondari [5], [6].

1.1 Young diagrams and tableaux. We start by giving the definitions and elementary properties of the basic objects in the theory. The symmetric group S_n is the group of all permutations of the set $\{1, \ldots, n\}$. We write $\lambda \vdash n$ to indicate that λ is a partition of n; that is, $\lambda = (n_1, \ldots, n_k)$ where $n = n_1 + \cdots + n_k$ and $n_1 \geq \cdots \geq n_k \geq 1$. If $n \leq 9$ then we write unambiguously $\lambda = n_1 \cdots n_k$.

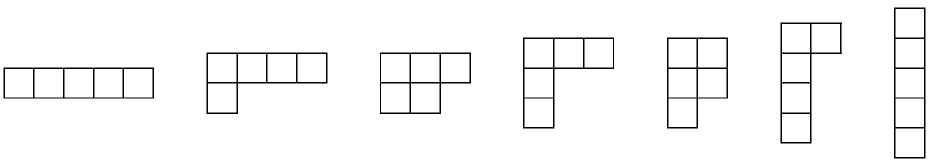
Definition 1.1. The **Young diagram** Y^λ of the partition $\lambda = (n_1, \ldots, n_k)$ consists of k left-justified rows of empty square boxes where row i contains n_i boxes.

Example 1.2. Young diagrams for some partitions of $n = 9$:

$Y^{531} =$  $Y^{4221} =$  $Y^{32211} =$ 

Definition 1.3. Suppose that $\lambda = (n_1, \dots, n_k)$ and $\lambda' = (n'_1, \dots, n'_\ell)$ are partitions of n. We say that $\lambda \prec \lambda'$ (equivalently, $Y^\lambda \prec Y^{\lambda'}$) if and only if either $n_1 < n'_1$ or there exists $i \geq 1$ such that $n_1 = n'_1, \dots, n_i = n'_i$ but $n_{i+1} < n'_{i+1}$.

Example 1.4. The seven Young diagrams for $n = 5$ in decreasing order:



Definition 1.5. A **Young tableau** T^λ of shape λ where $\lambda \vdash n$ consists of a bijective assignment of the numbers $1, \dots, n$ to the boxes in the Young diagram Y^λ. The number in row i and column j will be denoted $T(i,j)$. The sequence of numbers from left to right in row i will be denoted $T(i,-)$; the sequence of numbers from top to bottom in column j will be denoted $T(-,j)$. A tableau is **standard** if all the sequences $T(i,-)$ and $T(-,j)$ are increasing.

Remark 1.6. The hook formula for the number d_λ of standard tableaux for the Young diagram Y^λ follows; $|h_{ij}|$ counts the boxes in the hook with corner (i,j):

$$d_\lambda = \frac{n!}{\prod_{i,j} |h_{ij}|}, \qquad h_{ij} = \{\, (i,j') \mid j \leq j' \,\} \cup \{\, (i',j) \mid i \leq i' \,\}. \tag{3}$$

There is another version [4, Theorem 4.2], easier to implement on a computer:

$$d_\lambda = n! \, \frac{\prod_{i<j}(m_i - m_j)}{\prod_i m_i!}, \tag{4}$$

where $m_i = n_i + k - i$ and $i = 1, \dots, k$ for $\lambda = (n_1, \dots, n_k)$. For further details, see [35, §5.1.4, Theorem H] and [39].

Definition 1.7. Given two tableaux T and T' of shape $\lambda \vdash n$, let i be the least row index for which $T(i,-) \neq T'(i,-)$, and let j be the least column index for which $T(i,j) \neq T'(i,j)$. The **lexicographical order** (lex order) on tableaux is defined by $T \prec T'$ if and only if $T(i,j) < T'(i,j)$.

Example 1.8. The standard tableaux for $n = 5$, $\lambda = 32$ in lex order:

$$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3&5\\\cline{1-2}\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&2&5\\\hline 3&4\\\cline{1-2}\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&5\\\cline{1-2}\end{array}\qquad\begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2}\end{array}$$

Definition 1.9. For each partition $\lambda \vdash n$, the group S_n acts on the tableaux of shape λ by permuting the numbers in the boxes. For $p \in S_n$ and tableau T, the result will be denoted pT: that is, if $T(i,j) = x$ then $(pT)(i,j) = px$.

1.2 Horizontal and vertical permutations. Each tableau of shape $\lambda \vdash n$ determines certain subgroups of S_n which play an essential role in the theory.

Definition 1.10. Given a tableau T of shape $\lambda = (n_1, \dots, n_k) \vdash n$, we write $G_H(T)$ for the subgroup of S_n consisting of all **horizontal permutations** for T. These are the permutations $h \in S_n$ which leave the rows fixed as sets: for all $i = 1, \dots, k$, if $x \in T(i,-)$ then $hx \in T(i,-)$. Similarly, the subgroup $G_V(T)$ of **vertical permutations** of T consists of all permutations $v \in S_n$ which leave the columns fixed as sets: for all $j = 1, \dots, n_1$, if $x \in T(-,j)$ then $vx \in T(-,j)$.

Remark 1.11. If we regard the rows $T(i,-)$ and columns $T(-,j)$ as sets, then $G_H(T)$ and $G_V(T)$ can be defined as direct products:

$$G_H(T) = \prod_{i=1}^{k} S_{T(i,-)}, \qquad G_V(T) = \prod_{j=1}^{n_1} S_{T(-,j)},$$

where S_X denotes the group of all permutations of the set X.

Lemma 1.12. *If T is a tableau of shape $\lambda \vdash n$ then $G_H(T) \cap G_V(T) = \{\iota\}$ where $\iota \in S_n$ is the identity permutation. It follows that if $h, h' \in G_H(T)$ and $v, v' \in G_V(T)$ with $hv = h'v'$ then $h = h'$ and $v = v'$.*

PROOF: If $hv = h'v'$ then $(h')^{-1}h = v'v^{-1}$ and so both equal ι. $\square$

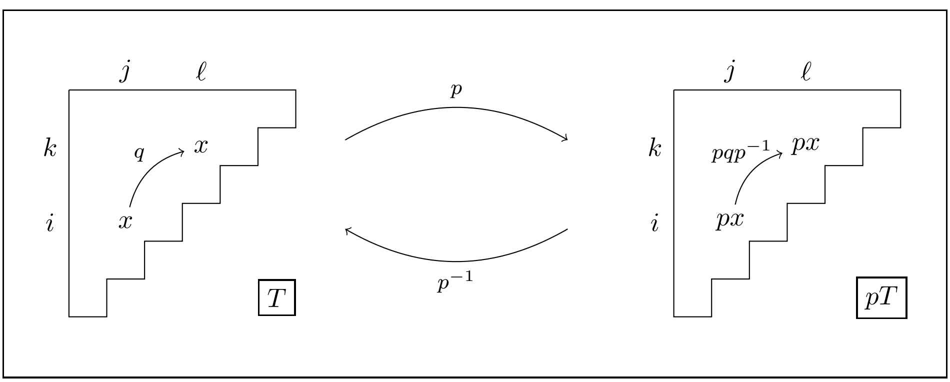


FIGURE 1. Tableaux for the proof of Lemma 1.13

Lemma 1.13. *Assume that T is a tableau of shape $\lambda \vdash n$ and $p \in S_n$.*

(a) If $h \in G_H(T)$ then $php^{-1} \in G_H(pT)$. Since conjugation by p is invertible, it is a bijection from $G_H(T)$ to $G_H(pT)$.

(b) If $v \in G_V(T)$ then $pvp^{-1} \in G_V(pT)$. Since conjugation by p is invertible, it is a bijection from $G_V(T)$ to $G_V(pT)$.

Proof: We refer to Figure 1. Suppose the permutation $q \in S_n$ moves the number x from position (i,j) of the tableau T to position (k,ℓ); this is represented by the arrow labelled q in the left tableau. Following the lower curved arrow labelled p^{-1}, then the arrow in the left tableau labelled q, and finally the upper curved arrow labelled p, we see that the permutation pqp^{-1} moves $x' = px$ from position (i,j) of the tableau pT to position (k,ℓ). This is represented by the arrow labelled pqp^{-1} in the right tableau. In particular, if $q = h \in G_H(T)$ then $i = k$, and so php^{-1} is a horizontal permutation for pT. Similarly, if $q = v \in G_V(T)$ then $j = \ell$, and so pvp^{-1} is a vertical permutation for pT. $\square$

Remark 1.14. The notation hvT indicates that we apply the vertical permutation $v \in G_V(T)$ to T and then apply the horizontal permutation $h \in G_H(T)$ to vT. However, h may not be a horizontal permutation for vT. We can rewrite this using permutations which are horizontal or vertical for the tableaux on which they act: we have $hvT = (hvh^{-1})hT$ and hvh^{-1} is a vertical permutation for hT.

1.3 Row and column intersections. The next few results investigate the intersection $T(i,-) \cap T'(-,j)$ for tableaux T and T' of shapes λ and μ.

Proposition 1.15. *Assume that $\lambda, \mu \vdash n$ with $Y^\lambda \succ Y^\mu$. For any tableaux T^λ, T^μ there exist i,j for which $T^\lambda(i,-) \cap T^\mu(-,j)$ contains at least two numbers. Thus there exist two numbers in one row of T^λ which appear in one column of T^μ.*

Proof: Write $\lambda = (n_1, \dots, n_k)$ and $\mu = (n'_1, \dots, n'_\ell)$. We make the contrary assumption that $T^\lambda(i,-) \cap T^\mu(-,j)$ contains at most one number for all $1 \leq i \leq k$ and $1 \leq j \leq n'_1$. In particular, for $i = 1$ we see that the n_1 numbers in $T^\lambda(1,-)$ belong to different columns of T^μ, and so $n_1 \leq n'_1$. But $Y^\lambda \succ Y^\mu$ implies $n_1 \geq n'_1$, and so $n_1 = n'_1$. The contrary assumption is not affected if we apply a vertical permutation to T^μ, and so there exists $v \in G_V(T^\mu)$ for which $T^\lambda(1,-) = (vT^\mu)(1,-)$ as sets; these rows contain the same numbers, possibly in different order.

We now delete the first rows of T^λ and vT^μ, obtaining tableaux $T^{\lambda'} \succ T^{\mu'}$ where λ', μ' are partitions of $n - n_1$. Both tableaux contain the numbers $\{a_1, \dots, a_{n-n_1}\} \subset \{1, \dots, n\}$ which we can identify with $\{1, \dots, n - n_1\}$. Repeating the argument of the first paragraph, we see that $n_2 = n'_2, \dots, n_k = n'_\ell$; at the end we must have $k = \ell$. This implies that $Y^\lambda = Y^\mu$, which is a contradiction. $\square$

Lemma 1.16. *Let T be a tableau of shape $\lambda = (n_1, \dots, n_k) \vdash n$. A permutation $p \in S_n$ has the form $p = hv$ for $h \in G_H(T)$ and $v \in G_V(T)$ if and only if $T(i,-) \cap (pT)(-,j)$ contains at most one number for all $i = 1, \dots, k$ and $j = 1, \dots, n_1$.*

Proof: Assume that $p = hv$ for some $h \in G_H(T)$ and $v \in G_V(T)$. Following Remark 1.14, we have $pT = hvT = (hvh^{-1})hT$ where $hvh^{-1} \in G_V(hT)$. If x, y are distinct numbers in the same row of T, then they are in the same row but different columns of hT; hence they are in different columns of $(hvh^{-1})hT = pT$.

Conversely, assume that $T(i,-) \cap (pT)(-,j)$ contains at most one number for all $i = 1, \ldots, k$ and $j = 1, \ldots, n_1$. Then the numbers in the first column of pT must appear in different rows of T. We can apply a horizontal permutation $h_1 \in G_H(T)$ so that $(h_1T)(-,1)$ is a permutation of $(pT)(-,1)$. Similarly, the numbers in the second column of pT must appear in different rows of h_1T and columns $j \geq 2$. Keeping the numbers in $(h_1T)(-,1)$ fixed, we can apply $h_2 \in G_H(T)$ so that $(h_2h_1T)(-,2)$ is a permutation of $(pT)(-,2)$. Continuing, we obtain permutations $h_1, h_2, \ldots, h_{n_1} \in G_H(T)$ so that every number in hT (where $h = h_{n_1} \cdots h_1$) is in the same column as in pT. We now apply a vertical permutation $v' \in G_V(hT)$ to obtain $v'hT = pT$. By Lemma 1.13, we have $v' = hvh^{-1}$ for some $v \in G_V(T)$. Therefore $pT = v'hT = hvh^{-1}hT = hvT$, as required. $\square$

$$
\begin{array}{c|cccccc}
T_r & \cdots & j'' & \cdots & j' & \cdots \\
\hline
\vdots & & \vdots & & \vdots & \\
i' & \cdots & z & \cdots & x & \cdots \\
\vdots & & \vdots & & \vdots & \\
i'' & \cdots & y & \cdots & & \\
\vdots & & \vdots & & &
\end{array}
\qquad\qquad
\begin{array}{c|cccccc}
T_s & \cdots & j'' & \cdots & j' & \cdots \\
\hline
\vdots & & \vdots & & \vdots & \\
i' & \cdots & z & \cdots & y & \cdots \\
\vdots & & \vdots & & \vdots & \\
i'' & \cdots & & \cdots & & \\
\vdots & & \vdots & & &
\end{array}
$$

Figure 2. Diagram for the proof of Proposition 1.17

Proposition 1.17. *Assume that $\lambda = (n_1, \ldots, n_k) \vdash n$, and let $T_1, \ldots, T_{d_\lambda}$ be the standard tableaux of shape λ in lex order. If $r > s$ then there exist $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, n_1\}$ such that $T_r(-,j) \cap T_s(i,-)$ contains at least two elements.*

Proof: Let (i', j') be the first position in which T_r and T_s have a different number. Let x, y be the numbers in position (i', j') in T_r, T_s respectively. Since $r > s$ we have $x > y$. In a standard tableau, each number in the first column is the least number that has not appeared in previous rows. Hence $j' \geq 2$. Suppose that y occurs in position (i'', j'') in T_r. Since T_r and T_s are equal up to position (i', j'), we have two cases: either $i'' = i'$ and $j'' > j'$ (y is in the same row as x but to the right), or $i'' > i'$ (y is in a lower row than x). Since $x > y$ and T_r is standard, the first case is impossible. In the second case, $x > y$ implies $j'' < j'$ (y must be in a column to the left of x). We illustrate this situation with the diagram of Figure 2. Since position (i', j'') occurs before (i', j'), the number z in this position must be the same in both T_r and T_s. Hence y, z are the two numbers in the same column of T_r and the same row of T_s. $\square$

1.4 Symmetric and alternating sums. We construct special elements of $\mathbb{F}S_n$ which will be used to define idempotents in the group algebra.

Definition 1.18. Given a tableau T of shape $\lambda \vdash n$ we define the following elements of $\mathbb{F}S_n$, where $\epsilon\colon S_n \to \{\pm 1\}$ is the sign homomorphism:

$$H_T = \sum_{h \in G_H(T)} h, \qquad\qquad V_T = \sum_{v \in G_V(T)} \epsilon(v)\, v.$$

(Classically these were called the "positive and negative symmetric groups" for T.)

Lemma 1.19. *If T is a tableau of shape $\lambda \vdash n$, and $h \in G_H(T)$, $v \in G_V(T)$, then*

$$hH_T = H_T = H_T h, \qquad vV_T = \epsilon(v)V_T = V_T v.$$

Proof: For a horizontal permutation h, the function $G_H(T) \to G_H(T)$ sending $h' \mapsto hh'$ is a bijection, and similarly for $h' \mapsto h'h$; this proves the claim for H_T. Analogous bijections hold for $G_V(T)$ and a vertical permutation v, so

$$vV_T = \sum_{v' \in G_V(T)} \epsilon(v')\, vv' = \epsilon(v)^{-1} \sum_{v' \in G_V(T)} \epsilon(v)\epsilon(v')\, vv' = \epsilon(v) \sum_{v' \in G_V(T)} \epsilon(vv')\, vv' = \epsilon(v)V_T.$$

The proof that $\epsilon(v)V_T = V_T v$ is similar. □

Proposition 1.20. *If T is a tableau of shape $\lambda \vdash n$, and $p \in S_n$, then*

$$H_{pT} = p\, H_T\, p^{-1}, \qquad V_{pT} = p\, V_T\, p^{-1}.$$

Proof: Since $\epsilon(p) = \epsilon(p^{-1})$, the result follows from Lemma 1.13. □

1.5 Idempotents and orthogonality in the group algebra. We construct idempotent elements in $\mathbb{F}S_n$ and study their orthogonality properties.

