A Constructive Approach to the Structure Theory for the Group Algebra of the Symmetric Group S_n

Murray R. Bremner* (Canada),
together with
Sara Madariaga (Spain), and Luiz A. Peresi (Brazil)

*Department of Mathematics and Statistics,
University of Saskatchewan,
Saskatoon, Canada
bremner@math.usask.ca

13–24 November 2017,
Süleyman Demirel University,
Almaty, Kazakhstan

Four countries are involved with this short course . . .

Canada

Spain

Brazil

Kazakhstan

These lecture notes are based on . . .

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.

Comment. Math. Univ. Carolin. 57 (2016), no. 4, 413–452.

`cmuc.karlin.mff.cuni.cz` `arXiv:1407.3810[math.RA]`

MR 3583300: "The paper under review surveys results, both classical and new, concerning applications of the theory of representations of the symmetric groups to the study of polynomial identities in algebras. . . . . [T]he authors recall the structure of the group algebra $\mathbb{F}S_n$ in characteristic 0, providing a lot of historical details as well as proofs of the main statements. . . . The authors discuss in detail several efficient methods for analyzing the polynomial identities of a given algebra . . . They also give adequate computer implementations. . . . The most interesting part of the paper is its last section, where the authors study polynomial identities of the octonion algebras. . . . The paper is well written ☺☺☺ and gives a lot of material of interest to specialists in PI algebras as well as in combinatorics and computational algebra." (Plamen Koshlukov)

Abstract

We discuss applications of representation theory of symmetric groups S_n to polynomial identities for associative and nonassociative algebras:

- Section 1 presents complete proofs of the classical structure theory for the group algebra $\mathbb{F}S_n$ over a field $\mathbb{F}$ of characteristic 0 (or $p > n$). We obtain a constructive version of the Wedderburn decomposition (λ is a partition of n, and d_λ counts the standard tableaux of shape λ):
$$\psi\colon \bigoplus_\lambda M_{d_\lambda}(\mathbb{F}) \longrightarrow \mathbb{F}S_n.$$
Alfred Young showed how to compute ψ; to compute ψ^{-1}, we use an efficient algorithm for representation matrices discovered by Clifton.
- Section 2 discusses constructive methods which allow us to analyze the polynomial identities satisfied by a specific (non)associative algebra: the fill and reduce algorithm, the module generators algorithm, and Bondari's algorithm for finite dimensional algebras.
- Section 3 applies these methods to the study of multilinear identities satisfied by octonion algebras over a field of characteristic 0.

Section 1: Structure Theory for the Group Algebra $\mathbb{F}S_n$

- We study the Wedderburn decomposition of the group algebra $\mathbb{F}S_n$ of the symmetric group S_n on n letters usually assumed to be $\{1, 2, \dots, n\}$.
- As a vector space, $\mathbb{F}S_n$ has basis $\{\sigma \mid \sigma \in S_n\}$ over $\mathbb{F}$; multiplication is defined on basis elements by the product in S_n and extended bilinearly.
- We assume that $\mathrm{char}(\mathbb{F}) = 0$; most results hold also if $\mathrm{char}(\mathbb{F}) = p > n$.
- Maschke's Theorem for group algebras implies that if $\mathrm{char}(\mathbb{F}) = 0$ or $\mathrm{char}(\mathbb{F}) = p > n$ then $\mathbb{F}S_n$ is semisimple (since $|S_n| = n!$).
- The structure theory of finite-dimensional associative algebras implies that $\mathbb{F}S_n$ is isomorphic to the direct sum of full matrix algebras with entries in division algebras over $\mathbb{F}$.
- In fact, for $\mathbb{F}S_n$, each of these division algebras is isomorphic to $\mathbb{F}$.
- The Wedderburn decomposition is given by two isomorphisms:

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

- The sum is over all partitions λ of n.
- d_λ is the dimension of the irreducible representation corresponding to λ.

- For the first isomorphism, the matrices $R_\lambda(\pi)$ obtained by restricting ϕ to $\pi \in S_n$, and taking the component of ϕ for partition λ, have entries in $\{0, \pm 1\}$, and form what is called the *natural representation* of S_n:

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

- We will show how to compute the matrices $R_\lambda(\pi)$ for all λ, all $\pi \in S_n$.
- Each matrix algebra $M_{d_\lambda}(\mathbb{F})$ has the standard basis of matrix units E^λ_{ij} for $i, j = 1, \dots, d_\lambda$ which multiply according to the standard relations:

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

- The second isomorphism ψ produces elements $\psi(E^\lambda_{ij})$ in $\mathbb{F}S_n$ (linear combinations of permutations) which obey the same relations:

$$\psi\colon \bigoplus_\lambda M_{d_\lambda}(\mathbb{F}) \longrightarrow \mathbb{F}S_n.$$

- We will show how to calculate the elements $\psi(E^\lambda_{ij})$ of the group algebra.
- 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.

Historical references (1)

- The structure theory of $\mathbb{F}S_n$ was developed by Alfred Young in a series of papers entitled "On quantitative substitutional analysis" appearing in *Proceedings of the London Mathematical Society* from 1900 to 1934.
- These papers were reprinted in *The Collected Papers of Alfred Young (1873–1940)*: Mathematical Expositions, No. 21; University of Toronto Press, 1977; 684 pages. For a scientific biography of Young, see: `rsbm.royalsocietypublishing.org/content/royobits/3/10/761.full.pdf`
- The proofs in Young's papers were simplified by D. E. Rutherford in *Substitutional Analysis*: Edinburgh University Press, 1948; 103 pages.
- The theory was reformulated in modern terminology by H. Boerner, following suggestions by J. von Neumann and B. L. van der Waerden, in *Representation of Groups with Special Consideration for the Needs of Modern Physics*: North-Holland Publishing Co., Amsterdam, and Interscience Publishers, New York, 1963; 325 pages.

Historical references (2)

- A substantial simplification of the algorithms for computing the matrices of the natural representation was introduced in J. Clifton's Ph.D. thesis: *Complete Sets of Orthogonal Tableaux*: Iowa State University, 1980.

– 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.

- Our exposition is based on S. Bondari's Ph.D. thesis: *Constructing the Identities and the Central Identities of Degree Less Than 9 of the $n \times n$ matrices*. Iowa State University, 1993.

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

- Clifton and Bondari were students of I. R. Hentzel; see his papers:

– Processing identities by group representation. *Computers in Nonassociative Rings and Algebras*, 13–40. Academic Press, New York, 1977.

– Applying group representation to nonassociative algebras. *Ring Theory*, 133–141. Lecture Notes Pure Appl. Math., 25. Dekker, New York, 1977.

Young diagrams and tableaux

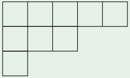
- We start by giving the definitions of the basic objects in the theory.
- The symmetric group S_n is the group of all permutations of $\{1, \ldots, n\}$.
- We write $\lambda \vdash n$ to indicate that λ is a partition of n; that is, $\lambda = (n_1, \ldots, n_k)$ where $n = n_1 + \cdots + n_k$ and $n_1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

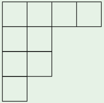
Definition

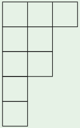
The **Young diagram** Y^λ of the partition $\lambda = (n_1, \ldots, n_k)$ consists of k left-justified rows of empty square boxes where row i contains n_i boxes.