Definition 1.21. Let $T_1, \dots, T_{n!}$ be the tableaux of shape $\lambda \vdash n$ in lex order. For $1 \le i, j \le n!$ we define $s_{ij} \in S_n$ by $s_{ij}T_j = T_i$; clearly $s_{ji} = s_{ij}^{-1}$ and $s_{ij}s_{jk} = s_{ik}$. We also define these group algebra elements (omitting λ if it is understood):

$$D_i^\lambda = H_{T_i} V_{T_i} = \sum_{h \in G_H(T_i)} \sum_{v \in G_V(T_i)} \epsilon(v)\, hv.$$

Proposition 1.22. *If T is a tableau of shape $\lambda \vdash n$ then*

$$D_j = s_{ji} D_i s_{ij}. \tag{5}$$

Equivalently,

$$s_{ij} D_j = D_i s_{ij}. \tag{6}$$

Proof: Using Proposition 1.20 and the definition of s_{ij}, we obtain

$$s_{ji}D_i s_{ij} = s_{ji}H_{T_i}V_{T_i}s_{ji}^{-1} = \big[s_{ji}H_{T_i}s_{ji}^{-1}\big]\big[s_{ji}V_{T_i}s_{ji}^{-1}\big] = H_{s_{ji}T_i}V_{s_{ji}T_i} = H_{T_j}V_{T_j} = D_j.$$

This proves the first equation, and the second follows from $s_{ji} = s_{ij}^{-1}$. □

Proposition 1.23. *If $\lambda, \mu \vdash n$ with $\lambda \neq \mu$, then $D_i^\lambda D_j^\mu = 0$ for all tableaux T_i^λ, T_j^μ.*

Proof: We first assume $Y^\lambda \prec Y^\mu$. Proposition 1.15 shows that there exist two numbers k, ℓ in the same row of T^μ and the same column of T^λ; we now omit the subscripts i, j. For the transposition $t = (k, \ell)$ we have $t \in G_V(T^\lambda)$ and $t \in G_H(T^\mu)$. Using Lemma 1.19 we obtain

$$\begin{aligned} D^\lambda D^\mu &= H_{T^\lambda}V_{T^\lambda}H_{T^\mu}V_{T^\mu} = H_{T^\lambda}V_{T^\lambda}\,t^2 H_{T^\mu}V_{T^\mu} = H_{T^\lambda}(V_{T^\lambda}t)(tH_{T^\mu})V_{T^\mu} \\ &= H_{T^\lambda}(-V_{T^\lambda})(H_{T^\mu})V_{T^\mu} = -H_{T^\lambda}V_{T^\lambda}H_{T^\mu}V_{T^\mu} = -D^\lambda D^\mu. \end{aligned}$$

Hence $D^\lambda D^\mu = 0$. On the other hand, if $Y^\lambda \succ Y^\mu$, then Proposition 1.20 implies that for any $p \in S_n$ we have

$$H_{T^\lambda}pV_{T^\mu} = H_{T^\lambda}\big(pV_{T^\mu}p^{-1}\big)p = H_{T^\lambda}V_{pT^\mu}p.$$

Proposition 1.15 shows that there exist two numbers k, ℓ in the same row of T^λ and the same column of pT^μ. Then $t = (k, \ell) \in G_V(pT^\mu) \cap G_H(T^\lambda)$ and so

$$H_{T^\lambda}V_{pT^\mu}p = H_{T^\lambda}t^2V_{pT^\mu}p = (H_{T^\lambda}t)(tV_{pT^\mu})p = (H_{T^\lambda})(-V_{pT^\mu})p = -H_{T^\lambda}V_{pT^\mu}p.$$

Hence $H_{T^\lambda}V_{pT^\mu}p = 0$ and so $H_{T^\lambda}pV_{T^\mu} = 0$, for all $p \in S_n$. Therefore

$$D^\lambda D^\mu = H_{T^\lambda}V_{T^\lambda}H_{T^\mu}V_{T^\mu} = H_{T^\lambda}\bigg(\sum_{p \in S_n} x_p\, p\bigg)V_{T^\mu} = \sum_{p \in S_n} x_p\big(H_{T^\lambda}pV_{T^\mu}\big) = 0,$$

where $x_p \in \mathbb{F}$ for all $p \in S_n$. This completes the proof. □

Corollary 1.24. *Let $\lambda \vdash n$, and let $T_1, \dots, T_{d_\lambda}$ be the standard tableaux in lex order. If $i > j$ then $D_iD_j = 0$.*

Proof: We write H_i, V_i for H_{T_i}, V_{T_i}. By Proposition 1.17, there exist two numbers k, ℓ in the same column of T_i and the same row of T_j. Using the transposition $t = (k, \ell)$ and Lemma 1.19 we obtain

$$D_iD_j = H_iV_iH_jV_j = H_iV_it^2H_jV_j = H_i(V_it)(tH_j)V_j = H_i(-V_i)(H_j)V_j = -D_iD_j.$$

Therefore $D_iD_j = 0$. □

Proposition 1.25 (von Neumann's Theorem)**.** *Let $\lambda \vdash n$. For $i = 1, \dots, n!$ we have $D_i^2 = c_iD_i$ where $c_i = n!/f_i$, and f_i is the dimension of the left ideal $\mathbb{F}S_nD_i$.*

PROOF: For scalars $x_p \in \mathbb{F}$ which we will determine, we write

$$D_i^2 = \sum_{p \in S_n} x_p \, p.$$

For any $h \in G_H(T_i)$ and $v \in G_V(T_i)$ we have

$$hD_i^2 v = h\Big(\sum_{p \in S_n} x_p \, p \Big) v = \sum_{p \in S_n} x_p \, hpv,$$
$$hD_i^2 v = (hH_i)V_iH_i(V_iv) = \epsilon(v)H_iV_iH_iV_i = \epsilon(v)D_i^2.$$

Therefore

$$\sum_{p \in S_n} x_p \, hpv = \epsilon(v) \sum_{p \in S_n} x_p \, p. \tag{7}$$

Each permutation in S_n occurs once and only once on each side of this equation.

First, consider the coefficient in D_i^2 of a permutation of the form hv. On the left side of (7) take $p = \iota$, on the right side take $p = hv$, and compare coefficients:

$$x_\iota = \epsilon(v)x_{hv}.$$

Hence $x_{hv} = \epsilon(v)x_\iota$. Second, consider the coefficient in D_i^2 of a permutation q not of the form hv. Lemma 1.16 implies that there are two numbers k, ℓ in the same row of T_i and the same column of qT_i. For the transposition $t = (k, \ell)$, we have $t \in G_H(T_i)$ and $q^{-1}tq \in G_V(T_i)$. We can take $h = t$ and $v = q^{-1}tq$ in equation (7):

$$\sum_{p \in S_n} x_p \, tpq^{-1}tq = \epsilon(q^{-1}tq) \sum_{p \in S_n} x_p \, p.$$

Setting $p = q$ on both sides, we obtain

$$x_q \, tqq^{-1}tq = \epsilon(q^{-1}tq)x_q \, q,$$

and this simplifies to $x_q q = -x_q q$, implying $x_q = 0$. Combining the results of the two cases, we obtain $D_i^2 = c_i D_i$ where $c_i = x_\iota$.

It remains to show that $x_\iota = n!/f_i$. We choose a basis for the left ideal $\mathbb{F}S_nD_i$ consisting of elements $p_1D_i, \dots, p_{f_i}D_i$ where $p_1, \dots, p_{f_i} \in S_n$, and extend this to a basis of $\mathbb{F}S_n$. We regard D_i as a linear operator on $\mathbb{F}S_n$, acting by right multiplication. The matrix representing D_i with respect to our basis has the form

$$\begin{bmatrix} x_\iota I_{f_i} & * \\ 0 & 0 \end{bmatrix},$$

where $*$ indicates irrelevant entries. Hence $\mathrm{trace}(D_i) = x_\iota f_i$. On the other hand, since $\mathrm{trace}(q) = 0$ for $q \neq \iota$, we have

$$\mathrm{trace}(D_i) = \mathrm{trace}\Big(\sum_{h,v} \epsilon(v) hv\Big) = \sum_{h,v} \epsilon(v)\mathrm{trace}(hv) = \mathrm{trace}\big(I_{\mathbb{F}S_n}\big) = n!.$$

Now we have $x_\iota f_i = n!$, so $c_i = x_\iota = n!/f_i$. □

Definition 1.26. Let $T_1^\lambda, \ldots, T_{n!}^\lambda$ be all the tableaux of shape $\lambda \vdash n$. We define

$$E_i^\lambda = \frac{f_i}{n!} D_i^\lambda \qquad (i = 1, \ldots, n!).$$

Corollary 1.27. *Every E_i^λ is an idempotent: $(E_i^\lambda)^2 = E_i^\lambda$.*

1.6 Two-sided ideals in the group algebra. The results in this subsection lead us towards an explicit description of the isomorphism ψ in the Wedderburn decomposition (1).

Definition 1.28. If T_i, T_j are tableaux of shape $\lambda \vdash n$ then we define

$$\xi_{ij} = \begin{cases} \epsilon(v) & \text{if } s_{ji} = vh \text{ for } h \in G_H(T_i) \text{ and } v \in G_V(T_i) \\ 0 & \text{otherwise} \end{cases}$$

Lemma 1.29. *If T_i, T_j are tableaux of shape $\lambda \vdash n$ then $E_i E_j = \xi_{ij} E_i s_{ij}$.*

Proof: First, assume that $s_{ji} = vh$ for some $h \in G_H(T_i)$, $v \in G_V(T_i)$. Proposition 1.22, equation (5) and von Neumann's Theorem imply

$$\begin{aligned} E_i E_j &= E_i(s_{ji} E_i s_{ij}) = \frac{1}{c_i^2} H_i(V_i v)(h H_i) V_i s_{ij} = \frac{1}{c_i^2} \epsilon(v) H_i V_i H_i V_i s_{ij} \\ &= \epsilon(v) E_i^2 s_{ij} = \epsilon(v) E_i s_{ij}. \end{aligned}$$

Second, assume that $s_{ji} \neq vh$ for any $h \in G_H(T_i)$, $v \in G_V(T_i)$. Since $T_j = s_{ji} T_i$, Lemma 1.16 shows that there are numbers k, ℓ in the same column of T_i and the same row of T_j. Using the transposition $t = (k, \ell) \in G_V(T_i) \cap G_H(T_j)$ we obtain

$$E_i E_j = \frac{1}{c_i^2} H_i(V_i t)(t H_j) V_j = -\frac{1}{c_i^2} H_i V_i H_j V_j = -E_i E_j.$$

Hence $E_i E_j = 0$. □

Remark 1.30. From now on we will work only with standard tableaux.

Definition 1.31. Given a partition $\lambda \vdash n$ with standard tableaux $T_1, \ldots, T_{d_\lambda}$ in lex order, we write $\mathcal{E}^\lambda$ for the $d_\lambda \times d_\lambda$ matrix with (i,j) entry ξ_{ij}.

Lemma 1.32. *We have $\mathcal{E}^\lambda = I^\lambda + \mathcal{F}^\lambda$ where I^λ is the identity matrix and $\mathcal{F}^\lambda$ is a strictly upper triangular matrix. In particular, $\mathcal{E}^\lambda$ is invertible.*

Proof: If $i > j$ then Corollary 1.24 implies that $E_i E_j = 0$ and so $\xi_{ij} = 0$. If $i = j$ then $s_{ii} = \iota$ and so Lemma 1.29 gives $E_i = \xi_{ii} E_i$, hence $\xi_{ii} = 1$. □

Proposition 1.33. *If $\lambda \vdash n$ and T_i, T_j, T_k, T_ℓ are standard tableaux of shape λ then*

$$(E_i s_{ij})(E_k s_{k\ell}) = \xi_{jk} E_i s_{i\ell}.$$

Proof: Using Proposition 1.22, equation (6), and Lemma 1.29, we obtain

$$E_i s_{ij} E_k s_{k\ell} = s_{ij} E_j E_k s_{k\ell} = \xi_{jk} s_{ij} E_j s_{jk} s_{k\ell} = \xi_{jk} E_i s_{ij} s_{jk} s_{k\ell} = \xi_{jk} E_i s_{i\ell},$$

as required. □

Remark 1.34. If we replace the scalar ξ_{jk} in Proposition 1.33 by the Kronecker delta δ_{jk}, and write $E_{ij} = E_i s_{ij}$, then we obtain the matrix unit relations $E_{ij} E_{k\ell} = \delta_{jk} E_{i\ell}$. In order to construct the isomorphism ψ, we need to modify the elements $E_i s_{ij}$ to produce other elements which exactly satisfy the matrix unit relations.

Definition 1.35. We write N^λ for the subspace spanned by the $E_i^\lambda s_{ij}^\lambda$:

$$N^\lambda = \operatorname{span}\{E_i^\lambda s_{ij}^\lambda \mid 1 \leq i, j \leq d_\lambda\} \subset \mathbb{F} S_n.$$

We write N for the sum of the subspaces N^λ over all $\lambda \vdash n$.

Corollary 1.36. *For each $\lambda \vdash n$, the subspace N^λ is a subalgebra of $\mathbb{F} S_n$.*

We fix a partition $\lambda \vdash n$ with standard tableaux $T_1, \dots, T_{d_\lambda}$ in lex order. Let $A = (a_{ij})$ be any $d_\lambda \times d_\lambda$ matrix over $\mathbb{F}$, and consider the group algebra element

$$\alpha^\lambda(A) = \sum_{i=1}^{d_\lambda} \sum_{j=1}^{d_\lambda} a_{ij} E_i s_{ij}. \tag{8}$$

As usual, we write E_{ij} for the $d_\lambda \times d_\lambda$ matrix with 1 in position (i,j) and 0 elsewhere.

Lemma 1.37. *For all partitions $\lambda \vdash n$ and all $i, j, k, \ell \in \{1, \dots, d_\lambda\}$ we have*

$$\alpha^\lambda(E_{ij}) \alpha^\lambda(E_{k\ell}) = \alpha^\lambda(E_{ij} \mathcal{E}^\lambda E_{k\ell}).$$

Proof: We have $\alpha^\lambda(E_{ij}) \alpha^\lambda(E_{k\ell}) = E_i s_{ij} E_k s_{k\ell} = \xi_{jk} E_i s_{i\ell} = \alpha^\lambda(E_{ij} \mathcal{E}^\lambda E_{k\ell})$, using Proposition 1.33. □

Proposition 1.38. *The set $\{E_i^\mu s_{ij}^\mu \mid \mu \vdash n, 1 \leq i, j \leq d_\mu\}$ is linearly independent.*

Proof: A linear dependence relation among the $E_i^\mu s_{ij}^\mu$ can be written as

$$\sum_{\mu \vdash n} \alpha^\mu(A^\mu) = 0.$$

We fix a partition λ, and obtain

$$\alpha^\lambda\big(E_{ii}(\mathcal{E}^\lambda)^{-1}\big)\Big[\sum_{\mu\vdash n}\alpha^\mu(A^\mu)\Big]\alpha^\lambda\big((\mathcal{E}^\lambda)^{-1}E_{jj}\big)=0.$$

Using equation (6), Definition 1.26, and Proposition 1.23, we see that all terms vanish except for $\mu=\lambda$:

$$\alpha^\lambda\big(E_{ii}(\mathcal{E}^\lambda)^{-1}\big)\alpha^\lambda(A^\lambda)\alpha^\lambda\big((\mathcal{E}^\lambda)^{-1}E_{jj}\big)=0.$$

Lemma 1.37 gives $\alpha^\lambda(E_{ii}(\mathcal{E}^\lambda)^{-1}\mathcal{E}^\lambda A^\lambda\mathcal{E}^\lambda(\mathcal{E}^\lambda)^{-1}E_{jj})=0$, hence $\alpha^\lambda(E_{ii}A^\lambda E_{jj})=0$ and $\alpha^\lambda(a^\lambda_{ij}E_{ij})=0$, and so $a^\lambda_{ij}E_i s_{ij}=0$. Thus $a^\lambda_{ij}=0$ for all λ and all i,j. $\square$

Definition 1.39. Suppose that n has r distinct partitions $\lambda_1,\dots,\lambda_r$ in lex order. For $i=1,\dots,r$ let $d_i=d_{\lambda_i}$ be the number of standard tableaux of shape λ_i. Consider the direct sum of full matrix algebras

$$M=\bigoplus_{i=1}^r M_{d_i}(\mathbb{F}).$$

The linear map $\alpha\colon M\to\mathbb{F}S_n$ is the direct sum of the $\alpha^i=\alpha^{\lambda_i}$ from equation (8):

$$\alpha(A_1,\dots,A_r)=\alpha^1(A_1)+\cdots+\alpha^r(A_r).$$

Corollary 1.40. *The map α is injective. For every $\lambda\vdash n$ and $1\le i,j\le d_\lambda$, we have* $\dim N^\lambda=d_\lambda^2$. *The sum N of the N^λ is direct, and hence* $\dim N=\sum_\lambda d_\lambda^2$.

Proof: Injectivity of α is equivalent to the linear independence in Proposition 1.38. Since linear independence holds for each λ, the spanning set for N^λ is also a basis. The sum of the N^λ is direct by Proposition 1.23. $\square$

Remark. Since $N\subseteq\mathbb{F}S_n$, it follows that $\sum_\lambda d_\lambda^2\le n!$, so to prove $N=\mathbb{F}S_n$, it remains to show equality. Algorithms for insertion or deletion of a number to or from a standard tableau provide a bijection between S_n and the set of ordered pairs of standard tableaux of the same shape [35, §5.1.4, Theorem A].

1.7 Matrix units in the group algebra. We prove that the map ψ in (1) is an isomorphism by constructing elements of $\mathbb{F}S_n$ corresponding to matrix units.

Remark 1.41. The linear map $\alpha^\lambda\colon M_{d_\lambda}(\mathbb{F})\to\mathbb{F}S_n$ is not in general an algebra homomorphism. However, we can easily obtain an algebra homomorphism from it.

Definition 1.42. For all $\lambda\vdash n$ and $1\le i,j\le d_\lambda$, we define the following elements:

$$U^\lambda_{ij}=\alpha^\lambda\big(E^\lambda_{ij}(\mathcal{E}^\lambda)^{-1}\big)\in\mathbb{F}S_n.$$

Proposition 1.43. *For all $\lambda, \mu \vdash n$, $1 \le i, j \le d_\lambda$, $1 \le k, \ell \le d_\mu$ we have*

$$U^\lambda_{ij} U^\mu_{k\ell} = \delta_{\lambda\mu} \delta_{jk} U^\lambda_{i\ell}.$$

PROOF: If $\lambda = \mu$ then

$$\begin{aligned} U_{ij} U_{k\ell} &= \alpha(E_{ij}\mathcal{E}^{-1})\alpha(E_{k\ell}\mathcal{E}^{-1}) = \alpha(E_{ij}\mathcal{E}^{-1}\mathcal{E}E_{k\ell}\mathcal{E}^{-1}) = \alpha(E_{ij}E_{k\ell}\mathcal{E}^{-1}) \\ &= \alpha(\delta_{jk}E_{i\ell}\mathcal{E}^{-1}) = \delta_{jk}\alpha(E_{i\ell}\mathcal{E}^{-1}) = \delta_{jk}U_{i\ell}. \end{aligned}$$

The factor $\delta_{\lambda\mu}$ comes from the orthogonality of Proposition 1.23. □

Definition 1.44. We define the linear map $\psi\colon M \to \mathbb{F}S_n$ on matrix units as

$$\psi(E^\lambda_{ij}) = U^\lambda_{ij} \qquad (\lambda \vdash n;\ 1 \le i, j \le d_\lambda).$$

Theorem 1.45. *The map $\psi\colon M \to \mathbb{F}S_n$ is an isomorphism of associative algebras. In particular, $M_{d_i}(\mathbb{F})$ is isomorphic to N^{λ_i}.*

PROOF: This is an immediate corollary of the preceding results. □

Remark 1.46. Since the direct sum M of full matrix algebras is clearly semisimple, and simplicity is preserved by isomorphism, it follows that $\mathbb{F}S_n$ is semisimple, and moreover that it splits over $\mathbb{F}$: the structure theory of semisimple associative algebras implies that $\mathbb{F}S_n$ is isomorphic to the direct sum of simple two-sided ideals, and that each simple ideal is isomorphic to the endomorphism algebra of a vector space over a division ring $\mathbb{D}$ over $\mathbb{F}$. But our results show that $\mathbb{D} = \mathbb{F}$ for every λ. Since the scalar factors $d_\lambda/n!$ in Definition 1.26 are defined in characteristic $> n$, we also obtain the semisimplicity of $\mathbb{F}S_n$ in this case.

Example 1.47. For $n = 3$ we take the permutations 123, 132, 213, 231, 312, 321 in lex order — writing p as $p(1)p(2)p(3)$ — as our basis of $\mathbb{F}S_3$. The partitions $\lambda = 3$, $\mu = 21$, $\nu = 111$ have the following standard tableaux:

$$T^\lambda_1 = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline\end{array} \qquad T^\mu_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \qquad T^\mu_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \qquad T^\nu_1 = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array}$$

Thus $d_\lambda = 1$, $d_\mu = 2$, $d_\nu = 1$ and hence we have the isomorphism

$$\psi\colon M = \mathbb{F} \oplus M_2(\mathbb{F}) \oplus \mathbb{F} \longrightarrow \mathbb{F}S_3.$$

As ordered basis of M we take the matrix units E^λ_{11}, E^μ_{11}, E^μ_{12}, E^μ_{21}, E^μ_{22}, E^ν_{11}. We will compute the corresponding elements U^ρ_{ij} of $\mathbb{F}S_n$. The groups of horizontal and vertical permutations are as follows:

$$\begin{aligned} &G_H(T^\lambda_1) = S_3,\ \ G_V(T^\lambda_1) = \{123\},\ \ G_H(T^\mu_1) = \{123, 213\},\ \ G_V(T^\mu_1) = \{123, 321\}, \\ &G_H(T^\mu_2) = \{123, 321\},\ \ G_V(T^\mu_2) = \{123, 213\},\ \ G_H(T^\nu_1) = \{123\},\ \ G_V(T^\nu_1) = S_3. \end{aligned}$$

The symmetric and alternating sums over these subgroups are as follows:

$$H_{T_1^\lambda} = 123+132+213+231+312+321, \quad V_{T_1^\lambda} = 123,$$
$$H_{T_1^\mu} = 123+213, \quad V_{T_1^\mu} = 123-321, \quad H_{T_2^\mu} = 123+321, \quad V_{T_2^\mu} = 123-213,$$
$$H_{T_1^\nu} = 123, \quad V_{T_1^\nu} = 123-132-213+231+312-321.$$

The products D_{ij}^ρ are easily calculated; and scaling gives the idempotents:

$$E_1^\lambda = \tfrac{1}{6}(123+132+213+231+312+321), \qquad E_1^\mu = \tfrac{1}{3}(123+213-312-321),$$
$$E_2^\mu = \tfrac{1}{3}(123-213-231+321), \qquad E_1^\nu = \tfrac{1}{6}(123-132-213+231+312-321).$$

Clearly $s_{12}^\mu = s_{21}^\mu = 132$, and this is the only nontrivial case. Hence $s_{12} \neq vh$ for any $v \in G_V(T_2^\mu)$, $h \in G_H(T_2^\mu)$ (Lemma 1.29), and so every $\mathcal{E}^\rho$ is the identity matrix of size d_ρ. Therefore every $U_{ij}^\rho = \alpha^\rho(E_{ij}) = E_i^\rho s_{ij}^\rho$, which gives the following matrix units in the group algebra:

$$U_{11}^\lambda = E_1^\lambda, \qquad U_{11}^\mu = E_1^\mu, \qquad U_{12}^\mu = E_1^\mu s_{12} = \tfrac{1}{3}(132+231-312-321),$$
$$U_{21}^\mu = E_2^\mu s_{21} = \tfrac{1}{3}(132-213-231+312), \qquad U_{22}^\mu = E_2^\mu, \qquad U_{11}^\nu = E_1^\nu.$$

These equations can be summarized by the matrix representing ψ with respect to our ordered bases of M and $\mathbb{F}S_n$, and then we obtain the matrix representing ψ^{-1}:

$$(9) \qquad \psi \sim \frac{1}{6}\begin{bmatrix} 1 & 2 & 0 & 0 & 2 & 1 \\ 1 & 0 & 2 & 2 & 0 & -1 \\ 1 & 2 & 0 & -2 & -2 & -1 \\ 1 & 0 & 2 & -2 & -2 & 1 \\ 1 & -2 & -2 & 2 & 0 & 1 \\ 1 & -2 & -2 & 0 & 2 & -1 \end{bmatrix} \qquad \psi^{-1} \sim \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & -1 & 0 & -1 \\ 0 & 1 & -1 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 1 & 0 & -1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

For any $X \in \mathbb{F}S_3$, we have

$$\psi^{-1}(X) = x_1 E_{11}^\lambda + x_2 E_{11}^\mu + x_3 E_{12}^\mu + x_4 E_{21}^\mu + x_5 E_{22}^\mu + x_6 E_{11}^\nu = \left[x_1, \begin{bmatrix} x_2 & x_3 \\ x_4 & x_5 \end{bmatrix}, x_6 \right]$$

Therefore

$$(10) \quad \left\{ \begin{array}{ll} \psi^{-1}(123) = \left[1, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, 1\right] & \psi^{-1}(132) = \left[1, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, -1\right] \\ \psi^{-1}(213) = \left[1, \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, -1\right] & \psi^{-1}(231) = \left[1, \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}, 1\right] \\ \psi^{-1}(312) = \left[1, \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, 1\right] & \psi^{-1}(321) = \left[1, \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix}, -1\right] \end{array} \right.$$

These are the representation matrices for the irreducible representations of S_3.

1.8 Clifton's theorem on representation matrices. Our next goal is to compute explicitly the algebra homomorphism ϕ:

$$\phi\colon \mathbb{F}S_n \longrightarrow \bigoplus_{\lambda\vdash n} M_{d_\lambda}(\mathbb{F}). \tag{11}$$

We fix $\lambda \vdash n$ throughout the following discussion, and consider all the tableaux $T_1, \dots, T_{n!}$ of shape λ. Recall that for $1 \le i, j \le n!$ we define $s_{ij} \in S_n$ by the equation $s_{ij} T_j = T_i$. If $p \in S_n$ then $pT_j = T_r$ for some r, and so $p = s_{rj}$. As before, we write E_i for the idempotent corresponding to T_i. Proposition 1.22 and Lemma 1.29 show that $E_i E_j = \xi_{ij} E_i s_{ij} = \xi_{ij} s_{ij} E_j$. Therefore

$$E_i p E_j = E_i s_{rj} E_j = E_i E_r s_{rj} = \xi_{ir} s_{ir} E_r s_{rj} = \xi_{ir} s_{ir} s_{rj} E_j = \xi_{ir} s_{ij} E_j.$$

We define $\xi^p_{ij} = \xi_{ir}$ when $p = s_{rj}$, so that for all i, j, p we have

$$E_i p E_j = \xi^p_{ij} s_{ij} E_j. \tag{12}$$

We now restrict to the d_λ standard tableaux $T_1, \dots, T_{d_\lambda}$ in lex order.

Definition 1.48. For all $p \in S_n$ the **Clifton matrix** A^λ_p is defined by

$$(A^\lambda_p)_{ij} = \xi^p_{ij} \qquad (1 \le i, j \le d_\lambda).$$

The matrix previously denoted $\mathcal{E}^\lambda$ is the Clifton matrix A^λ_ι for $\iota \in S_n$.

Referring to the definition of ξ_{ij} in Lemma 1.29, we see that A^λ_p can be computed by the following steps, presented formally in Figure 3 [3], [13], [15].

- Apply p to the standard tableau T_j obtaining the (possibly nonstandard) tableau pT_j.
- If there exist two numbers that appear together both in a column of T_i and in a row of pT_j, then $(A^\lambda_p)_{ij} = 0$.
- Otherwise, there exists a vertical permutation $q \in G_V(T_i)$ which takes the numbers of T_i into the rows they occupy in pT_j. Then $(A^\lambda_p)_{ij} = \epsilon(q)$.

Figure 3 attempts to find q, and returns 0 if no such permutation exists.

Before proving Clifton's theorem, it is worth quoting in its entirety the review by G.D. James[1] of Clifton's paper [15]: "From his natural representation of the symmetric groups, A. Young produced representations known as the orthogonal form and the seminormal form and gave a straightforward method of calculating the matrices representing permutations. A disadvantage of these representations is that the matrix entries are not in general integers, and for many practical purposes, the natural representation is preferable. Most methods for working out the matrices for the natural representation are messy, but this paper gives an approach which is simple both to prove and to apply. Let $T_1, T_2, \dots, T_f$ be the standard tableaux. For each $\pi \in S_n$, form the $f \times f$ matrix A_π whose i, j entry

[1] MathSciNet: MR0624907

Input: A partition $\lambda = (n_1, \dots, n_\ell) \vdash n$ and a permutation $p \in S_n$.
Output: The Clifton matrix A_p^λ. For $j = 1, \dots, d_\lambda$ do:

(1) Compute pT_j.
(2) For $i = 1, \dots, d_\lambda$ do:
  (a) Set $e \leftarrow 1$, $k \leftarrow 1$, $\beta \leftarrow$ `false`.
  (b) While $k \leq n$ and not β do:
    (i) Set $r_i, c_i \leftarrow$ row, column indices of k in T_i.
    (ii) Set $r_j, c_j \leftarrow$ row, column indices of k in pT_j.
    (iii) If $r_i \neq r_j$ then [*k is not in the correct row*]
      - if $c_i > n_{r_j}$ then set $e \leftarrow 0$, $\beta \leftarrow$ `true` [*required position does not exist*]
      - else if $T_i(r_j, c_i) < T_i(r_i, c_i)$ then set $e \leftarrow 0$, $\beta \leftarrow$ `true` [*required position is already occupied*]
      - else set $e \leftarrow -e$, interchange $T_i(r_i, c_i) \leftrightarrow T_i(r_j, c_i)$ [*transpose k into the required position*]
    (iv) Set $k \leftarrow k + 1$
  (c) Set $(A_p^\lambda)_{ij} \leftarrow e$

FIGURE 3. Algorithm to compute the Clifton matrix A_p^λ

is given by the following rule. If two numbers lie in the same row of πT_j and in the same column of T_i, then the i, j entry in A_π is zero. Otherwise, the i, j entry equals the sign of the column permutation for T_i which takes the numbers of T_i to the correct rows they occupy in πT_j. The matrix representing π in the natural representation is then $A_I^{-1} A_\pi$, where I is the identity permutation of S_n."

The Wedderburn decomposition of $\mathbb{F}S_n$ shows that every permutation $p \in S_n$ is a sum of terms $p^\lambda \in \mathbb{F}S_n$ for $\lambda \vdash n$, and each p^λ is a linear combination of the U_{ij}^λ:

$$p = \sum_\lambda \left(\sum_{i=1}^{d_\lambda} \sum_{j=1}^{d_\lambda} r_{ij}^\lambda(p) U_{ij}^\lambda \right). \tag{13}$$

Definition 1.49. We define $R^\lambda(p)$ to be the $d_\lambda \times d_\lambda$ matrix with (i, j) entry $r_{ij}^\lambda(p)$. We call $R^\lambda(p)$ the **representation matrix** of $p \in S_n$ for $\lambda \vdash n$.