Example

Young diagrams for partitions $\lambda = 531, 4221, 32211$ of $n = 9$:

$Y^{531} =$ 

$Y^{4221} =$ 

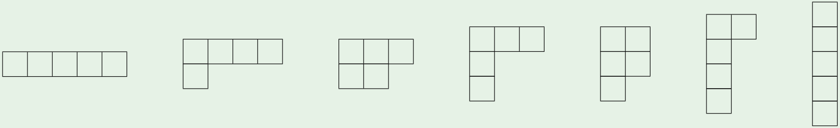
$Y^{32211} =$ 

Definition

Suppose that $\lambda = (n_1, \ldots, n_k)$ and $\lambda' = (n'_1, \ldots, n'_\ell)$ are partitions of n.
We write $\lambda \prec \lambda'$ (or $Y^\lambda \prec Y^{\lambda'}$) and say λ **precedes** λ' if and only if:
either (i) $n_1 < n'_1$ or (ii) $n_1 = n'_1, \ldots, n_i = n'_i$ but $n_{i+1} < n'_{i+1}$ for some i.

Example

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



Definition

A **Young tableau** T^λ of shape λ where $\lambda \vdash n$ is a bijection between the set $\{1, \ldots, n\}$ and the boxes in the Young diagram Y^λ.
The number corresponding to row i and column j will be denoted $T(i, j)$.

Standard tableaux and the hook formula

Definition

We write $T(i,-)$ for the sequence of numbers (left to right) in row i.
We write $T(-,j)$ for the sequence (top to bottom) in column j.
A tableau is **standard** if all sequences $T(i,-)$ and $T(-,j)$ are increasing.

Remark

The **hook formula** for the number d_λ of standard tableaux for Y^λ is as follows; $|h_{ij}|$ is the number of boxes in the hook with NW corner (i,j):

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

There is another version which is easier to implement on a computer; we write $m_i = n_i + k - i$ and $i = 1, \ldots, k$ for $\lambda = (n_1, \ldots, n_k)$:

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

Definition

Given two tableaux T and T' of the same 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 (lex) order** on tableaux is defined as follows:

- We say T **precedes** T' and write $T \prec T'$ if and only if $T(i,j) < T'(i,j)$.

Example

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

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

For this Young diagram the hook formula gives $5!/(4 \cdot 3 \cdot 1 \cdot 2 \cdot 1) = 5$.

Notation

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 $(pT)(i,j) = px$.

Horizontal and vertical permutations

Definition

Let T be a tableau of shape $\lambda = (n_1, \ldots, n_k) \vdash n$.

- The subgroup $G_H(T) \subseteq S_n$ consists of all **horizontal permutations** for the tableau T: the permutations $h \in S_n$ which leave the rows fixed as sets, meaning that for all $i = 1, \ldots, k$, if $x \in T(i, -)$ then $hx \in T(i, -)$.
- The subgroup $G_V(T) \subseteq S_n$ consists of all **vertical permutations** for T: the permutations $v \in S_n$ which leave the columns fixed as sets, meaning that for all $j = 1, \ldots, n_1$, if $x \in T(-, j)$ then $vx \in T(-, j)$.

Remark

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.

Lemmas on $G_H(T)$ and $G_V(T)$

Lemma

Let T be a tableau of shape $\lambda \vdash n$.

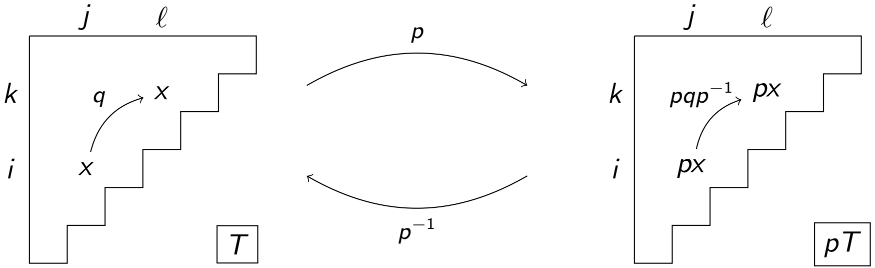
- *We have $G_H(T) \cap G_V(T) = \{\iota\}$ ($\iota \in S_n$ is the identity permutation).*
- *If $h, h' \in G_H(T)$, $v, v' \in G_V(T)$ with $hv = h'v'$ then $h = h'$, $v = v'$.*

Proof.

The first claim is clear. For the second, if $hv = h'v'$ then $(h')^{-1}h = v'v^{-1}$ which belongs to $G_H(T) \cap G_V(T)$ and so both sides equal ι. □

Lemma (conjugacy lemma)

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.

• Suppose the permutation $q \in S_n$ moves the number x from position (i, j) to position (k, ℓ) in the tableau T; see the arrow labelled q on the left.
• If we follow the big curved arrow labelled p^{-1}, then the arrow labelled q on the left, and finally the big curved arrow labelled p, then we see that pqp^{-1} moves $x' = px$ from position (i, j) to position (k, ℓ) of tableau pT.
• This is represented by the arrow labelled pqp^{-1} on the right. Therefore: if $q = h \in G_H(T)$ then $i = k$ and php^{-1} is horizontal permutation for pT; if $q = v \in G_V(T)$ then $j = \ell$ and pvp^{-1} is vertical permutation for pT. □

Row and column intersections

Remark

- We write hvT to indicate that we first apply the vertical permutation $v \in G_V(T)$ to T, then the horizontal permutation $h \in G_H(T)$ to vT.
- But note that h may not be a horizontal permutation for vT!
- We rewrite this using permutations which are horizontal or vertical for the tableaux they act on: $hvT = (hvh^{-1})hT$ for $hvh^{-1} \in G_V(hT)$ (that is, hvh^{-1} is a vertical permutation for hT).

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

Proposition (intersection proposition)

Assume that λ, μ are partitions of n with $Y^\lambda \succ Y^\mu$. For any two tableaux T^λ, T^μ there is a row index i for Y^λ and a column index j for Y^μ such that the intersection $T^\lambda(i, -) \cap T^\mu(-, j)$ contains at least two elements: two numbers appear in the same row of T^λ and the same column of T^μ.

Proof.

- Write $\lambda = (n_1, \ldots, n_k)$ and $\mu = (n'_1, \ldots, n'_\ell)$.
- We proceed by contradiction and assume that $T^\lambda(i, -) \cap T^\mu(-, j)$ contains at most one element for all $1 \leq i \leq k$ and $1 \leq j \leq n'_1$.
- For $i = 1$ this implies that the n_1 numbers in row 1 of T^λ belong to different columns of T^μ, so $n_1 \leq n'_1$.
- But $Y^\lambda \succ Y^\mu$ implies $n_1 \geq n'_1$, so $n_1 = n'_1$.
- The assumption is not affected if we apply a vertical permutation to T^μ, 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 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 $\{a_1, \ldots, a_{n-n_1}\} \subset \{1, \ldots, n\}$ which we can identify with $\{1, \ldots, n - n_1\}$.
- Repeating the argument, we see that $n_2 = n'_2, \ldots, n_k = n'_\ell$; at the end we must have $k = \ell$.
- This implies that $Y^\lambda = Y^\mu$, which is a contradiction. □

Example

Find two numbers in the same row of A and the same column of B, then find two numbers in the same row of B and the same column of C:

$$A = \begin{array}{|c|c|c|c|c|}\hline 3 & 8 & 9 & 5 & 1 \\ \hline 7 & 2 & 0 \\ \cline{1-3} 4 & 6 \\ \cline{1-2}\end{array} \quad \succ_Y \quad B = \begin{array}{|c|c|c|c|}\hline 6 & 7 & 2 & 4 \\ \hline 0 & 1 & 5 \\ \cline{1-3} 9 & 3 & 8 \\ \cline{1-3}\end{array} \quad \succ_Y \quad C = \begin{array}{|c|c|c|c|}\hline 0 & 6 & 2 & 3 \\ \hline 5 & 7 & 4 \\ \cline{1-3} 8 & 9 \\ \cline{1-2} 1 \\ \cline{1-1}\end{array}$$

Lemma

Suppose that

- $\lambda = (n_1, \ldots, n_k)$ *is a partition of* n,
- T *is a tableau of shape* λ, *and*
- p *is a permutation in* S_n.