Lemma 1.50. *We have $U_{ii}^\lambda \, p \, U_{jj}^\lambda = r_{ij}^\lambda(p) U_{ij}^\lambda$.*

Proposition 1.51 (Clifton's Theorem)**.** *For all $\lambda \vdash n$ and $p \in S_n$ we have*

$$R^\lambda(p) = (A_\iota^\lambda)^{-1} A_p^\lambda.$$

PROOF: We write $\mathcal{E} = A_\iota^\lambda$ and denote the entries of $\mathcal{E}^{-1}$ by η_{ij}. We have

$$
\begin{aligned}
U_{ii}^\lambda \, p \, U_{jj}^\lambda &= \alpha(E_{ii}\mathcal{E}^{-1}) \, p \, \alpha(E_{jj}\mathcal{E}^{-1}) = \left(\sum_{k=1}^{d_\lambda} \eta_{ik} E_i s_{ik} \right) p \left(\sum_{\ell=1}^{d_\lambda} \eta_{j\ell} E_j s_{j\ell} \right) \\
&= \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik} \, \eta_{j\ell} \, E_i \, s_{ik} \, p \, E_j \, s_{j\ell} = \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik} \, \eta_{j\ell} \, s_{ik} \, E_k \, p \, E_j \, s_{j\ell} \\
&\overset{(12)}{=} \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik} \, \eta_{j\ell} \, s_{ik} \, \xi_{kj}^p \, s_{kj} \, E_j \, s_{j\ell} = \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik} \, \eta_{j\ell} \, \xi_{kj}^p \, s_{ik} \, s_{kj} \, E_j \, s_{j\ell} \\
&= \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik} \, \eta_{j\ell} \, \xi_{kj}^p \, s_{ik} \, E_k \, s_{kj} \, s_{j\ell} = \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik} \, \eta_{j\ell} \, \xi_{kj}^p \, E_i \, s_{ik} \, s_{kj} \, s_{j\ell} \\
&= \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik} \, \eta_{j\ell} \, \xi_{kj}^p \, E_i \, s_{i\ell} = \left(\sum_{k=1}^{d_\lambda} \eta_{ik} \, \xi_{kj}^p \right) \left(\sum_{\ell=1}^{d_\lambda} \eta_{j\ell} \, E_i \, s_{i\ell} \right) \\
&= \left(\sum_{k=1}^{d_\lambda} \eta_{ik} \, \xi_{kj}^p \right) U_{ij} = (A_\iota^{-1} A_p)_{ij} \, U_{ij}.
\end{aligned}
$$

Therefore $r_{ij}^\lambda(p) = (A_\iota^{-1} A_p)_{ij}$ for all i, j and so $R^\lambda(p) = A_\iota^{-1} A_p$ as required. $\square$

Example 1.52. For $n = 3$ we have $A_\iota^\lambda = I_{d_\lambda}$ for all $\lambda \vdash 3$, so $R_p^\lambda = A_p^\lambda$ for all $p \in S_3$. Consider $\lambda = 21$ with $d_\lambda = 2$, and $p = 213$. For $i, j = 1, 2$ we write the tableaux T_i and pT_j, and the vertical permutation q (when it exists):

$$
\begin{array}{lllll}
(i,j) = (1,1) & T_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} & pT_1 = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} & q = \iota & \epsilon(q) = 1 \\[2ex]
(i,j) = (1,2) & T_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} & pT_2 = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} & q = 321 & \epsilon(q) = -1 \\[2ex]
(i,j) = (2,1) & T_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} & pT_1 = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} & q \text{ does not exist} & \\[2ex]
(i,j) = (2,2) & T_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} & pT_2 = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} & q = 213 & \epsilon(q) = -1
\end{array}
$$

We obtain $A_p^\lambda = \left[\begin{smallmatrix} 1 & -1 \\ 0 & -1 \end{smallmatrix}\right]$ which agrees with $(\psi^\lambda)^{-1}(213)$ from Example 1.47.

Example 1.53. Consider $n = 5$, the smallest n for which there exists $\lambda \vdash n$ such that $A_\iota^\lambda \neq I_{d_\lambda}$. We list the standard tableaux for $\lambda = 32$ in lex order:

$$
T_1, \, \ldots, \, T_5 \quad = \quad
\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2}\end{array} \qquad
\begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 5 \\ \cline{1-2}\end{array} \qquad
\begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array} \qquad
\begin{array}{|c|c|c|}\hline 1 & 3 & 4 \\ \hline 2 & 5 \\ \cline{1-2}\end{array} \qquad
\begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 \\ \cline{1-2}\end{array}
$$

Let $p = \iota$ and consider the $(i, j) = (1, 5)$ entry of $\mathcal{E} = A_\iota^\lambda$; we have $T_i = T_1$ and $pT_j = T_5$. The required vertical permutation is the transposition $q = 15342$

interchanging 2 and 5, so $(A_\iota^\lambda)_{15} = -1$. Similar calculations show that

$$A_\iota^\lambda = I_5 - E_{15}, \qquad (A_\iota^\lambda)^{-1} = I_5 + E_{15}.$$

To illustrate the difference between the Clifton matrix A_p^λ and the representation matrix $R_p^\lambda = (A_\iota^\lambda)^{-1} A_p^\lambda$, consider the 5-cycle $p = 23451$; in this case we obtain

$$A_p^\lambda = \begin{bmatrix} -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 \end{bmatrix} \qquad R_p^\lambda = \begin{bmatrix} -1 & -1 & 1 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 & 0 \end{bmatrix}$$

2. Computational methods for studying polynomial identities

2.1 Preliminary definitions.

Definition 2.1. Let A be an algebra, not necessarily associative, of finite dimension d over $\mathbb{F}$. The multiplication in A is a bilinear map $m\colon A \times A \to A$ denoted $(x, y) \mapsto xy$. If $v_1, \dots, v_d$ is an ordered basis of A then the multiplication can be defined in terms of **structure constants** c_{ij}^k with respect to this basis:

$$v_i v_j = \sum_{k=1}^{d} c_{ij}^k v_k. \tag{14}$$

Definition 2.2 ([53])**.** Let $X = \{x_1, x_2, \dots\}$ be a finite or countably infinite set of variables. We construct the set $M(X)$ of **nonassociative monomials** inductively: $X \subset M(X)$; if $x, y \in X$ and $v, w \in M(X) \backslash X$ then xy, $x(v)$, $(v)x$, $(v)(w) \in M(X)$. With this composition $M(X)$ is the **free magma** generated by X. We write $M(X)_n$ for the subset consisting of monomials of degree n.

Definition 2.3. The monomials in $M(X)$ include various placements of parentheses called **association types**, which have a natural total order, defined inductively, based on the unique factorization $m = m_1 m_2$ of nonassociative monomials. If we fix a symbol $*$ then the association type of a monomial is obtained by replacing every variable by $*$. If $m = m_1 m_2$ and $m' = m_1' m_2'$ are monomials of the same degree, then $m \prec m'$ if and only if either (i) $m_1 \prec m_1'$ or (ii) $m_1 = m_1'$ and $m_2 \prec m_2'$. We assume that lower degrees precede higher degrees.

Example 2.4. For $n = 3$ we have two association types: $(**)*$ and $*(**)$. For $n = 4$ we have five: $((**)*)*$, $(*(**))*$, $(**)(**)$, $*((**)*)$, $*(*(**))$.

Lemma 2.5. *The number of association types of degree n equals the number of rooted binary plane trees with n leaves, which is the (shifted) Catalan number:*

$$C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}.$$

Example 2.6. The numbers C_n grow very rapidly:

n	1	2	3	4	5	6	7	8	9	10	11	12
C_n	1	1	2	5	14	42	132	429	1430	4862	16796	58786

Definition 2.7. We write $\mathbb{F}\{X\}$ for the vector space over $\mathbb{F}$ with basis $M(X)$. The **free nonassociative algebra** generated by X over $\mathbb{F}$ equals $\mathbb{F}\{X\}$ with multiplication extended bilinearly from $M(X)$. The elements of $\mathbb{F}\{X\}$ are **nonassociative polynomials** in the variables X with coefficients in $\mathbb{F}$. The **homogeneous component** $\mathbb{F}\{X\}_n$ is the subspace whose monomial basis is the ordered set $M(X)_n$. A **T-ideal** in $\mathbb{F}\{X\}$ is an ideal $\mathcal{R} \subseteq \mathbb{F}\{X\}$ such that $f(\mathcal{R}) \subseteq \mathcal{R}$ for any algebra endomorphism f of $\mathbb{F}\{X\}$.

Definition 2.8. A **polynomial identity** satisfied by the algebra A is an equation $I \equiv 0$ where $I \in \mathbb{F}\{X\}$ is a nonassociative polynomial which vanishes when arbitrary elements of A are substituted for the variables in X. The symbol $\equiv$ indicates that the identity holds for all substitutions. If $I \in \mathbb{F}\{X\}_n$ then we say I is **homogeneous** of degree n. If $I \in \mathbb{F}\{X\}_n$ and $x_1, \dots, x_n$ each occur exactly once in every monomial of I then we say I is **multilinear** of degree n. We denote by $T_X(A)$ the set of all polynomial identities in the variables X satisfied by the algebra A.

Lemma 2.9 ([53]). *The set of polynomial identities $T_X(A) \subseteq \mathbb{F}\{X\}$ satisfied by the algebra A does not depend on the basis of A, and $T_X(A)$ is a T-ideal in $\mathbb{F}\{X\}$.*

2.1.1 Historical remarks. Algebras which satisfy polynomial identities (also known as PI-algebras) constitute an important class of algebras, and therefore research on polynomial identities is of great interest to algebraists [36]. Investigation of this area was initiated in 1922 by Dehn [16] who was motivated by problems in geometry. Wagner [51] in 1937 found identities for the quaternions and for matrix algebras. However, the vigorous development of the theory of PI-algebras began with the works of Jacobson [29] and Kaplansky [31] in the late 1940's. In particular, the following is a classical problem for PI-algebras.

Problem 2.10 (Specht [48]). For a given variety of algebras, determine whether every algebra in this variety has a finite basis of identities, in the sense that its T-ideal $T_X(A) \subseteq \mathbb{F}\{X\}$ is finitely generated.

Specht originally posed this problem for associative algebras over fields of characteristic 0. The complete solution was found by Kemer.

Theorem 2.11 (Kemer [32], [33]). *Every associative algebra over a field of characteristic 0 has a finite basis of identities.*

Similar results were obtained by Vais & Zelmanov [49] for finitely generated Jordan algebras, and by Iltyakov [26], [27] for f.g. alternative and Lie algebras.

2.1.2 Minimal identities. We can study $T_X(A)_n = T_X(A) \cap \mathbb{F}\{X\}_n$, the homogeneous component of degree n of the T-ideal. The nonzero elements of $T_X(A)_n$ are the polynomial identities of degree n for A. An important problem is to find the smallest n for which $T(A)_n \neq 0$; in this case, the nonzero elements of $T_X(A)_n$ are called **minimal identities** for A. For the simple matrix algebras $M_k(\mathbb{F})$, the minimal identities were found by Amitsur and Levitzki.

Theorem 2.12 ([1]). *Apart from associativity, the minimal degree of a polynomial identity of $M_n(\mathbb{F})$ is $2n$. Every multilinear polynomial identity of degree $2n$ for $M_n(\mathbb{F})$ is a scalar multiple of the standard polynomial:*

$$s_{2n}(x_1, \ldots, x_{2n}) = \sum_{\sigma \in S_{2n}} \epsilon(\sigma) x_{\sigma(1)} \cdots x_{\sigma(2n)}.$$

Leron [37] proved that if $\mathrm{char}(\mathbb{F}) = 0$ and $n > 2$ then every polynomial identity of degree $2n{+}1$ for $M_n(\mathbb{F})$ is a consequence of s_{2n}. In particular, the identities of degree 7 for $M_3(\mathbb{F})$ are consequences of s_6. Drensky & Kasparian [19] found all identities of degree 8 for $M_3(\mathbb{F})$ when $\mathrm{char}(\mathbb{F}) = 0$, and showed that they are consequences of s_6; see also Bondari [5], [6]. The T-ideal of identities for $M_2(\mathbb{F})$ has been studied by many authors [42]. Computational methods used to study polynomial identities of matrices are discussed by Benanti et al. [2].

Problem 2.13. Given an ordered basis and structure constants c_{ij}^k for a finite-dimensional algebra A over $\mathbb{F}$, determine the polynomial identities of degree $\leq n$ satisfied by A. In particular, find the minimal identities satisfied by A.

2.2 Multilinear polynomial identities.

Lemma 2.14 ([53]). *Over a field of characteristic 0, every polynomial identity is equivalent to a set of multilinear identities.*

Thus in characteristic 0, we may restrict our study to multilinear identities $I(x_1, \ldots, x_n) \equiv 0$ where each term of $I(x_1, \ldots, x_n)$ consists of a coefficient from $\mathbb{F}$, a permutation of $x_1, x_2, \ldots, x_n$, and an association type indicating the order in which the multiplications are performed.

If there are $t = t(n)$ association types in degree n then $I(x_1, \ldots, x_n)$ can be written as $I_1 + I_2 + \cdots + I_t$ where the terms in the i-th summand all have the i-th association type. In each summand, the monomials differ only in the permutation of the variables, and so we may identify each summand with an element of $\mathbb{F}S_n$. We may therefore regard multilinear identities $I(x_1, \ldots, x_n)$ as elements of $(\mathbb{F}S_n)^t$, the direct sum of t copies of $\mathbb{F}S_n$.

This approach to polynomial identities, using the representation theory of the symmetric group (that is, the structure of the group algebra $\mathbb{F}S_n$), was introduced in 1950 independently by Malcev [38] and Specht [48]. In the 1970's, this theory was developed further by Regev [43], [44], with a particular focus on associative PI algebras, and the computational implementation was initiated by Hentzel[2] [21], [22].

Example 2.15. If A is an associative algebra, then the placement of parentheses does not affect the product, and so we only need to choose one association type in each degree as the normal form. If necessary, we choose the right-normed product $x_1(x_2(\cdots(x_{n-1}x_n)\cdots))$, using the identity permutation of the variables. But usually we can omit the parentheses and write simply $x_1x_2\cdots x_{n-1}x_n$. In an associative algebra, any two multilinear monomials of degree n in n variables differ only by the permutation of the variables, and so a multilinear polynomial identity in degree n can be regarded as an element of the group algebra $\mathbb{F}S_n$.

Example 2.16. If A is commutative (such as a Jordan algebra) or anticommutative (such as a Lie algebra), then the association types are not independent; for example, $(ab)c = \pm c(ab)$. In these cases, the association types are enumerated by the Wedderburn-Etherington numbers $1, 1, 1, 2, 3, 6, 11, 23, 46, \ldots$ (OEIS A001190).

2.3 Fill and reduce algorithm. We explain the algorithm used to find the multilinear polynomial identities of degree n for an algebra A of dimension d over $\mathbb{F}$.

We choose a basis for A and express elements of A as vectors in $\mathbb{F}^d$. In degree n there are $t = t(n)$ association types and $n!$ permutations of the variables, for a total of $tn!$ distinct monomials; we fix once and for all a total order on these monomials. A polynomial identity $I(x_1, \ldots, x_n)$ is a linear combination of these $tn!$ monomials, with coefficients in $\mathbb{F}$. This method is only practical when the number $tn!$ of monomials is relatively small.

Let $E(n)$ be a matrix with $tn!$ columns and $tn!+d$ rows, consisting of a $tn!\times tn!$ upper block and a $d \times tn!$ lower block. We generate n pseudorandom elements $a_1, \ldots, a_n \in A$. We evaluate the $tn!$ monomials by setting $x_i = a_i$ $(i = 1, \ldots, n)$ and obtain a sequence r_j $(j = 1, \ldots, tn!)$ of elements of A. For each j we put the coefficient vector of r_j into the jth column of the lower block. The d rows of the lower block consist of linear constraints on the coefficients of the general multilinear polynomial identity $I(x_1, \ldots, x_n)$. We compute the row canonical form $\mathrm{RCF}(E(n))$, so the lower block becomes zero.