The following two statements are equivalent:

- $p = hv$ *for some* $h \in G_H(T)$ *and some* $v \in G_V(T)$.
- *The set* $T(i, -) \cap (pT)(-, j)$ *has at most one element for all* $i = 1, \ldots, k$ *and all* $j = 1, \ldots, n_1$.

Proof.

Assume that $p = hv$ for some $h \in G_H(T)$ and $v \in G_V(T)$.

- We have $pT = hvT = (hvh^{-1})hT$ where $hvh^{-1} \in G_V(hT)$.
- Suppose x, y are distinct numbers in the same row of T.
- Then x, y are in the same row but different columns of hT.
- Hence x, y are in different columns of $(hvh^{-1})hT = pT$.

Conversely, assume that the set $T(i,-) \cap (pT)(-,j)$ contains at most one element for all $i = 1, \dots, k$ and $j = 1, \dots, n_1$.

- The numbers in the first column of pT appear in different rows of T.
- Apply $h_1 \in G_H(T)$ so that $(h_1 T)(-,1)$ is a permutation of $(pT)(-,1)$.
- Numbers in $(pT)(-,2)$ appear in distinct rows of $h_1 T$ (columns $j \geq 2$).
- Keeping the numbers in $(h_1 T)(-,1)$ fixed, apply $h_2 \in G_H(T)$ so that $(h_2 h_1 T)(-,2)$ is a permutation of $(pT)(-,2)$.
- Continuing, we obtain $h_1, h_2, \dots, h_{n_1} \in G_H(T)$ so that every number in hT (where $h = h_{n_1} \cdots h_1$) is in the same column as in pT.
- We apply a vertical permutation $v' \in G_V(hT)$ to obtain $v'hT = pT$.
- By the conjugacy lemma, we have $v' = hvh^{-1}$ for some $v \in G_V(T)$.
- Therefore $pT = v'hT = hvh^{-1}hT = hvT$. □

Proposition

Assume $\lambda = (n_1, \ldots, n_k) \vdash n$. Let $T_1, \ldots, T_{d_\lambda}$ be the standard tableaux of shape λ in lex order. If $d_\lambda \geq r > s \geq 1$ then for some $i \in \{1, \ldots, k\}$ and $j \in \{1, \ldots, n_1\}$ the set $T_r(-, j) \cap T_s(i, -)$ has at least two elements.

T_r	$\cdots$	j''	$\cdots$	j'	$\cdots$
$\vdots$		$\vdots$		$\vdots$	
i'	$\cdots$	z	$\cdots$	x	$\cdots$
$\vdots$		$\vdots$		$\vdots$	
i''	$\cdots$	y	$\cdots$		
$\vdots$		$\vdots$			

T_s	$\cdots$	j''	$\cdots$	j'	$\cdots$
$\vdots$		$\vdots$		$\vdots$	
i'	$\cdots$	z	$\cdots$	y	$\cdots$
$\vdots$		$\vdots$		$\vdots$	
i''	$\cdots$		$\cdots$		
$\vdots$		$\vdots$			

Proof.

- Let (i', j') be the first position in which T_r, T_s have a different number.
- Let x and 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 smallest number that has not appeared in a previous row. 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).
- First case: Impossible, since $x > y$ and T_r is standard.
- Second case: $x > y$ implies $j'' < j'$ (y is in a column to the left of x). (We illustrate this situation with the diagram on the previous slide.)
- 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, by the choice of (i', j').
- Hence y, z are in the same column of T_r and the same row of T_s. □

Symmetric and alternating sums

We construct elements of $\mathbb{F}S_n$ which we use to define idempotents in $\mathbb{F}S_n$.

Definition

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:

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

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

Lemma (commutativity lemma)

If T is a tableau of shape $\lambda \vdash n$ with $h \in G_H(T)$ and $v \in G_V(T)$ 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 functions $G_H(T) \to G_H(T)$ sending $h' \mapsto hh'$ and $h' \mapsto h'h$ are bijections; this proves the claim for H_T. Similar 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 (conjugacy proposition)

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

$$H_{pT} = pH_T p^{-1}, \qquad V_{pT} = pV_T p^{-1}.$$

Proof.

Since $\epsilon(p) = \epsilon(p^{-1})$, the result follows from the conjugacy lemma. □

Idempotents and orthogonality in the group algebra

Definition

For $\lambda \vdash n$ let $T_1^\lambda, \ldots, T_{n!}^\lambda$ be all tableaux of shape λ in lex order.
Define $s_{ij}^\lambda \in S_n$ by $s_{ij}^\lambda T_j^\lambda = T_i^\lambda$ so $s_{ji}^\lambda = (s_{ij}^\lambda)^{-1}$, $s_{ij}^\lambda s_{jk}^\lambda = s_{ik}^\lambda$ $(1 \leq i, j \leq n!)$.
Define elements $D_i^\lambda \in \mathbb{F}S_n$ (we omit the superscript 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 (proposition on s_{ij})

If T is a tableau of shape λ then $D_j = s_{ji} D_i s_{ij}$; equivalently, $s_{ij} D_j = D_i s_{ij}$.

Proof.

Using the conjugacy proposition and the definition of s_{ij}, we obtain

$$s_{ji} D_i s_{ij} = s_{ji} H_{T_i} V_{T_i} s_{ji}^{-1} = s_{ji} H_{T_i} s_{ji}^{-1} s_{ji} V_{T_i} s_{ji}^{-1} = H_{s_{ji} T_i} V_{s_{ji} T_i} = H_{T_j} V_{T_j} = D_j,$$

from which the second equation follows immediately. □

Proposition (zero product proposition)

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.

First, assume $Y^\lambda \prec Y^\mu$.
By the intersection proposition, there are numbers $k \neq \ell$ in the same row of $T^\mu = T_i^\mu$ and the same column of $T^\lambda = T_j^\lambda$.
For the transposition $t = (k\ell)$ we have $t \in G_V(T^\lambda)$ and $t \in G_H(T^\mu)$.
Using the commutativity lemma, we obtain

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

Hence $D^\lambda D^\mu = 0$.

Proof (continued).