We repeat this process of generating pseudorandom elements of A, filling the lower block, and reducing the matrix until the rank of $E(n)$ stabilizes. At this point, we write a for the nullity; the nullspace consists of the coefficient vectors of a canonical set of generators for the multilinear polynomial identities satisfying the constraints imposed at each step, that is, the multilinear polynomial identities in degree n satisfied by A.

We compute the canonical basis of the nullspace by setting the free variables equal to the standard basis vectors and solving for the leading variables. We then

[2] Two of the present authors (MB, LP) learned about the application of this theory to polynomial identities through working with Hentzel.

put these canonical basis vectors into another matrix of size $a \times tn!$, and compute its RCF, which we denote by $[\mathrm{All}(n)]$. We call the row space of this matrix $\mathrm{All}(n)$; this is the vector space of all multilinear identities of degree n satisfied by A.

Example 2.17. We find the polynomial identities of degree 4 for $A = M_2(\mathbb{F})$, the 4-dimensional associative algebra of 2×2 matrices over $\mathbb{F}$. We construct a 28×24 zero matrix $E(4)$ and repeat the following process:

- generate pseudorandom 2×2 matrices a_1, a_2, a_3, a_4 over $\mathbb{F}$,
- evaluate $m^{(j)} = a_{p_j(1)} a_{p_j(2)} a_{p_j(3)} a_{p_j(4)}$ for all $p_j \in S_4 = \{p_1, \dots, p_{24}\}$,
- for $1 \leq j \leq 24$, store $m^{(j)}$ in the last 4 positions of column j of $E(4)$: $E(4)_{25,j} \leftarrow m^{(k)}_{11}$, $E(4)_{26,j} \leftarrow m^{(k)}_{12}$, $E(4)_{27,j} \leftarrow m^{(k)}_{21}$, $E(4)_{28,j} \leftarrow m^{(k)}_{22}$,
- compute the row canonical form $\mathrm{RCF}(E(4))$.

The first 6 iterations produce ranks 4, 8, 12, 16, 20, 23 and the rank remains 23 for the next 10 iterations. Hence the nullity is 1, and a basis for the nullspace consists of the coefficient vector of the standard identity of degree 4:

$$s_4(x_1, x_2, x_3, x_4) = \sum_{p \in S_4} \epsilon(p)\, x_{p(1)} x_{p(2)} x_{p(3)} x_{p(4)} \equiv 0. \tag{15}$$

This is the Amitsur-Levitzki identity from Theorem 2.12 in the case $n = 2$.

2.4 Consequences of polynomial identities. When computing the multilinear polynomial identities satisfied by an algebra A, we often find that many of the identities in degree n are consequences of known identities of lower degrees, so they do not provide any new information. We want the identities in degree n which cannot be expressed in terms of known identities of lower degrees.

Definition 2.18. Let $I(x_1, \dots, x_n)$ be a multilinear nonassociative polynomial of degree n. There are $n+2$ **consequences** of this polynomial in degree $n+1$, namely n substitutions obtained by replacing x_i by $x_i x_{n+1}$ $(i = 1, \dots, n)$ and two multiplications of I by x_{n+1} (on the right and the left):

$$\begin{array}{lll} I(x_1 x_{n+1}, \dots, x_n), \quad \dots & I(x_1, \dots, x_i x_{n+1}, \dots, x_n), \quad \dots & I(x_1, \dots, x_n x_{n+1}), \\ I(x_1, \dots, x_i, \dots, x_n) x_{n+1}, & x_{n+1} I(x_1, \dots, x_i, \dots, x_n). & \end{array}$$

If $I \equiv 0$ is a polynomial identity for an algebra A, then so are its consequences.

Lemma 2.19. *Every multilinear polynomial of degree $n+1$ in the T-ideal generated by $I(x_1, \dots, x_n) \in \mathbb{F}\{X\}$ is a linear combination of permutations of the $n+2$ consequences in Definition* 2.18.

PROOF: By definition, the T-ideal generated by I in $\mathbb{F}\{X\}$ is the ideal containing I which is invariant under all endomorphisms of $\mathbb{F}\{X\}$. The n substitutions correspond to invariance under endomorphisms, and the two multiplications correspond to invariance under right and left multiplication. $\square$

Example 2.20. The algebra $\mathbb{O}$ of octonions which we will study in detail later is an alternative algebra. These algebras are defined by the left and right alternative identities $(x,x,y) \equiv 0$ and $(x,y,y) \equiv 0$, where $(x,y,z) = (xy)z - x(yz)$ is the associator. Over a field of characteristic $\neq 2$, these two identities are equivalent to their linearized forms: $(x,z,y) + (z,x,y) \equiv 0$ and $(x,y,z) + (x,z,y) \equiv 0$. Each of these has five consequences in degree 4; the left alternative identity produces

$$(xw,z,y) + (z,xw,y) \equiv 0,\ (x,z,yw) + (z,x,yw) \equiv 0,\ (x,zw,y) + (zw,x,y) \equiv 0,$$
$$(x,z,y)w + (z,x,y)w \equiv 0,\ w(x,z,y) + w(z,x,y) \equiv 0.$$

2.4.1 Algebras, S_n-modules, operads. The vector space $\mathrm{All}(n)$ of all multilinear polynomial identities of degree n satisfied by an algebra A is a subspace of the multilinear space of degree n in the free nonassociative algebra $\mathbb{F}\{X\}$ where $X = \{x_1, \dots, x_n\}$. Since $\mathrm{All}(n)$ is invariant under permutations of the variables, we can regard $\mathrm{All}(n)$ as a left S_n-module with action given by permuting the subscripts of the variables: $\sigma \cdot f(x_1, \dots, x_n) = f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$. We can also consider $\mathrm{All}(n)$ as a submodule of $\mathrm{Bin}(n)$, the arity n component of the free symmetric operad Bin generated by one nonassociative binary operation with no symmetry [8]. The operad Bin is the symmetrization of a free nonsymmetric operad, but the operad governing the algebra A has the quotient module $\mathrm{Bin}(n)/\mathrm{All}(n)$ as its S_n-module in arity n, and hence may or may not be the symmetrization of a nonsymmetric operad, depending on the properties of the identities satisfied by A.

2.5 Module generators algorithm. For an algebra A, the consequences in degree n of the identities of degrees $< n$ generate a submodule $\mathrm{Old}(n) \subseteq \mathrm{All}(n)$. We explain the algorithm used to find a set of S_n-module generators for $\mathrm{Old}(n)$.

We assume by induction that we have already determined a set of S_{n-1}-module generators for $\mathrm{All}(n{-}1)$. The consequences of these generators in degree n form a set $O(n)$ of S_n-module generators for $\mathrm{Old}(n)$.

We construct a $(tn! + n!) \times tn!$ matrix $C(n)$ consisting of a $tn! \times tn!$ upper block and a $n! \times tn!$ lower block (as before, $t = t(n)$ is the number of association types in degree n). Using the lex order on permutations, we write σ_i for the i-th element of S_n. We take an identity $I \in O(n)$ and for $i = 1, \dots, n!$ we put the coefficient vector of $\sigma \cdot I$ into the i-th row of the lower block. The $n!$ rows of the lower block then contain all the permutations of I, and hence they span the S_n-module generated by I. We compute $\mathrm{RCF}(C(n))$ so the lower block becomes zero.

We repeat this process for each $I \in O(n)$. At the end, the nonzero rows of $\mathrm{RCF}(C(n))$ form a matrix $[\mathrm{Old}(n)]$ which contains the coefficient vectors of a canonical set of S_n-module generators for $\mathrm{Old}(n)$.

2.5.1 Existence of new identities. We compare the S_n-modules $\mathrm{Old}(n)$ and $\mathrm{All}(n)$ to determine whether there exist new multilinear identities in degree n satisfied by A; that is, identities which do not follow from those of degrees $< n$. To do this, we compare the reduced matrices $[\mathrm{Old}(n)]$ and $[\mathrm{All}(n)]$; we denote their ranks by r_{old} and r_{all}. If $r_{\mathrm{old}} = r_{\mathrm{all}}$ then we must have $[\mathrm{Old}(n)] = [\mathrm{All}(n)]$:

every identity in degree n satisfied by A follows from identities of lower degrees. If $r_{\mathrm{old}} \neq r_{\mathrm{all}}$ then since $\mathrm{Old}(n) \subseteq \mathrm{All}(n)$ we must have $r_{\mathrm{old}} < r_{\mathrm{all}}$, and the row space of $[\mathrm{Old}(n)]$ must be a subspace of the row space of $[\mathrm{All}(n)]$. The difference $r_{\mathrm{all}} - r_{\mathrm{old}}$ is the dimension of the S_n-module of new identities in degree n.

Definition 2.21. If X is a matrix in RCF, we write $\texttt{leading}(X)$ for the set of ordered pairs (i, j) such that X has a leading 1 in row i and column j. We write $\texttt{jleading}(X) = \{\, j \mid (i, j) \in \texttt{leading}(X)\}$.

Definition 2.22. The **new identities** satisfied by A in degree n are the nonzero elements of the quotient module $\mathrm{New}(n) = \mathrm{All}(n)/\mathrm{Old}(n)$. We find S_n-module generators for $\mathrm{New}(n)$, by calculating the set difference

$$\texttt{jleading}([\mathrm{All}(n)]) \setminus \texttt{jleading}([\mathrm{Old}(n)]) = \{j_1, \ldots, j_r\} \qquad (r = r_{\mathrm{all}} - r_{\mathrm{old}}).$$

Lemma 2.23. *For $s = 1, \ldots, r$ define i_s by $(i_s, j_s) \in \texttt{leading}([\mathrm{All}(n)])$. Rows $i_1, \ldots, i_r$ of $[\mathrm{All}(n)]$ are the coefficient vectors of the canonical generators of* $\mathrm{New}(n)$.

Example 2.24. To illustrate these concepts, we extend the results of Example 2.17 regarding 2×2 matrices from degree 4 to degree 5.

To find all the identities we proceed as before, with some obvious changes: the matrix $E(5)$ has size 124×120; each iteration generates 5 pseudorandom matrices; there are 120 permutations to evaluate. We find that the rank increases by 4 for each of the first 22 iterations, but the next iteration produces rank 91, and this remains constant for the next 10 iterations. Thus the nullspace of $E(5)$ has dimension 29; this is the S_5-module $\mathrm{All}(5)$, consisting of the coefficient vectors of all identities in degree 5 satisfied by 2×2 matrices.

To find which of these identities are new, we need to generate the consequences in degree 5 of the standard identity (15). Every consequence is a linear combination of permutations of these 6 generators: $s_4(x_1x_5, x_2, x_3, x_4)$, $s_4(x_1, x_2x_5, x_3, x_4)$, $s_4(x_1, x_2, x_3x_5, x_4)$, $s_4(x_1, x_2, x_3, x_4x_5)$, $x_5 s_4(x_1, x_2, x_3, x_4)$, $s_4(x_1, x_2, x_3, x_4)x_5$. By the skew-symmetry of s_4, we only need one of the first four generators. We construct a 240×120 zero matrix $C(5)$ and do the following for each generator.

- Set $i \leftarrow 120$.
- For each permutation $p \in S_5$ do:
  - Set $i \leftarrow i + 1$.
  - For each term cm in the generator, where $c = \pm 1$, $m = x_{q(1)} \cdots x_{q(5)}$, let j be the index of pq in the lex-ordering on S_5, and set $C(5)_{ij} \leftarrow c$.
- Compute the row canonical form $\mathrm{RCF}(C(5))$.

After all 6 generators have been processed, the rank of $C(5)$ is 24; its row space is the S_5-module $\mathrm{Old}(5)$. Combining this result with that of the previous paragraph, we see that the quotient module $\mathrm{New}(5)$ has dimension 5.

It remains to find generators for $\mathrm{New}(5)$. From $\mathrm{RCF}(E(5))$ we extract a basis for its nullspace, and sort these 29 vectors by increasing Euclidean norm (from 18 to 74). Starting with $\mathrm{RCF}(C(5))$ we apply the same module generators algorithm to these 29 vectors, and find that the first vector increases the rank from 24 to 29. Hence (the coset of) this single vector is a generator for New(5); this vector has 18 (nonzero) terms, and all coefficients are ± 1.

Remark 2.25. We can obtain slightly better results using the LLL algorithm for lattice basis reduction [12]. This depends on the fact that the nonzero entries of $\mathrm{RCF}(E(5))$ are all integers ($\pm 1, \pm 2$). We compute a 120×120 integer matrix U with determinant ± 1 such that $UE(5)^t$ is the Hermite normal form of the transpose of $E(5)$. The bottom 29 rows of U form a lattice basis for the integer nullspace of $E(5)$. We sort these vectors by increasing Euclidean norm (from 16 to 34), and proceed as in Example 2.24. The first vector increases the rank from 24 to 29, and is the coefficient vector of the linearized Hall identity $[\,[x_1, x_2] \circ [x_3, x_4], x_5\,] \equiv 0$, where $[x, y] = xy - yx$ is the Lie bracket and $x \circ y = xy + yx$ is the Jordan product. Drensky [18] has shown that $s_4 \equiv 0$ and the Hall identity $[[x, y]^2, z] \equiv 0$ generate the T-ideal of identities satisfied by 2×2 matrices over fields of characteristic 0.

2.6 Representations of S_n and multilinear identities in degree n. We use the representation theory of the symmetric group to split the preceding computations into smaller pieces, one for each irreducible representation. This significantly reduces the sizes of the matrices involved. Fix a partition λ of n with irreducible representation of dimension d_λ. Let E^λ_{ij} for $i, j = 1, \dots, d_\lambda$ be the matrix units.

Lemma 2.26. *It suffices to consider only the matrix units in the first row, in the following sense. Let M_λ be an irreducible submodule of type λ in the left regular representation $\mathbb{F}S_n$. Then there exists a generator $f \in M_\lambda$ such that its matrix form $\phi_\lambda(f)$ is in RCF and has rank 1 (the only nonzero row is the first).*

Proof: In the left regular representation, row i can be moved to row 1 by left-multiplying by the element of $\mathbb{F}S_n$ which is the image under ψ of the elementary matrix which transposes row 1 and row i. Recall that the matrix units in row i are linear combinations of the elements $E_i s_{ij}$: combine Definitions 1.31, 1.42, 1.44 and equation (8). We can left-multiply by any $p \in S_n$ and obtain another element in the same matrix algebra. In particular, if $p = s_{1i}$ then by Proposition 1.22 we obtain $s_{1i} E_i s_{ij} = E_1 s_{1i} s_{ij} = E_1 s_{1j}$. Thus left-multiplication by s_{1i} moves the matrix units in row i to row 1. The other rows are zero by irreducibility. $\square$

Recall that $d = \dim(A)$ and $t = t(n)$ is the number of association types in degree n. In the direct sum of t copies of the left regular representation, the λ-component is isomorphic to the direct sum of t copies of the full matrix algebra $M_{d_\lambda}(\mathbb{F})$. We construct a matrix M of size $(td_\lambda + d) \times td_\lambda$, consisting of an upper block of size $td_\lambda \times td_\lambda$ and a lower block of size $d \times td_\lambda$.

Notation 2.27. Recall the multilinear associative polynomial U^λ_{1j} of degree n: the image under ψ of the matrix unit E^λ_{1j}. For $k = 1, \dots, t$ write $[U^\lambda_{1j}]_k$ for the nonassociative result of applying association type k to every term of U^λ_{1j}.

2.6.1 Fill and reduce, with representation theory. Given n pseudorandom elements of the algebra A, we can evaluate $[U^\lambda_{1j}]_k$ using the structure constants of A to obtain another element of A. We do this for each $k = 1, \dots, t$ and each $j = 1, \dots, d_\lambda$ to obtain a sequence of td_λ elements of A, which we regard as column vectors of dimension d. We store each of these column vectors in the corresponding column of the lower block of M, and then compute $\mathrm{RCF}(M)$.