Second, assume $Y^\lambda \succ Y^\mu$.
The conjugacy proposition implies that

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

By the intersection proposition, there exist $k \neq \ell$ in the same row of T^λ and the same column of pT^μ, so $t = (k\ell) \in G_V(pT^\mu) \cap G_H(T^\lambda)$, and

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

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

$$D^\lambda D^\mu = H_{T^\lambda} V_{T^\lambda} H_{T^\mu} V_{T^\mu} = H_{T^\lambda}\left[\sum_{p \in S_n} x_p\, p\right] V_{T^\mu} = \sum_{p \in S_n} x_p \big(H_{T^\lambda} p V_{T^\mu}\big),$$

and this is 0, where $x_p \in \mathbb{F}$ for all $p \in S_n$. This completes the proof. □

Corollary (zero product corollary)

Assume $\lambda \vdash n$, and let $T_1, \ldots, T_{d_\lambda}$ be the standard tableaux in lex order. If $i > j$ then $D_i D_j = 0$.

Proof.

Write H_i and V_i for H_{T_i} and V_{T_i}.
By the intersection proposition, there exist numbers $k \neq \ell$ in the same column of T_i and the same row of T_j.
Using the transposition $t = (k\ell)$ and the lemma on signs we obtain

$$D_i D_j = H_i V_i H_j V_j = H_i V_i t^2 H_j V_j = H_i V_i t\, t H_j V_j = -H_i V_i H_j V_j = -D_i D_j.$$

Therefore $D_i D_j = 0$. □

Proposition (von Neumann's theorem)

If $\lambda \vdash n$ then for $i = 1, \ldots, n!$ we have $D_i^2 = c_i D_i$ where $c_i = n!/f_i$, and f_i is the dimension of the left ideal $\mathbb{F}S_n D_i$.

Proof.

For some scalars $x_p \in \mathbb{F}$ which we will determine during this proof, we have

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

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

$$hD_i^2 v = h\Big(\sum_{p \in S_n} x_p \, p \Big) v = \sum_{p \in S_n} x_p \, hpv,$$

$$hD_i^2 v = (hH_i)V_i H_i(V_i v) = \epsilon(v) H_i V_i H_i V_i = \epsilon(v) D_i^2.$$

Therefore

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

Each $p \in S_n$ occurs once and only once on each side of this equation.

Proof (continued).

First, consider the coefficient in D_i^2 of a permutation of the form hv.
On the left side of the last equation, take $p = \iota$, on the right side $p = hv$:

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

Second, consider the coefficient of a permutation q not of the form hv.
There are numbers $k \neq \ell$ in same row of T_i and same column of qT_i.
For the transposition $t = (k\ell)$, we have

$$t \in G_H(T_i) \qquad \text{and} \qquad q^{-1}tq \in G_V(T_i).$$

Take $h = t$ and $v = q^{-1}tq$ in the equation at bottom of previous page:

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

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 $x_q = 0$.

Proof (continued).

Combining the last two results, we obtain $D_i^2 = c_i D_i$ where $c_i = x_\iota$, and so it remains to show that $x_\iota = n!/f_i$.
Choose a basis for the left ideal $\mathbb{F}S_n D_i$ consisting of $p_1 D_i, \ldots, p_{f_i} D_i$ for some permutations $p_1, \ldots, p_{f_i} \in S_n$, and extend this to a basis of $\mathbb{F}S_n$.
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 chosen basis has the form

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

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

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

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

Definition

Let $T_1^\lambda, \dots, T_{n!}^\lambda$ be all the tableaux of shape $\lambda \vdash n$, and define

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

Corollary

Every E_i^λ is an idempotent in the group algebra $\mathbb{F}S_n$:

$$(E_i^\lambda)^2 = E_i^\lambda.$$

Example

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline \end{array} \qquad E_{123} = \tfrac{1}{6}\big(123 + 132 + 213 + 231 + 312 + 321\big)$$

$$\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} \end{array} \qquad E_{12.3} = \tfrac{1}{3}\big(123 + 213\big)\big(123 - 132\big)$$

$$\begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline \end{array} \qquad E_{1.2.3} = \tfrac{1}{6}\big(123 - 132 - 213 + 231 + 312 - 321\big)$$

Two-sided ideals in the group algebra

In this section we obtain an explicit description of the isomorphism ψ:

$$\psi \colon \bigoplus_\lambda M_{d_\lambda}(\mathbb{F}) \longrightarrow \mathbb{F}S_n.$$

Definition

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 (xi lemma)

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

Exercise

Work out the other products of the idempotents in the previous example:

$E_{123}E_{12.3}$, $E_{12.3}E_{123}$, $E_{123}E_{1.2.3}$, $E_{1.2.3}E_{123}$, $E_{12.3}E_{1.2.3}$, $E_{1.2.3}E_{12.3}$

Proof.

First, assume that $s_{ji} = vh$ for some $h \in G_H(T_i)$, $v \in G_V(T_i)$. The proposition on s_{ij} 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 $T_j = s_{ji} T_i$, 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

From now on we will work only with standard tableaux.

Definition

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

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; therefore $\mathcal{E}^\lambda$ is invertible.

Proof.

If $i > j$ then the zero product corollary implies $E_i E_j = 0$ and so $\xi_{ij} = 0$.
If $i = j$ then $s_{ii} = \iota$ and so the xi lemma gives $E_i = \xi_{ii} E_i$, hence $\xi_{ii} = 1$. □

Proposition

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 the proposition on s_{ij} and the xi lemma 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

If we replace the scalar ξ_{jk} in the last proposition by the Kronecker delta δ_{jk} (1 if $j = k$, 0 otherwise), and write E_{ij} instead of $E_i s_{ij}$, then we obtain the matrix unit relations $E_{ij} E_{k\ell} = \delta_{jk} E_{i\ell}$.
Therefore, in order to construct the isomorphism ψ, we need to modify slightly the elements $E_i s_{ij}$ to produce other elements which exactly satisfy the matrix unit relations.

Definition

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

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

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

Corollary

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

Proof.

This follows immediately from the last proposition. □

Remark

Our next goal is to show that each subalgebra N^λ is in fact isomorphic to a full matrix algebra. We will do this by constructing a basis for N^λ which satisfies the matrix unit relations.

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

$$\alpha^\lambda(A) = \sum_{i=1}^{d_\lambda} \sum_{j=1}^{d_\lambda} a_{ij} E_i s_{ij} \in \mathbb{F} S_n.$$

Lemma (alpha lemma)

For all partitions $\lambda \vdash n$ and all $i, j, k, \ell \in \{1, \ldots, d_\lambda\}$ we have

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

where E_{ij} is the $d_\lambda \times d_\lambda$ matrix with 1 in position (i, j) and 0 elsewhere.

Proof.

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

Proposition (independence proposition)

The set $\{\, E_i^\mu s_{ij}^\mu \mid \mu \vdash n, 1 \le i, j \le d_\mu \,\}$ *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

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

Using the proposition on s_{ij} and the zero product proposition, we see that all terms vanish except for $\mu = \lambda$:

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

Proof (continued).

Now the alpha lemma gives

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

Hence $\alpha^\lambda(E_{ii}A^\lambda E_{jj}) = 0$ and $\alpha^\lambda(a^\lambda_{ij}E_{ij}) = 0$, and so $a^\lambda_{ij}E_is_{ij} = 0$.
Therefore $a^\lambda_{ij} = 0$ for all λ and all i, j. □

Definition

Suppose that n has r distinct partitions $\lambda_1, \dots, \lambda_r$ in lex order.
Let $d_i = d_{\lambda_i}$ $(1 \le i \le r)$ 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 $\alpha^i = \alpha^{\lambda_i}$:

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

Corollary

The map α is injective.
For every $\lambda \vdash n$ and $1 \leq i,j \leq d_\lambda$, 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 stated in the independence proposition.
Linear independence holds for each λ, so the set spanning N^λ is a basis.
The sum of the N^λ is direct by the zero product proposition.
Since $N \subseteq \mathbb{F}S_n$, it follows that $\sum_\lambda d_\lambda^2 \leq n!$.
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; see Knuth, §5.1.4, Theorem A. □

Matrix units in the group algebra

We prove that the map ψ is an isomorphism by constructing elements of $\mathbb{F}S_n$ corresponding to matrix units.

Remark

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

Definition

For all $\lambda \vdash n$ and $1 \leq i, j \leq d_\lambda$, 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

For all $\lambda, \mu \vdash n$, $1 \leq i, j \leq d_\lambda$, $1 \leq k, \ell \leq d_\mu$ we have

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

Proof.

If $\lambda = \mu$ then

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

The factor $\delta_{\lambda\mu}$ comes from the zero product proposition. □

Definition

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

$$\psi(E_{ij}^{\lambda}) = U_{ij}^{\lambda} \qquad (\lambda \vdash n;\ 1 \leq i, j \leq d_\lambda).$$

Theorem

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

The direct sum M of full matrix algebras is clearly semisimple.
Simplicity is preserved by isomorphism, and therefore $\mathbb{F}S_n$ is semisimple.
Moreover, $\mathbb{F}S_n$ 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
- 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 $d_\lambda/n!$ are defined in characteristic $p > n$, we also obtain the semisimplicity of $\mathbb{F}S_n$ in this case.

Example

$$\mathbb{F}S_2 = \mathbb{F} \oplus \mathbb{F}, \qquad \mathbb{F} \cong M_1(\mathbb{F}), \qquad \mathrm{char}(\mathbb{F}) \neq 2$$

$$\mathbb{F}S_3 = \mathbb{F} \oplus M_2(\mathbb{F}) \oplus \mathbb{F}, \qquad \mathrm{char}(\mathbb{F}) \neq 2, 3$$

$$\mathbb{F}S_4 = \mathbb{F} \oplus M_3(\mathbb{F}) \oplus M_2(\mathbb{F}) \oplus M_3(\mathbb{F}) \oplus \mathbb{F}, \qquad \mathrm{char}(\mathbb{F}) \neq 2, 3$$

Example

For $n = 3$ we take the permutations in lex order as our basis of $\mathbb{F}S_3$:

$$123, \quad 132, \quad 213, \quad 231, \quad 312, \quad 321 \qquad \text{writing } p \text{ as } p(1)p(2)p(3)$$

The partitions $\lambda = 3$, $\mu = 21$, $\nu = 111$ have these standard tableaux:

$$T_1^\lambda = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline\end{array} \qquad T_1^\mu = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \qquad T_2^\mu = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \qquad T_1^\nu = \begin{array}{|c|}\hline 1 \\ \hline 2 \\ \hline 3 \\ \hline\end{array}$$

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:

$$E_{11}^\lambda,\ E_{11}^\mu,\ E_{12}^\mu,\ E_{21}^\mu,\ E_{22}^\mu,\ E_{11}^\nu.$$

We will compute the corresponding elements U_{ij}^ρ of $\mathbb{F}S_n$.

Example (continued)

The groups of horizontal and vertical permutations are as follows:

$$\begin{aligned} G_H(T_1^\lambda) &= S_3, & G_V(T_1^\lambda) &= \{123\}, \\ G_H(T_1^\mu) &= \{123, 213\}, & G_V(T_1^\mu) &= \{123, 321\}, \\ G_H(T_2^\mu) &= \{123, 321\}, & G_V(T_2^\mu) &= \{123, 213\}, \\ G_H(T_1^\nu) &= \{123\}, & G_V(T_1^\nu) &= S_3. \end{aligned}$$

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

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

Example (continued)

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

$$
\begin{aligned}
E^\lambda_1 &= \tfrac{1}{6}(123 + 132 + 213 + 231 + 312 + 321),\\
E^\mu_1 &= \tfrac{1}{3}(123 + 213 - 312 - 321),\\
E^\mu_2 &= \tfrac{1}{3}(123 - 213 - 231 + 321),\\
E^\nu_1 &= \tfrac{1}{6}(123 - 132 - 213 + 231 + 312 - 321).
\end{aligned}
$$

Clearly $s^\mu_{12} = s^\mu_{21} = 132$, and this is the only nontrivial case.
Hence $s_{12} \neq vh$ for any $v \in G_V(T^\mu_2)$ and $h \in G_H(T^\mu_2)$ (by the xi lemma).
Therefore every $\mathcal{E}^\rho$ is the identity matrix of size d_ρ.
Hence every $U^\rho_{ij} = \alpha^\rho(E_{ij}) = E^\rho_i s^\rho_{ij}$.
We obtain the following matrix units in the group algebra:

$$
\begin{aligned}
&U^\lambda_{11} = E^\lambda_1, \quad U^\mu_{11} = E^\mu_1, \quad U^\mu_{12} = E^\mu_1 s_{12} = \tfrac{1}{3}(132 + 231 - 312 - 321),\\
&U^\mu_{21} = E^\mu_2 s_{21} = \tfrac{1}{3}(132 - 213 - 231 + 312), \quad U^\mu_{22} = E^\mu_2, \quad U^\nu_{11} = E^\nu_1.
\end{aligned}
$$

Example (continued)

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

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

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

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

Example (concluded)

Putting this all together gives the representation matrices for the three irreducible representations of S_3:

$$\psi^{-1}(123) = \left[1, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, 1\right] \qquad \psi^{-1}(132) = \left[1, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, -1\right]$$

$$\psi^{-1}(213) = \left[1, \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix}, -1\right] \qquad \psi^{-1}(231) = \left[1, \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}, 1\right]$$

$$\psi^{-1}(312) = \left[1, \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, 1\right] \qquad \psi^{-1}(321) = \left[1, \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix}, -1\right]$$

Clifton's theorem on representation matrices

- Our next goal is to compute explicitly the algebra homomorphism ϕ:

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

- We fix $\lambda \vdash n$ and consider all the tableaux $T_1, \ldots, T_{n!}$ 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.
- The proposition on s_{ij} and the xi lemma 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 $\xi_{ij}^p = \xi_{ir}$ when $p = s_{rj}$, so that for all i, j, p we have

$$E_i p E_j = \xi_{ij}^p s_{ij} E_j.$$

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

Definition

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

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

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

By definition of ξ_{ij}, we see that $(A^\lambda_p)_{ij}$ can be computed as follows:

- Apply p to the standard tableau T_j obtaining the tableau pT_j (which may not be standard).
- If there exist two numbers that appear together in a column of T_i and in a row of pT_j, then $(A^\lambda_p)_{ij} = 0$.
- Otherwise, there is a vertical permutation $q \in G_V(T_i)$ which takes each number in T_i into the row it occupies in pT_j, and $(A^\lambda_p)_{ij} = \epsilon(q)$.