We repeat this process until the rank of M stabilizes; at this point, the nullspace of M contains the coefficient vectors of the polynomial identities satisfied by A in the component of $(\mathbb{F}S_n)^t$ corresponding to partition λ. We compute the canonical basis of the nullspace, and call its dimension a_λ. We put the basis vectors into another matrix of size $a_\lambda \times td_\lambda$, and compute its RCF. This matrix, denoted $\texttt{allmat}(\lambda)$, contains the canonical form of the polynomial identities for A in partition λ.

2.6.2 New identities, with representation theory. We need to compare $\texttt{allmat}(\lambda)$ with the representation matrix for the consequences of known identities of lower degrees. We construct a matrix of size $\ell d_\lambda \times td_\lambda$ consisting of $d_\lambda \times d_\lambda$ blocks where ℓ is the number of consequences. The block in position (i, j) where $i = 1, \dots, \ell$ and $j = 1, \dots, t$ is the representation matrix for the terms of i-th consequence in association type j. We compute the RCF of this matrix, and call its rank o_λ. We denote the resulting $o_\lambda \times td_\lambda$ matrix of full rank by $\texttt{oldmat}(\lambda)$; this is the canonical form of the consequences in partition λ.

Since the row space of $\texttt{oldmat}(\lambda)$ is a subspace of the row space of $\texttt{allmat}(\lambda)$ we have $o_\lambda \leq a_\lambda$. Furthermore, $\texttt{oldmat}(\lambda) = \texttt{allmat}(\lambda)$ if and only if $o_\lambda = a_\lambda$; in this case, there are no new identities for the algebra A in partition λ. We have

$$\texttt{jleading}(\,\texttt{oldmat}(\lambda)\,) \subseteq \texttt{jleading}(\,\texttt{allmat}(\lambda)\,).$$

The rows of $\texttt{allmat}(\lambda)$ whose leading 1s occur in the columns with indices in

$$\texttt{jleading}(\,\texttt{allmat}(\lambda)\,) \setminus \texttt{jleading}(\,\texttt{oldmat}(\lambda)\,),$$

represent new identities for the algebra A in partition λ. (This is the representation theoretic version of Lemma 2.23.)

2.6.3 Explicit identities, with representation theory. Consider one of the matrix rows which represents a new identity for the algebra A:

$$[\, c^\lambda_{11}, \dots, c^\lambda_{1d_\lambda}, \quad \dots, \quad c^\lambda_{k1}, \dots, c^\lambda_{kd_\lambda}, \quad \dots, \quad c^\lambda_{t1}, \dots, c^\lambda_{td_\lambda} \,] \qquad (1 \leq k \leq t).$$

As explained, we may assume that this is row 1 of the matrix, and so we may regard it as representing a linear combination of the elements $[U^\lambda_{1j}]_k$ where $1 \leq k \leq t$ and $1 \leq j \leq d_\lambda$. From this we obtain an explicit form of the new identity:

$$\sum_{k=1}^{t} \sum_{j=1}^{d_\lambda} c^\lambda_{k,j}\, [U^\lambda_{1j}]_k \equiv 0. \tag{16}$$

In general, identities of this form have a very large number of terms, when fully expanded as elements of $\mathbb{F}S_n$, especially when n becomes large.

2.7 The membership problem for T-ideals. A basic question about polynomial identities satisfied by an algebra is the following.

Problem 2.28. Let $f^1, \dots, f^k$ and f be multilinear polynomial identities of degree n satisfied by an algebra A. Does f belong to the S_n-module generated by $f^1, \dots, f^k$? Equivalently, is f a linear combination of permutations of $f^1, \dots, f^k$?

Let $\phi_\lambda\colon \mathbb{F}\,S_n \to M_{d_\lambda}(\mathbb{F})$ be the projection onto the λ-component in the Wedderburn decomposition (1). Let $f = f_1 + \cdots + f_t$ be the decomposition of $f \in (\mathbb{F}S_n)^t$ into terms corresponding to the $t = t(n)$ association types.

Definition 2.29. The **representation matrix** of f for λ equals:

$$\phi_\lambda(f) = \left[\begin{array}{c|c|c|c|c} \phi_\lambda(f_1) & \phi_\lambda(f_2) & \cdots & \phi_\lambda(f_{t-1}) & \phi_\lambda(f_t) \end{array} \right].$$

More generally, the representation matrix for a sequence of identities $f^1, \dots, f^k$ is obtained by stacking the matrices $\phi_\lambda(f^1), \dots, \phi_\lambda(f^k)$:

$$\phi_\lambda(f^1, \dots, f^k) = \begin{bmatrix} \phi_\lambda(f^1) \\ \phi_\lambda(f^2) \\ \vdots \\ \phi_\lambda(f^k) \end{bmatrix} = \begin{bmatrix} \phi_\lambda(f^1_1) & \phi_\lambda(f^1_2) & \cdots & \phi_\lambda(f^1_{t-1}) & \phi_\lambda(f^1_t) \\ \phi_\lambda(f^2_1) & \phi_\lambda(f^2_2) & \cdots & \phi_\lambda(f^2_{t-1}) & \phi_\lambda(f^2_t) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_\lambda(f^k_1) & \phi_\lambda(f^k_2) & \cdots & \phi_\lambda(f^k_{t-1}) & \phi_\lambda(f^k_t) \end{bmatrix}$$

Proposition 2.30. *Let $f^1, \dots, f^k$ and f be multilinear polynomial identities of degree n. Then the following conditions are equivalent:*

- *f belongs to the S_n-module generated by $f^1, \dots, f^k$,*
- *the matrices $\phi_\lambda(f^1, \dots, f^k)$ and $\phi_\lambda(f^1, \dots, f^k, f)$ have the same row space,*
- *the matrices $\phi_\lambda(f^1, \dots, f^k)$ and $\phi_\lambda(f^1, \dots, f^k, f)$ have the same RCF,*
- *the matrices $\phi_\lambda(f^1, \dots, f^k)$ and $\phi_\lambda(f^1, \dots, f^k, f)$ have the same rank.*

Example 2.31. Every alternative algebra A satisfies the multilinear identity

$$f(x,y,z,t) = (xy,z,t) + (x,y,[z,t]) - x(y,z,t) - (x,z,t)y \equiv 0.$$

To prove this we need to verify that f is a consequence of the alternative laws. Assuming $\mathrm{char}(\mathbb{F}) \neq 2$, the alternative laws are equivalent to their linearizations (Example 2.20). The consequences of these identities in degree 4 are as follows; some follow from others using the alternative laws:

$$\begin{aligned} f^1 &= (xt,y,z) + (y,xt,z) \equiv 0, & f^6 &= (xt,y,z) + (xt,z,y) \equiv 0, \\ f^2 &= (x,yt,z) + (yt,x,z) \equiv 0, & f^7 &= (x,yt,z) + (x,z,yt) \equiv 0, \\ f^3 &= (x,y,zt) + (y,x,zt) \equiv 0, & f^8 &= (x,y,zt) + (x,zt,y) \equiv 0, \\ f^4 &= (x,y,z)t + (y,x,z)t \equiv 0, & f^9 &= (x,y,z)t + (x,z,y)t \equiv 0, \\ f^5 &= t(x,y,z) + t(y,x,z) \equiv 0, & f^{10} &= t(x,y,z) + t(x,z,y) \equiv 0. \end{aligned}$$

In degree 4, there are $t = 5$ association types. For each $\lambda \vdash 4$ we use Clifton's algorithm to calculate the matrices

$$M_\lambda = \phi_\lambda(f^1, \dots, f^{10}), \qquad N_\lambda = \phi_\lambda(f^1, \dots, f^{10}, f),$$

and compute their RCFs. For example, when $\lambda = 22$ we have $d_\lambda = 2$ and so the matrix M_λ has size 20×10 and N_λ has size 22×10. In Figure 4 we display N_λ and its RCF, which coincides with the RCF of M_λ. Further calculations show that for all $\lambda \vdash 4$ the ranks of M_λ and N_λ are equal:

λ	4	31	22	211	1111
d_λ	1	3	2	3	1
rank	4	12	8	10	2

We conclude that $f(x, y, z, t)$ belongs to the S_4-module generated by the consequences in degree 4 of the linearized forms of the alternative laws.

$$
\left[\begin{array}{rr|rr|rr|rr|rr}
-1 & 1 & 1 & 0 & 1 & -1 & -1 & 0 & 0 & 0 \\
-1 & 0 & 0 & 1 & 1 & 0 & 0 & -1 & 0 & 0 \\
\hline
0 & 1 & 1 & -1 & 0 & -1 & -1 & 1 & 0 & 0 \\
1 & 0 & 0 & -1 & -1 & 0 & 0 & 1 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 2 & -1 & 0 & 0 & -2 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
2 & -1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
-2 & 1 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 \\
-2 & 1 & 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 1 & -1 & 0 & 1 & -1 & 1 & 0 & -1 \\
0 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 \\
\hline
0 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 1 & -1 & 0 & 1 & -1 & 1 & 0 & -1 \\
\hline
1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & -2 & 1 & 2 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & -2 & 1 & 2 & -1 \\
\hline
1 & 1 & 0 & -1 & -1 & 1 & -1 & 0 & 1 & -1 \\
-1 & 2 & 1 & -1 & 0 & 1 & 0 & -1 & 0 & -1
\end{array}\right]
\qquad
\left[\begin{array}{rr|rr|rr|rr|rr}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1
\end{array}\right]
$$

FIGURE 4. The matrix N_λ and its RCF from Example 2.31

2.8 Bondari's algorithm for finite-dimensional algebras. Bondari [5], [6] introduced an algorithm using the representation theory of S_n which computes an independent generating set for the multilinear identities (and central identities) of the full matrix algebra $M_k(\mathbb{F})$ with $\operatorname{char}\mathbb{F} = 0$ or $\operatorname{char}\mathbb{F} = p > n$ where n is the degree of the identities under consideration. He constructed all the multilinear identities of degrees ≤ 8 for $M_3(\mathbb{F})$, confirming existing results in the literature and discovering a new central identity in degree 8. Bondari's algorithm can be used to find multilinear polynomial identities up to a certain degree (depending on computational limitations) for any algebra A over $\mathbb{F}$ of dimension $d < \infty$. This algorithm involves evaluating matrix units in $\mathbb{F}S_n$ using the structure constants of A with respect to a chosen basis.

Definition 2.32. Fix $\lambda \vdash n$ and $f = f_1 + \cdots + f_t \in (\mathbb{F}S_n)^t$. The rank of the matrix $\phi_\lambda(f)$ is called the **rank of f for λ**. If this rank is 1, then we say that f is **irreducible for λ**. (That is, the isotypic component of type λ in the submodule generated by f is irreducible.)

Consider $f \in (\mathbb{F}S_n)^t$ and let r be the rank of the matrix $\operatorname{RCF}(\phi_\lambda(f))$. Each of the r nonzero rows $g_1, \dots, g_r$ generates an irreducible submodule of type λ, and the isotypic component of type λ is the direct sum of these r isomorphic submodules; in other words, r is the multiplicity of λ in the submodule generated by f. Extending Lemma 2.26 to the case of $t > 1$ association types, we see that each g_i can be regarded independently as an irreducible identity for λ in the first row of the matrix.

Lemma 2.33. *Every polynomial identity $f \in (\mathbb{F}S_n)^t$ is equivalent to a finite set of identities, each of which is irreducible for some $\lambda \vdash n$.*

Proof: This is another way of saying that every finite dimensional S_n-module over $\mathbb{F}$ is the direct sum of irreducible modules. □

Recall the images of the matrix units, $U^\lambda_{1j} = \psi(E^\lambda_{1j}) \in \mathbb{F}S_n$. The general element $h \in \mathbb{F}S_n$ which is irreducible for $\lambda \vdash n$ has the form

$$h = \sum_{k=1}^{t} \sum_{j=1}^{d_\lambda} x^k_{1j} [U^\lambda_{1j}]_k \qquad (x^k_{1j} \in \mathbb{F}). \tag{17}$$

Suppose that A has basis $b_1, \dots, b_d$. We describe one iteration of Bondari's algorithm. We choose arbitrary elements $a_1, \dots, a_n \in A$ and evaluate the $[U^\lambda_{1j}]_k$:

$$[U^\lambda_{1j}]_k(a_1, \dots, a_n) = \sum_{i=1}^{d} c^i_{kj} b_i. \tag{18}$$

(This step can be very time-consuming, since the number of terms in the elements $U_{1j}^{\lambda} \in \mathbb{F}S_n$ is roughly $n!$.) Combining the last two equations we obtain

$$h(a_1, \ldots, a_n) = \sum_{i=1}^{d} \left[\sum_{k=1}^{t} \sum_{j=1}^{d_\lambda} c_{kj}^{i} x_{1j}^{k} \right] b_i. \tag{19}$$

If h is an identity for A then the coefficient of each b_i must be 0 for all $a_1, \ldots, a_n \in A$:

$$\sum_{k=1}^{t} \sum_{j=1}^{d_\lambda} c_{kj}^{i} x_{1j}^{k} = 0 \qquad (1 \leq i \leq d). \tag{20}$$

This is a homogeneous linear system of d equations in the td_λ coefficients x_{1j}^{k} of the identity. We compute the RCF of the coefficient matrix, and find its rank.

After s iterations, we have a linear system of sd equations. We repeat this process until the rank stabilizes. We then solve the system by computing the nullspace of the RCF. The nonzero vectors in the nullspace are (probably) coefficient vectors of identities satisfied by A. We need to check these identities by further computations.

Example 2.34. We describe an application of Bondari's method to loops and loop algebras. Recall that a loop L is a set with a binary operation $*$ which has a two-sided unit element and in which the equation $a * b = c$ has a unique solution whenever any two of the three elements are specified. If $\mathbb{F}$ is a field of characteristic $\neq 2$ (resp. $= 2$) then L is called an RA loop (resp. RA2 loop) if the loop algebra $\mathbb{F}L$ is alternative; the initials RA stand for "ring alternative". There are three RA2 loops of order 16 which are not RA loops [30]; in each case, the loop algebra is isomorphic to $\mathbb{F}^8 \oplus A$ where A is a simple algebra. Bondari's algorithm was used to calculate the minimal identities for these 8-dimensional algebras; these identities have degree 4 and are in fact the same in all three cases. Further investigations [23] showed that these three simple 8-dimensional algebras A are in fact isomorphic, and all their identities are satisfied by a large class of loop algebras.

2.9 Rational and modular arithmetic. In general, we prefer to do all linear algebra computations over the field $\mathbb{Q}$ of rational numbers. However, even if a large matrix is very sparse and its entries are very small, computing its RCF can produce exponential increases in the entries. Even if enough computer memory is available to store the intermediate results, the calculations can take far too long. It is therefore often convenient to use modular arithmetic, so that each entry uses a fixed small amount of memory. This leads to the issue of rational reconstruction: recovering correct results over $\mathbb{Q}$ or $\mathbb{Z}$ from known results over $\mathbb{F}_p$.

2.9.1 Rational reconstruction. This process is not well-defined: we want to compute the inverse of a partially defined ∞-to-1 map $\mathbb{Q} \to \mathbb{F}_p$. It is only possible when we have a good theoretical understanding of the expected results. For our purposes, Remark 1.46 justifies our assumption that the correct rational coefficients have denominators which are divisors of $n!$, where n is the degree of the identities [11, Lemma 8]. If $p > n!$ then we can guess the common denominator b of the nonzero rational coefficients a/b from the distribution of the congruence classes modulo p: the modular coefficients will be clustered near the congruence classes representing a/b for $1 \leq a \leq b{-}1$. This allows us to recover rational coefficients which are correct with high probability; we then multiply by the LCM of the denominators to get integers, and finally divide by the GCD to make the identity primitive.