The algorithm on the next page attempts to find q, and returns 0 if no such permutation exists.

Algorithm to compute the Clifton matrix A_p^λ

Input: A partition $\lambda = (n_1, \ldots, n_\ell) \vdash n$; a permutation $p \in S_n$. Output: A_p^λ.

For $j = 1, \ldots, d_\lambda$ do:

- Compute pT_j.
- For $i = 1, \ldots, d_\lambda$ do:
  - Set $e \leftarrow 1$, $k \leftarrow 1$, $\beta \leftarrow$ `false`.
  - While $k \leq n$ and not β do:
    - Set $r_i, c_i \leftarrow$ row, column indices of k in T_i.
    - Set $r_j, c_j \leftarrow$ row, column indices of k in pT_j.
    - 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*]
    - Set $k \leftarrow k + 1$
  - Set $(A_p^\lambda)_{ij} \leftarrow e$

Before proving Clifton's theorem, it is worth quoting in its entirety the review by G. D. James (MathSciNet: MR0624907) of Clifton's paper: "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 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."

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

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

Definition

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

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

Proposition (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

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

Proof (continued).

$$
\begin{aligned}
&= \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik}\, \eta_{j\ell}\, \xi^p_{kj}\, s_{ik}\, E_k\, s_{kj}\, s_{j\ell} \\
&= \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik}\, \eta_{j\ell}\, \xi^p_{kj}\, E_i\, s_{ik}\, s_{kj}\, s_{j\ell} \\
&= \sum_{k=1}^{d_\lambda} \sum_{\ell=1}^{d_\lambda} \eta_{ik}\, \eta_{j\ell}\, \xi^p_{kj}\, E_i\, s_{i\ell} \\
&= \left(\sum_{k=1}^{d_\lambda} \eta_{ik}\, \xi^p_{kj} \right) \left(\sum_{\ell=1}^{d_\lambda} \eta_{j\ell}\, E_i\, s_{i\ell} \right) \\
&= \left(\sum_{k=1}^{d_\lambda} \eta_{ik}\, \xi^p_{kj} \right) U_{ij} \\
&= (A_\iota^{-1} A_p)_{ij}\, U_{ij}.
\end{aligned}
$$

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

Example

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 T_i and pT_j, and the vertical permutation q:

$$(i,j) = (1,1) \quad T_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad pT_1 = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \quad q = \iota \quad \epsilon(q) = 1$$

$$(i,j) = (1,2) \quad T_1 = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1}\end{array} \quad pT_2 = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} \quad q = 321 \quad \epsilon(q) = -1$$

$$(i,j) = (2,1) \quad T_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \quad pT_1 = \begin{array}{|c|c|}\hline 2 & 1 \\ \hline 3 \\ \cline{1-1}\end{array} \quad q \text{ does not exist}$$

$$(i,j) = (2,2) \quad T_2 = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1}\end{array} \quad pT_2 = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline 1 \\ \cline{1-1}\end{array} \quad q = 213 \quad \epsilon(q) = -1$$

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

Example

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

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

Let $p = \iota$ (identity) and consider the $(i,j) = (1,5)$ entry of $\mathcal{E} = A_\iota^\lambda$:

$$T_i = T_1, \qquad pT_j = T_5.$$

The required vertical permutation is the transposition $q = 15342 = (25)$:

$$(A_\iota^\lambda)_{15} = -1.$$

Similar calculations give the remaining entries of the matrix:

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

Example

This illustrates 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:

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

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

Definition

Let A be an algebra (not necessarily associative) with $d = \dim_{\mathbb{F}} A < \infty$. The multiplication in A is a bilinear map $A \times A \to A$ 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:

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

Definition

Let $X = \{x_1, x_2, \dots\}$ be a finite or countably infinite set of variables. The **free magma** $M(X)$ generated by X consists of all (nonassociative) **monomials** constructed inductively from $X \subset M(X)$ by the condition

$$x, y \in X,\ v, w \in M(X) \setminus X \implies xy, x(v), (v)x, (v)(w) \in M(X).$$

We write $M(X)_n$ for the subset consisting of monomials of degree n.

Definition

Each monomial in $m \in M(X)$ has a placement of parentheses called an **association type**: if we fix an argument symbol $*$ then the association type of $m \in M(X)$ is obtained by replacing every variable in m by $*$. Association types have a total order, defined inductively by degree, based on unique factorization $m = m_1 m_2$ of nonassociative monomials. If $m \in M(X)_n$, $m' \in M(X)_{n'}$ $(n \neq n')$ then $m \prec m'$ if and only if $n < n'$. If $m, m' \in M(X)_n$ then we write $m = m_1 m_2$ and $m' = m'_1 m'_2$ and define $m \prec m'$ if and only if either (i) $m_1 \prec m'_1$ or (ii) $m_1 = m'_1$ and $m_2 \prec m'_2$.

Example

For $n = 3$ we have two association types: $(**)*$ and $*(**)$.
For $n = 4$ we have five: $((**)*)*$, $(*(**))*$, $(**)(**)$, $*((**)*)$, $*(*(**))$.
For $n = 5$ we have fourteen:

$(((**)*)*)*$, $((*(**))*)*$, $((**)(**))*$, $(*((**)*))*$, $(*(*(**)))*$,
$((**)*)(**)$, $(*(**))(**)$, $(**)((**)*)$, $(**)(*(**))$,
$(((**)*)*)*$, $((*(**))*)*$, $((**)(**))*$, $(*((**)*))*$, $(*(*(**)))*$.

Lemma

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

The numbers C_n grow very rapidly:

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

Definition

We write $\mathbb{F}\{X\}$ for the vector space over $\mathbb{F}$ with basis $M(X)$.
The **free nonassociative algebra** generated by X over $\mathbb{F}$ is $\mathbb{F}\{X\}$ with multiplication extended bilinearly from $M(X)$.
The elements of $\mathbb{F}\{X\}$ are (nonassociative) **polynomials** in the variables X with coefficients in $\mathbb{F}$.

Definition

The **homogeneous component** $\mathbb{F}\{X\}_n$ is the subspace whose monomial basis is the ordered set $M(X)_n$. Clearly $\mathbb{F}\{X\}_n\mathbb{F}\{X\}_{n'} \subseteq \mathbb{F}\{X\}_{n+n'}$.
A **T-ideal** in $\mathbb{F}\{X\}$ is a two-sided ideal $\mathcal{R} \subseteq \mathbb{F}\{X\}$ such that $f(\mathcal{R}) \subseteq \mathcal{R}$ for any algebra endomorphism $f\colon \mathbb{F}\{X\} \to \mathbb{F}\{X\}$ (that is, $\mathcal{R}$ is closed under arbitrary substitutions).

Definition

A **polynomial identity** satisfied by an algebra A is a polynomial $I \in \mathbb{F}\{X\}$ such that $I \equiv 0$ when elements of A are substituted for the variables X. We write $\equiv$ to mean that the identity holds for all values of the arguments. If $I \in \mathbb{F}\{X\}_n$ then we say I is **homogeneous** of degree n; if $x_1, \ldots, x_n$ each occur exactly once in every monomial then we say I is **multilinear**. We write $T_X(A)$ for the set of polynomial identities in X satisfied by A.

Lemma

$T_X(A)$ is a T-ideal in $\mathbb{F}\{X\}$ which does not depend on the basis of A.

Historical remarks

- Algebras which satisfy polynomial identities (also known as PI-algebras) constitute a very important class of algebras with a large literature.
- Investigation of this topic was initiated in 1922 by Dehn, motivated by problems in geometry.
- Wagner in 1937 found identities for the quaternions and matrix algebras.
- The vigorous development of the theory of PI-algebras began with the work of Jacobson and Kaplansky in the late 1940s.
- In particular, the following is a famous classical problem for PI-algebras.

Problem (Specht)

For a given variety of algebras (associative, Lie, Jordan, alternative, etc.), determine whether every algebra A in this variety has a finite basis of polynomial identities, where "finite basis" means that the T-ideal $T_X(A)$ is finitely generated.

Kemer's theorem and generalizations

Specht originally posed this problem for associative algebras over fields of characteristic 0. The complete solution was found by Kemer around 1990.

Theorem (Kemer)

Every finitely generated associative algebra over a field of characteristic 0 has a finite basis of polynomial identities.

Analogous results were obtained by:

- Vais & Zelmanov for finitely generated Jordan algebras,
- Iltyakov for finitely generated Lie and alternative algebras.

The subspace $T_X(A)_n = T_X(A) \cap \mathbb{F}\{X\}_n$ is the homogeneous component of degree n in the T-ideal of polynomial identities for the algebra A. The nonzero elements of $T_X(A)_n$ are the identities of degree n for A.

Definition

If n is as small as possible such that $T_X(A)_n \neq 0$ then the nonzero elements of $T_X(A)_n$ are called **minimal identities** for A.

Minimal identities

For the simple matrix algebras $M_k(\mathbb{F})$, we have the following result.

Theorem (Amitsur-Levitzki)

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 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 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.
The T-ideal of identities for $M_2(\mathbb{F})$ has been studied by many authors; see the book by Razmyslov. For computational methods, see Benanti et al.

Multilinear polynomial identities

Problem

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

Lemma

Every polynomial identity (not necessarily multilinear or homogeneous) over a field of characteristic 0 is equivalent to a set of multilinear identities.

Proof.

See Chapter 1 of Zhevlakov et al., *Rings That Are Nearly Associative.* □

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 for $1 \leq i \leq t$.
- In each summand, the monomials differ only in the permutation of the variables, so we may identify each summand with an element of $\mathbb{F}S_n$.
- We may therefore regard a multilinear identity $I(x_1, \ldots, x_n)$ as an element 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 and Specht.
- In the 1970's, this theory was developed further by Regev, with a particular focus on associative PI algebras.
- Implementation of these algorithms on a computer was initiated by Hentzel in the 1970's.

Example

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.
We usually choose the right-normed product $x_1(x_2(\cdots(x_{n-1}x_n)\cdots))$ (or the left-normed product), here using the identity permutation.
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.
So a multilinear polynomial identity in degree n can be regarded as an element of the group algebra $\mathbb{F}S_n$.

Example

If the binary operation is commutative (as for Jordan algebras) or anticommutative (as for Lie algebras), then the association types are not independent; for example, $(ab)c = \pm c(ab)$.
In these two cases, the association types are enumerated by what are known as Wedderburn-Etherington numbers $1, 1, 1, 2, 3, 6, 11, 23, 46, \ldots$.

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 $tn!$ is 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 of the matrix $E(n)$.

- 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 $\mathrm{RCF}(E(n))$ (also called reduced row-echelon form), 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 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: set the free variables equal to the standard basis vectors and solve for the leading variables.
- We then 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

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 these steps:

- 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 $1 \le j \le 24$, 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.
The rank remains 23 for the next 10 iterations; hence the nullity is 1.
A basis for the nullspace consists of the coefficient vector of the standard identity of degree 4 (Amitsur-Levitzki theorem):

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

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.
- These consequences do not provide any new information.
- We only want the new identities in degree n: those which cannot be expressed in terms of known identities of lower degrees.

Definition

Let $I(x_1, \ldots, x_n)$ 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 $x_i x_{n+1}$ $(i = 1, \ldots, n)$ and two multiplications of I by x_{n+1} (on the right and the left):

$$I(x_1 x_{n+1}, \ldots, x_n), \ \ldots \ I(x_1, \ldots, x_i x_{n+1}, \ldots, x_n), \ \ldots \ I(x_1, \ldots, x_n x_{n+1}),$$
$$I(x_1, \ldots, x_i, \ldots, x_n) x_{n+1}, \quad x_{n+1} I(x_1, \ldots, x_i, \ldots, x_n).$$

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

Lemma

Every multilinear polynomial of degree $n+1$ in the T-ideal generated by I is a linear combination of permutations of the $n+2$ consequences of I.

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.
The two multiplications correspond to the definition of an ideal. □

Example (alternative laws)

The 8-dimensional algebra $\mathbb{O}$ of octonions is an alternative algebra.
These algebras are defined by the left and right alternative identities $(x, x, y) \equiv 0$, $(x, y, y) \equiv 0$ for $(x, y, z) = (xy)z - x(yz)$ (associator).
Over a field of characteristic $\neq 2$, these two identities are equivalent to their linearized forms: $(x, z, y) + (z, x, y) \equiv 0$, $(x, y, z) + (x, z, y) \equiv 0$.

Algebras, S_n-modules, operads

Example (continues)

Each has 5 consequences in degree 4; the left alternative identity gives

$$\begin{aligned}
&(xw, z, y) + (z, xw, y) \equiv 0, \quad (x, z, yw) + (z, x, yw) \equiv 0, \\
&(x, zw, y) + (zw, x, y) \equiv 0, \\
&(x, z, y)w + (z, x, y)w \equiv 0, \quad w(x, z, y) + w(z, x, y) \equiv 0.
\end{aligned}$$

- 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 $\mathbb{F}\{X\}$, $X = \{x_1, \dots, x_n\}$.
- 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 regard $\mathrm{All}(n)$ as a submodule of $\Sigma\mathrm{Bin}(n)$, the degree n component of the symmetrization of the nonsymmetric operad Bin generated by one nonassociative binary operation with no symmetry.

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 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 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.

Existence of new identities

- 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 $\mathrm{Old}(n)$.
- 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 $r_{\mathsf{old}} = r_{\mathsf{all}}$ 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 $r_{\mathsf{old}} \neq r_{\mathsf{all}}$ then since $\mathrm{Old}(n) \subseteq \mathrm{All}(n)$ we must have $r_{\mathsf{old}} < r_{\mathsf{all}}$, and the row space of $[\mathrm{Old}(n)]$ is a subspace of the row space of $[\mathrm{All}(n)]$.
- The difference $r_{\mathsf{all}} - r_{\mathsf{old}}$ is the dimension of the S_n-module of new identities in degree n.