2.9.2 Probability of error. By an error we mean that Gaussian elimination over $\mathbb{Q}$ produces a row with leading nonzero entry a/b (before normalizing to 1) which is 0 modulo p: that is, $\gcd(a,b) = 1$ and $p \mid a$. On average, the probability of error is $1/p$. We can make the leading entry 1 over $\mathbb{Q}$, but it will remain 0 over $\mathbb{F}_p$. If the algebra A has dimension d, then each iteration of fill and reduce produces another d linear constraints on the coefficients of the identity, and we expect to perform d operations of scalar multiplication of a row during the iteration. Hence the chance that no error occurs is $(1 - 1/p)^d$. The chance that an error does occur before the rank stabilizes, and remains for s iterations after it stabilizes, is therefore $X(p,d,s) = (1 - (1{-}1/p)^d)^s$. For example, if we use $p = 101$ for the octonions ($d = 8$) and wait only $s = 10$ iterations after stabilization before terminating, then $X \approx 0.688 \cdot 10^{-11}$, which for practical purposes is indistiguishable from 0.

2.9.3 Nonexistence of identities. Suppose that $f \equiv 0$ is an identity with rational coefficients satisfied by the algebra A which has integral structure constants with respect to a given basis. We multiply f by the LCM of the denominators of its coefficients, obtaining a polynomial f' with integral coefficients; we then divide f' by the GCD of its coefficients, obtaining a polynomial f'' whose coefficients are integers with no common prime factor. Clearly $f'' \equiv 0$ is an identity satisfied by A, and the reduction of f'' modulo p is nonzero for every prime p. Thus existence of identities over $\mathbb{Q}$ implies existence of identities over $\mathbb{F}_p$ for all p. Equivalently, nonexistence over $\mathbb{F}_p$ for a single p implies nonexistence over $\mathbb{Q}$. In this way, we can verify nonexistence of rational identities using computations with modular arithmetic.

2.9.4 Lattice basis reduction. Most of our computations require finding a basis of integer vectors for the nullspace of an integer matrix. In many cases, modular methods give good results: the reconstructed coefficient vectors have small Euclidean lengths. But in other cases, we obtain better results by combining the Hermite normal form (HNF) of an integer matrix with the LLL algorithm for lattice basis reduction. If M is an $s \times t$ matrix over $\mathbb{Z}$ then computing the HNF of its transpose produces integer matrices H $(t \times s)$ and U $(t \times t)$ such that $\det(U) = \pm 1$ and $UM^t = H$. If $\operatorname{rank}(M) = r$ then the bottom $t{-}r$ rows of U

form a lattice basis for the left integer nullspace of M^t, which is the right integer nullspace of M. We then apply the LLL algorithm to obtain shorter basis vectors [12, §3], [7].

3. Polynomial identities of Cayley-Dickson algebras

3.1 Alternative algebras. The standard reference for the theory of alternative algebras is the "Russian book" [53]. The most important alternative algebra is the division algebra $\mathbb{O}$ of real octonions, which arises from the Cayley-Dickson doubling process [53, §2.2]: $\mathbb{R} \subset \mathbb{C} \subset \mathbb{H} \subset \mathbb{O}$. Cayley-Dickson algebras (also called generalized octonion algebras) are 8-dimensional alternative algebras $C(\alpha, \beta, \gamma)$ depending on parameters $\alpha, \beta, \gamma \in \mathbb{F} \setminus \{0\}$. Cayley-Dickson algebras are quadratic algebras, in the sense that they are unital algebras C over $\mathbb{F}$ such that every $x \in C$ satisfies $x^2 - t(x)x + n(x)1 = 0$, where the trace $t\colon C \to \mathbb{F}$ is a linear map and the norm $n\colon C \to \mathbb{F}$ is a quadratic form. If $x = a \cdot 1 + \sum_{i=1}^{7} a_i e_i \in C$ has conjugate $\overline{x} = a \cdot 1 - \sum_{i=1}^{7} a_i e_i \in C$ then the trace and the norm are as follows:

$$
\begin{aligned}
t(x) &= x + \overline{x} = 2a, \\
n(x) &= x\overline{x} = a^2 - \alpha a_1^2 - \beta a_2^2 + \alpha\beta a_3^2 - \gamma a_4^2 + \alpha\gamma a_5^2 + \beta\gamma a_6^2 - \alpha\beta\gamma a_7^2.
\end{aligned}
$$

Kleinfeld [34] classified simple alternative algebras that are not nilalgebras in terms of Cayley-Dickson algebras. Furthermore, we have:

Theorem 3.1 ([53]). *A simple nonassociative alternative algebra is a Cayley-Dickson algebra over its center* (*which is a field*).

If $\mathbb{F} = \mathbb{R}$ then $C(-1, -1, -1) = \mathbb{O}$. If $\operatorname{char} \mathbb{F} \neq 2$ then it is possible to choose a basis $1, e_1, \ldots, e_7$ of $C(\alpha, \beta, \gamma)$ so that its multiplication table is given by Figure 5.

	1	e_1	e_2	e_3	e_4	e_5	e_6	e_7
1	1	e_1	e_2	e_3	e_4	e_5	e_6	e_7
e_1	e_1	α	e_3	αe_2	e_5	αe_4	$-e_7$	$-\alpha e_6$
e_2	e_2	$-e_3$	β	$-\beta e_1$	e_6	e_7	βe_4	βe_5
e_3	e_3	$-\alpha e_2$	βe_1	$-\alpha\beta$	e_7	αe_6	$-\beta e_5$	$-\alpha\beta e_4$
e_4	e_4	$-e_5$	$-e_6$	$-e_7$	γ	$-\gamma e_1$	$-\gamma e_2$	$-\gamma e_3$
e_5	e_5	$-\alpha e_4$	$-e_7$	$-\alpha e_6$	γe_1	$-\alpha\gamma$	γe_3	$\alpha\gamma e_2$
e_6	e_6	e_7	$-\beta e_4$	βe_5	γe_2	$-\gamma e_3$	$-\beta\gamma$	$-\beta\gamma e_1$
e_7	e_7	αe_6	$-\beta e_5$	$\alpha\beta e_4$	γe_3	$-\alpha\gamma e_2$	$\beta\gamma e_1$	$\alpha\beta\gamma$

FIGURE 5. Multiplication table of the generalized octonions

Problem 3.2. Find a basis for the T-ideal $T(C)$ of polynomial identities of a Cayley-Dickson algebra C. (This is Problem 1.55 in the Dniester Notebook [17].)

Isaev [28] found a finite basis of $T(C)$ when $\mathbb{F}$ is finite. Iltyakov [25] proved that $T(C)$ is finitely generated when $\operatorname{char}\mathbb{F} = 0$ but did not give a set of generators.

We now mention some other results obtained "by hand" (that is, by theoretical insight) without any reliance on computer algebra. These proofs show what properties of the algebra yield the identity in question, and give more information about its structure. In contrast, while computer algebra can be extremely effective as an exploratory tool, it usually gives us no understanding of "why" the algebra satisfies the identity. In characteristic 0, Shestakov & Zhukavets [47] found a basis of three identities (one of degree 5, two of degree 6) for the skew-symmetric identities of $\mathbb{O}$. In characteristic $\neq 2, 3, 5$, Shestakov [46] found a basis of identities for split Cayley-Dickson algebras C modulo the associator ideal of a free alternative algebra; that is, a basis for a homomorphic image $T'(C)$ in the free associative algebra of the T-ideal $T(C)$ of identities of C. Henry [20] found a basis for the $\mathbb{Z}_2^2$-graded and $\mathbb{Z}_2^3$-graded identities for Cayley-Dickson algebras (the latter case requires characteristic $\neq 2$).

Using computer algebra, Bremner & Hentzel [9] studied identities for alternative algebras which are built out of associators; in degree 7, they found two identities satisfied by the associator in every alternative algebra, and five identities satisfied by the associator in $\mathbb{O}$.

3.2 Review of identities of degree ≤ 6. The identities for C of degree ≤ 5 when $\operatorname{char}\mathbb{F} \neq 2, 3, 5$ are well-known [53, Lemma 2.8 & Corollary]; see also Racine [40], [41]. One of the simplest of these identities is this:

$$[[x, y]^2, x] \equiv 0. \tag{21}$$

Every quadratic algebra, and hence every Cayley-Dickson algebra, satisfies:

$$V(t^2) - V(t) \circ t \equiv 0, \quad \text{where } V_x(y) = x \circ y, \;\; V = \sum_{\sigma \in S_3} \epsilon(\sigma) V_{x^\sigma} V_{y^\sigma} V_{z^\sigma}. \tag{22}$$

For every $a, b \in C$, we have $a \circ b - t(a)b - t(b)a + q(a, b)1 = 0$ and $t([a, b]) = 0$. It follows that $[x, y] \circ [z, t] \in \mathbb{F}1$ for every $x, y, z, t \in C$, implying this identity:

$$[\,[x, y] \circ [z, t],\, w\,] \equiv 0. \tag{23}$$

A new identity of degree 6 satisfied by Cayley-Dickson algebras was found by Hentzel & Peresi using Bondari's algorithm. We summarize all of these results as follows.

Theorem 3.3 ([24])**.** *The identities of degree $n \leq 6$ of Cayley-Dickson algebras are as follows, where either* $\operatorname{char}\mathbb{F} = 0$ *or* $\operatorname{char}\mathbb{F} = p > n$*:*

$n \leq 2$	*no identities*		
$n = 3$	$(x, x, y) \equiv 0,$	$(x, y, y) \equiv 0$	(*alternative laws*)
$n = 4$	*no new identities*		

$$n=5 \qquad V(t^2)-V(t)\circ t\equiv 0, \qquad [[x,y]\circ[z,t],w]\equiv 0$$

$$n=6 \qquad \Big[\sum_{\sigma\in S_5}\epsilon(\sigma)\big(24x(y(z(tw)))+8x((y,z,t)w)-11(x,y,(z,t,w))\big),\ u\Big]\equiv 0,$$

where σ permutes x,y,z,t,w and $\epsilon\colon S_5\to\{\pm1\}$ is the sign.

(*We list only the identities which are not consequences of those of lower degrees.*)

3.3 Computational study of multilinear identities of degree ≤ 7. We apply the computational techniques described in previous sections to the multilinear polynomial identities satisfied by the algebra $\mathbb{O}$ of octonions. We recover all the existing results in the literature on identities in degree ≤ 6, and then show that there are no new identities in degree 7. As basis for $\mathbb{O}$ over $\mathbb{F}$ we take the symbols $1, e_1, \dots, e_7$. The structure constants depend on parameters $\alpha,\beta,\gamma\in\mathbb{F}$ (Figure 5). If $\mathbb{F}=\mathbb{R}$ and $\alpha=\beta=\gamma=-1$ then we obtain the alternative division algebra of real octonions, which is the case we consider in what follows.

3.3.1 Degree 3. Every multilinear identity of degree 3 satisfied by $\mathbb{O}$ follows from the linearizations of the alternative laws [10, §9, Example 1]. Example 2.20 gives the alternative laws, their linearizations, and their consequences in degree 4.

3.3.2 Degree 4. Every multilinear identity of degree 4 satisfied by $\mathbb{O}$ follows from the consequences of the alternative laws [41]. We will verify this result using our computational methods. The partitions $\lambda\vdash 4$ are 4, 31, 22, 211, 1111 with corresponding dimensions $d_\lambda = 1, 3, 2, 3, 1$. The $t=5$ association types are $((**)*)*$, $(*(**))*$, $(**)(**)$, $*((**)*)$, $*(*(**))$. We give details for $\lambda=22$; the other cases are similar. The standard tableaux are $\begin{array}{|c|c|}\hline 1&2\\\hline 3&4\\\hline\end{array}$, $\begin{array}{|c|c|}\hline 1&3\\\hline 2&4\\\hline\end{array}$ and the elements $U^\lambda_{11}, U^\lambda_{12}\in\mathbb{Q}S_4$ corresponding to the first row matrix units are

$$\begin{aligned}
U^\lambda_{11}=\psi(E^\lambda_{11}) &= 1234-1432-3214+3412+1243-1342-4213+4312\\
&\quad+2134-2431-3124+3421+2143-2341-4123+4321,\\
U^\lambda_{12}=\psi(E^\lambda_{12}) &= 1324-1342-3124+3142+1423-1432-4123+4132\\
&\quad+2314-2341-3214+3241+2413-2431-4213+4231.
\end{aligned}$$

We create an 18×10 matrix consisting of 2×2 blocks, with a 10×10 upper block and an 8×10 lower block. The columns correspond to the following elements of the direct sum of $t=5$ copies of $\mathbb{F}S_4$, where the subscripts give the association types:

$$[U^\lambda_{11}]_1\ [U^\lambda_{12}]_1\quad [U^\lambda_{11}]_2\ [U^\lambda_{12}]_2\quad [U^\lambda_{11}]_3\ [U^\lambda_{12}]_3\quad [U^\lambda_{11}]_4\ [U^\lambda_{12}]_4\quad [U^\lambda_{11}]_5\ [U^\lambda_{12}]_5$$

Any identity for $\mathbb{O}$ of type λ can be expressed as a linear combination of these elements. The fill-and-reduce algorithm converges after one iteration to this matrix:

$$\left[\begin{array}{cc|cc|cc|cc|cc}
1&0&1&0&1&0&1&0&1&0\\
0&1&0&1&0&1&0&1&0&1
\end{array}\right]$$

We find a basis for the nullspace and calculate its RCF, obtaining the matrix whose rows represent identities of type λ satisfied by $\mathbb{O}$:

$$\texttt{allmat}(\lambda) = \left[\begin{array}{cc|cc|cc|cc|cc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{array}\right]$$

Using Clifton's algorithm we obtain the matrix representing the 10 consequences in degree 4 of the alternative laws for partition λ; this is $M_\lambda = \phi_\lambda(f^1, \ldots, f^{10})$ from Example 2.31, whose RCF equals $\texttt{allmat}(\lambda)$.

3.3.3 Degree 5. The multilinear form of identity (22) can be written as

$$x^2 s_3^+(y,z,t) - x s_3^+(y,z,t) \circ x \equiv 0, \qquad s_3^+(x,y,z) = s_3(R_\circ(x), R_\circ(y), R_\circ(z)),$$

where $s_3^+(x,y,z)$ is an operator acting on the right, s_3 is the standard polynomial of degree 3, and $R_\circ$ is the (right) multiplication operator using $\circ$: $xR_\circ(y) = x \circ y$; this follows the notation of [41]. Identities (21) and (23) are satisfied by $\mathbb{O}$, but are not quite sufficient to generate New(5). The S_5-module New(5) is generated by (22) and (23). Using our computational techniques, we obtained the results summarized in Figure 6. Column r_{all} gives the multiplicity of the irreducible S_5-module $[\lambda]$ in the module of all multilinear identities satisfied by $\mathbb{O}$. Column r_{old} gives the multiplicity of $[\lambda]$ in the module of all consequences of the alternative laws. Column $r_{\text{old}+21+22}$ gives the multiplicity of $[\lambda]$ in the module generated by the consequences of the alternative laws and the two identities (21) and (22). From this we see that (21) and (22) are sufficient in the first four representations, but in each of the last three representations, the multiplicities are one less than required. Column $r_{\text{old}+22+23}$ gives the multiplicity of $[\lambda]$ in the module generated by the consequences of the alternative laws together with the identities (22) and (23); these values are the same as r_{all} for all λ, and the corresponding matrices are equal. The last two columns verify that, modulo the consequences of the alternative laws, neither of the identities (22) or (23) generates New(5) by itself, and that these two identities are independent (neither is implied by the other).

We conclude this discussion by presenting explicit matrices to illustrate how we can obtain new identities from the matrix units in the group algebra. For the last partition $\lambda = 11111$ with dimension $d_\lambda = 1$, we obtain the matrices allmat(λ) and oldmat(λ) displayed in Figure 7, with ranks of 11 and 10 respectively.

The row space of oldmat(λ) is a subspace of the row space of allmat(λ). Row 9 of allmat(λ) has a leading 1 in column 9, but oldmat(λ) has no leading 1 in this column. Therefore row 9 of allmat(λ) represents an identity satisfied by $\mathbb{O}$ which is not a consequence of the alternative laws. In terms of matrix units, this row is

λ	d_λ	r_{all}	r_{old}	$r_{\text{old}+21+22}$	$r_{\text{old}+22+23}$	$r_{\text{old}+22}$	$r_{\text{old}+23}$
5	1	13	13	13	13	13	13
41	4	52	52	52	52	52	52
32	5	66	65	66	66	65	66
311	6	76	75	76	76	76	75
221	5	64	63	63	64	63	64
2111	4	48	46	47	48	47	47
11111	1	11	10	10	11	10	11

FIGURE 6. Multiplicities of irreducible modules in degree 5

$$\text{allmat}(\lambda) = \left[\begin{array}{cccccccc|ccc|ccc}
1&0&0&0&0&0&0&0&0&0&0&-2&1&0\\
0&1&0&0&0&0&0&0&0&0&0&-2&0&1\\
0&0&1&0&0&0&0&0&0&0&0&-1&0&0\\
0&0&0&1&0&0&0&0&0&0&0&-3&1&1\\
0&0&0&0&1&0&0&0&0&0&0&0&-1&0\\
0&0&0&0&0&1&0&0&0&0&0&-1&0&0\\
0&0&0&0&0&0&1&0&0&0&0&-1&0&0\\
0&0&0&0&0&0&0&1&0&0&0&-1&0&0\\
0&0&0&0&0&0&0&0&1&0&0&-1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&-2&0&1\\
0&0&0&0&0&0&0&0&0&0&1&1&-1&-1
\end{array}\right]$$

$$\text{oldmat}(\lambda) = \left[\begin{array}{cccccccccccccc}
1&0&0&0&0&0&0&0&-3&0&0&1&1&0\\
0&1&0&0&0&0&0&0&-1&0&0&-1&0&1\\
0&0&1&0&0&0&0&0&-2&0&0&1&0&0\\
0&0&0&1&0&0&0&0&-2&0&0&-1&1&1\\
0&0&0&0&1&0&0&0&-1&0&0&1&-1&0\\
0&0&0&0&0&1&0&0&-1&0&0&0&0&0\\
0&0&0&0&0&0&1&0&0&0&0&-1&0&0\\
0&0&0&0&0&0&0&1&-2&0&0&1&0&0\\
0&0&0&0&0&0&0&0&0&1&0&-2&0&1\\
0&0&0&0&0&0&0&0&0&0&1&1&-1&-1
\end{array}\right]$$

FIGURE 7. Matrices for new octonion identities ($\lambda = 11111$)

$E_{9,9} - E_{9,12}$ and therefore represents the following identity:

$$\sum_{\sigma\in S_5} \epsilon(\sigma)\Big[(x_{\sigma(1)}x_{\sigma(2)})(x_{\sigma(3)}(x_{\sigma(4)}x_{\sigma(5)})) - x_{\sigma(1)}((x_{\sigma(2)}x_{\sigma(3)})(x_{\sigma(4)}x_{\sigma(5)}))\Big] \equiv 0.$$

3.3.4 Degree 6. Hentzel & Peresi [24] discovered a multilinear central polynomial of degree 5 for $\mathbb{O}$, which produces the following polynomial identity where

S_5 permutes x, y, z, t, w:

$$\Big[\sum_{\sigma\in S_5} \epsilon(\sigma)\big(24x(y(z(tw))) + 8x((y,z,t)w) - 11(x,y,(z,t,w))\big),\; u\Big] \equiv 0. \tag{24}$$

Shestakov & Zhukavets [47] found a somewhat simpler central polynomial which produces the following polynomial identity:

$$\Big[\sum_{\sigma\in S_5} \epsilon(\sigma)\big(12([x,y][z,t])w - [[[[x,y],z],t],w]\big),\; u\Big] \equiv 0. \tag{25}$$

Using our computational techniques, we obtained the results in Figure 8.

λ	d_λ	r_{all}	r_{alt}	r_{old}
6	1	41	41	41
51	5	205	205	205
42	9	372	369	372
411	10	409	406	409
33	5	207	205	207
321	16	660	652	660
3111	10	407	400	407
222	5	204	202	204
2211	9	368	360	368
21111	5	202	194	202
111111	1	40	36	39

FIGURE 8. Irreducible multiplicities in degree 6

Column r_{all} gives the multiplicity of the irreducible S_6-module $[\lambda]$ in the module of all multilinear identities satisfied by $\mathbb{O}$. Column r_{alt} gives the multiplicity of $[\lambda]$ in the module of all consequences of the alternative laws. Column r_{old} gives the multiplicity of $[\lambda]$ in the module generated by the consequences of the alternative laws and the identities (22) and (23). From this we see that $r_{\text{old}} = r_{\text{all}}$ except for $\lambda = 111111$ where the difference is 1; hence there is a new identity which alternates in all 6 variables. We further checked that the multiplicities for the alternative laws, (22) and (23) together with either (24) or (25) are equal to r_{all} for all λ; hence either (24) or (25) can be taken as the new generator in degree 6.

Our computations led us to the following new identity in degree 6, which involves only two of the 42 association types, alternates in all 6 variables, and does not have the form $[f(v,w,x,y,z),u] \equiv 0$ where f is a central polynomial:

$$\sum_{\sigma\in S_6} \epsilon(\sigma)\Big(5x_1(x_2((x_3x_4)(x_5x_6))) - x_1(x_2(x_3(x_4(x_5x_6))))\Big) \equiv 0. \tag{26}$$

We can use this identity instead of (24) or (25) as the new generator in degree 6.

3.3.5 Degree 7. Our computations show that degree 7 has no new identities.

Theorem 3.4. *Every multilinear polynomial identity of degree* ≤ 7 *satisfied by the octonion algebra* $\mathbb{O}$ *is implied by the consequences of the alternative laws, the identities* (22) *and* (23)*, and either identity* (24) *or* (25) *or* (26)*.*

We therefore conclude this paper with the following conjecture.

Conjecture 3.5. *The alternative laws together with the identities* (22)*,* (23)*, and either* (24) *or* (25) *or* (26)*, generate the* T*-ideal of polynomial identities satisfied by the octonion algebra* $\mathbb{O}$*.*

Acknowledgments. We thank the two referees for constructive suggestions which significantly improved the paper. Murray Bremner was supported by a Discovery Grant from NSERC, the Natural Sciences and Engineering Research Council of Canada. Sara Madariaga was supported by a Postdoctoral Fellowship from PIMS, the Pacific Institute for the Mathematical Sciences. Luiz Peresi thanks the Department of Mathematics and Statistics at the University of Saskatchewan for its hospitality and financial support during his visits in summer 2012 and spring 2014.

References

[1] Amitsur A., Levitzki J., *Minimal identities for algebras*, Proc. Amer. Math. Soc. **1** (1950), 449–463.

[2] Benanti F., Demmel J., Drensky V., Koev P., *Computational approach to polynomial identities of matrices—a survey*, Polynomial Identities and Combinatorial Methods (Pantelleria, 2001), 141–178, Lecture Notes in Pure and Appl. Math., 235, Dekker, New York, 2003.

[3] Bergdolt G., *Tilted irreducible representations of the permutation group*, Comput. Phys. Comm. **86** (1995), no. 1–2, 97–104.

[4] Boerner H., *Representations of Groups. With Special Consideration for the Needs of Modern Physics*, (second English edition), North-Holland Publishing Co., Amsterdam-London, American Elsevier Publishing Co., Inc., New York, 1970.

[5] Bondari S., *Constructing the Identities and the Central Identities of Degree* < 9 *of the* $n \times n$ *Matrices*, Ph.D. Thesis, Iowa State University, 1993.
`http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=11403&context=rtd`

[6] Bondari S., *Constructing the polynomial identities and central identities of degree* < 9 *of* 3×3 *matrices*, Linear Algebra Appl. **258** (1997), 233–249.

[7] Bremner M., *Lattice Basis Reduction: An Introduction to the LLL Algorithm and Its Applications*, Pure and Applied Mathematics, 300, CRC Press, Boca Raton, 2012.

[8] Bremner M., Dotsenko V., *Algebraic Operads: An Algorithmic Companion*, Chapman and Hall/CRC, Boca Raton, 2016.

[9] Bremner M., Hentzel I., *Identities for the associator in alternative algebras*, J. Symbolic Comput. **33** (2002), no. 3, 255–273.

[10] Bremner M., Murakami L., Shestakov I., *Nonassociative algebras*, Chapter 69 of Handbook of Linear Algebra, edited by Leslie Hogben, Chapman & Hall/CRC, Boca Raton, 2007.

[11] Bremner M., Peresi L., *Nonhomogeneous subalgebras of Lie and special Jordan superalgebras*, J. Algebra **322** (2009), no. 6, 2000–2026.

[12] Bremner M., Peresi L., *An application of lattice basis reduction to polynomial identities for algebraic structures*, Linear Algebra Appl. **430** (2009), no. 2–3, 642–659.

[13] Bremner M., Peresi L., *Special identities for quasi-Jordan algebras*, Comm. Algebra **39** (2011), no. 7, 2313–2337.

[14] Clifton J., *Complete sets of orthogonal tableaux*, Ph.D. Thesis, Iowa State University, 1980; `http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=7688&context=rtd`

[15] Clifton J., *A simplification of the computation of the natural representation of the symmetric group* S_n, Proc. Amer. Math. Soc. **83** (1981), no. 2, 248–250.

[16] Dehn M., *Über die Grundlagen der projektiven Geometrie und allgemeine Zahlsysteme*, Math. Ann. **85** (1922), no. 1, 184–194.

[17] Filippov V., Kharchenko V., Shestakov I. (editors), *Dniester Notebook: Unsolved Problems in the Theory of Rings and Modules*, Nonassociative Algebra and its Applications, 461–516, Lect. Notes Pure Appl. Math., 246, Chapman & Hall/CRC, Boca Raton, 2006; translated by Murray Bremner and Mikhail Kotchetov.
`http://math.usask.ca/ bremner/research/publications/dniester.pdf`

[18] Drensky V., *A minimal basis for identities of a second-order matrix algebra over a field of characteristic 0*, Algebra Logic **20** (1981), no. 3, 188–194.

[19] Drensky V., Kasparian A., *Polynomial identities of eighth degree for* 3×3 *matrices*, Annuaire Univ. Sofia Fac. Math. Méc. **77** (1983), no. 1, 175–195.

[20] Henry F., *Some graded identities of the Cayley-Dickson algebra*, arxiv.org/abs/1205.5057, (submitted on 22 May 2012).

[21] Hentzel I., *Processing identities by group representation*, Computers in Nonassociative Rings and Algebras, (Special Session, 82nd Annual Meeting, Amer. Math. Soc., San Antonio, Tex., 1976), pages 13–40, Academic Press, New York, 1977.

[22] Hentzel I., *Applying group representation to nonassociative algebras*, Ring Theory (Proc. Conf., Ohio Univ., Athens, Ohio, 1976), 133–141, Lecture Notes in Pure and Appl. Math., 25, Dekker, New York, 1977.

[23] Hentzel I., Juriaans S., Peresi L., *Polynomial identities of RA and RA2 loop algebras*, Comm. Algebra **35** (2007), no. 2, 589–595.

[24] Hentzel I., Peresi L., *Identities of Cayley-Dickson algebras*, J. Algebra **188** (1997), no. 1, 292–309.

[25] Iltyakov A., *The Specht property of ideals of identities of certain simple nonassociative algebras*, Algebra Logic **24** (1985), no. 3, 210–228.

[26] Iltyakov A., *Finiteness of the basis of identities of a finitely generated alternative PI-algebra over a field of characteristic zero*, Siberian Math. J. **32** (1991), no. 6, 948–961.

[27] Iltyakov A., *On finite basis of identities of Lie algebra representations*, Nova J. Algebra Geom. **1** (1992), no. 3, 207–259.

[28] Isaev I., *Identities of a finite Cayley-Dickson algebra*, Algebra Logic **23** (1984), no. 4, 407–418.

[29] Jacobson N., *Structure theory for algebraic algebras of bounded degree*, Ann. of Math. (2) **46** (1945), 695–707.

[30] Juriaans S., Peresi L., *Polynomial identities of RA2 loop algebras*, J. Algebra, **213** (1999), no. 2, 557–566.

[31] Kaplansky I., *Rings with a polynomial identity*, Bull. Amer. Math. Soc. **54** (1948), 575–580.

[32] Kemer A., *Finite basability of identities of associative algebras*, Algebra Logic **26** (1987), no. 5, 362–397.

[33] Kemer A., *Ideals of Identities of Associative Algebras*, Translations of Mathematical Monographs, 87, American Mathematical Society, Providence, 1991.

[34] Kleinfeld E., *Simple alternative rings*, Ann. of Math. **58** (1953), no. 2, 544–547.

[35] Knuth D., *The Art of Computer Programming*, vol. 3: Sorting and Searching, (second edition), Addison-Wesley, Reading, 1998.

[36] Koshlukov P., *Algebras with polynomial identities*, 15th School of Algebra, Canela, Brazil, 1998. Mat. Contemp. **16** (1999), 137–186.

[37] Leron U., *Multilinear identities of the matrix ring*, Trans. Amer. Math. Soc. **183** (1973), 175–202.
[38] Malcev A., *On algebras defined by identities*, Mat. Sbornik N.S. **26(68)** (1950), 19–33.
[39] Novelli J., Pak I., Stoyanovskii A., *A direct bijective proof of the hook-length formula*, Discrete Math. Theor. Comput. Sci. **1** (1997) no. 1, 53–67.
[40] Racine M., *Minimal identities for Jordan algebras of degree 2*, Comm. Algebra **13** (1985), no. 12, 2493–2506.
[41] Racine M., *Minimal identities of octonion algebras*, J. Algebra **115** (1988), no. 1, 251–260.
[42] Razmyslov Y., *Identities of Algebras and Their Representations*, Translations of Mathematical Monographs, 138, American Mathematical Society, Providence, 1994.
[43] Regev A., *The representations of S_n and explicit identities for P.I. algebras*, J. Algebra **51** (1978), no. 1, 25–40.
[44] Regev A., *On the Codimensions of Matrix Algebras*, Algebra — Some Current Trends, Varna, 1986, 162–172, Lecture Notes in Math., 1352, Springer, Berlin, 1988.
[45] Rutherford D., *Substitutional Analysis*, Edinburgh, at the University Press, 1948.
[46] Shestakov I., *Associative identities of octonions*, Algebra Logic **49** (2011), no. 6, 561–565.
[47] Shestakov I., N. Zhukavets, *Skew-symmetric identities of octonions*, J. Pure Appl. Algebra **213** (2009), no. 4, 479–492.
[48] Specht W., *Gesetze in Ringen I.*, Math. Z. **52** (1950), 557–589.
[49] Vaĭs A., Zel'manov E., *Kemer's theorem for finitely generated Jordan algebras*, Soviet Math. (Iz. VUZ) **33** (1989), no. 6, 38–47.
[50] van der Waerden B., *Algebra, Vol. II*, Based in part on lectures by E. Artin and E. Noether, Springer, New York, 1991.
[51] Wagner W., *Über die Grundlagen der projektiven Geometrie und allgemeine Zahlensysteme*, Math. Ann. **113** (1937), no. 1, 528–567.
[52] Young A., *The Collected Papers of Alfred Young (1873–1940)*, With a foreword by G. de B. Robinson and a biography by H.W. Turnbull, Mathematical Expositions, 21, University of Toronto Press, Toronto, Ont., Buffalo, N.Y., 1977.
[53] Zhevlakov K., Slin'ko A., Shestakov I., Shirshov A., *Rings That Are Nearly Associative*, Pure and Applied Mathematics, 104, Academic Press, Inc., New York-London, 1982; translated by Harry F. Smith.

Murray R. Bremner:
Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, Canada

E-mail: bremner@math.usask.ca

Sara Madariaga:
Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon, Canada
current address:
Departamento de Matemáticas y Computación, Universidad de La Rioja, Logroño, España

E-mail: sara.madariaga@unirioja.es

Luiz A. Peresi:
Instituto de Matemática e Estatística, Universidade de São Paulo, Brazil

E-mail: peresi@ime.usp.br

(*Received* July 14, 2015, *revised* October 18, 2016)