Definition

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

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

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

Lemma (new generators lemma)

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

Example

We extend the example on 2×2 matrices from degree 4 to degree 5.
We proceed as before, with obvious changes: matrix $E(5)$ is 124×120; each iteration generates 5 random matrices; there are 120 permutations.
The rank increases by 4 for each of the first 22 iterations, but the next iteration produces rank 91, and this remains constant for 10 iterations.
Thus the nullspace of $E(5)$ has dimension 29; this is the S_5-module $\mathrm{All}(5)$: the coefficient vectors of all identities in degree 5 for 2×2 matrices.
To find new identities, we first generate the degree 5 consequences of the standard identity: linear combinations of permutations of the generators:

$$
\begin{array}{lll}
s_4(x_1x_5, x_2, x_3, x_4), & s_4(x_1, x_2x_5, x_3, x_4), & s_4(x_1, x_2, x_3x_5, x_4), \\
s_4(x_1, x_2, x_3, x_4x_5), & x_5 s_4(x_1, x_2, x_3, x_4), & s_4(x_1, x_2, x_3, x_4)x_5.
\end{array}
$$

We construct a 240×120 zero matrix $C(5)$ and perform the following steps for each generator:

Example

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

The rank of $C(5)$ is 24; its row space is the S_5-module $\mathrm{Old}(5)$.
Combining this with the previous result, the quotient module $\mathrm{New}(5)$ has dimension $5 = 29 - 24$; it remains to find generators for $\mathrm{New}(5)$.
From $\mathrm{RCF}(E(5))$ we extract a basis for its nullspace.
We 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; the first vector increases the rank from 24 to 29.
Hence (the coset of) this single vector is a generator for $\mathrm{New}(5)$; this vector has 18 (nonzero) terms, and all coefficients are ± 1.

Remark

We can obtain better results using lattice basis reduction.
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 are a lattice basis for the left integer nullspace of $E(5)^t$, which is the right integer nullspace of $E(5)$.
We use the LLL algorithm to find a lattice basis of shorter vectors.
We sort these vectors by increasing Euclidean norm (from 16 to 34).
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 $[x, y] = xy - yx$ (Lie bracket) and $x \circ y = xy + yx$ (Jordan product).
Drensky has shown that the identity $s_4 \equiv 0$ and the Hall identity $[[x, y]^2, z] \equiv 0$ generate the T-ideal of identities satisfied by 2×2 matrices over a field of characteristic 0.

Representations of S_n; multilinear identities in degree n

We use the representation theory of the symmetric group to split the computation 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_{ij}^λ for $i, j = 1, \ldots, d_\lambda$ be the matrix units.

Lemma (first row lemma)

It suffices to consider only the matrix units in the first row, as follows. Let M_λ be an irreducible submodule corresponding to partition 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 row).

Proof.

In the left regular representation, row i can be moved to row 1 in the corresponding representation matrix 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.
The matrix units in row i are linear combinations of the elements $E_i s_{ij}$.
If we left-multiply by $p \in S_n$ then we obtain another element in the same matrix algebra; in particular, if $p = s_{1i}$ then by the proposition on s_{ij} 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. □

Recall that $d = \dim(A)$ and $t = t(n)$ is the number of association types.
In the direct sum of t copies of the regular representation, the component for partition λ is isomorphic to the direct sum of t copies of the full matrix algebra $M_{d_\lambda}(\mathbb{F})$ (by the Wedderburn decomposition of $\mathbb{F}S_n$).
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 $d \times td_\lambda$.

Fill and reduce, with representation theory

Notation

The multilinear associative polynomial $U^\lambda_{1j} = \psi(E^\lambda_{1j})$ of degree n is the image under ψ of the matrix unit E^λ_{1j}.
We write $[U^\lambda_{1j}]_k$ $(1 \leq k \leq t)$ for the multilinear nonassociative polynomial obtained by applying association type k to every monomial of U^λ_{1j}.

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 $k = 1, \ldots, t$ and each $j = 1, \ldots, d_\lambda$ to obtain a sequence of td_λ elements of A (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.

New identities, with representation theory

When the rank stabilizes, the nullspace of M consists of the coefficient vectors of the polynomial identities satisfied by A in the component of $(\mathbb{F}S_n)^t$ corresponding to partition λ.

Compute the canonical basis of the nullspace, and call its dimension a_λ.

Put the basis vectors into a matrix of size $a_\lambda \times td_\lambda$ and compute its RCF, $\texttt{allmat}(\lambda)$, the canonical form of the identities for A in partition λ.

We need to compare $\texttt{allmat}(\lambda)$ with the representation matrix for the ℓ consequences of known identities of lower degrees.

Construct a matrix of size $\ell d_\lambda \times td_\lambda$ consisting of $d_\lambda \times d_\lambda$ blocks.

The block in position (i, j) where $i = 1, \ldots, \ell$ and $j = 1, \ldots, t$ is the representation matrix for the terms of consequence i 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.

Row space of $\texttt{oldmat}(\lambda)$ is subspace of row space of $\texttt{allmat}(\lambda)$:

$$o_\lambda \leq a_\lambda.$$

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.

Since both matrices are in row canonical form, we have

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

The rows of $\texttt{allmat}(\lambda)$ whose leading 1s occur with column 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 the new generators lemma.)

Explicit identities, with representation theory

Consider a matrix row which represents a new identity for the algebra A:

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

As explained, we may assume that this is row 1 of the matrix.
So we may regard it as representing a linear combination of the elements

$$[U^\lambda_{1j}]_k \text{ where } 1 \le k \le t \text{ and } 1 \le j \le d_\lambda.$$

From this we obtain an explicit form of the new identity:

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

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

The membership problem for T-ideals

A basic question about polynomial identities is the following.

Problem

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

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.
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

The **representation matrix** of f for λ equals:

$$\phi_\lambda(f) = \left[\quad \phi_\lambda(f_1) \quad \middle| \quad \phi_\lambda(f_2) \quad \middle| \quad \cdots \quad \middle| \quad \phi_\lambda(f_{t-1}) \quad \middle| \quad \phi_\lambda(f_t) \quad \right]$$

Definition

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

$$\phi_\lambda(f^1, \ldots, f^k) = \begin{bmatrix} \phi_\lambda(f^1) \\ \phi_\lambda(f^2) \\ \vdots \\ \phi_\lambda(f^k) \end{bmatrix} = \begin{bmatrix} \phi_\lambda(f_1^1) & \phi_\lambda(f_2^1) & \cdots & \phi_\lambda(f_{t-1}^1) & \phi_\lambda(f_t^1) \\ \phi_\lambda(f_1^2) & \phi_\lambda(f_2^2) & \cdots & \phi_\lambda(f_{t-1}^2) & \phi_\lambda(f_t^2) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_\lambda(f_1^k) & \phi_\lambda(f_2^k) & \cdots & \phi_\lambda(f_{t-1}^k) & \phi_\lambda(f_t^k) \end{bmatrix}$$

Proposition

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

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

Example

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

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

Example (continued)

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

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

and compute their RCFs.
For example, when $\lambda = 22$ we have $d_\lambda = 2$:
the matrix M_λ has size 20×10 and N_λ has size 22×10.
On the next page we display N_λ and its RCF.
The RCF of N_λ coincides with the RCF of M_λ.
Further calculations show that for all λ 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.

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

Bondari's algorithm for finite-dimensional algebras

Bondari 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 $\mathrm{char}\,\mathbb{F} = 0$ or $\mathrm{char}\,\mathbb{F} = p > n$ where n is the degree of the identities.
He constructed all the multilinear identities of degrees ≤ 8 for $M_3(\mathbb{F})$, confirming known results and discovering a new central identity for $n = 8$.
Bondari's algorithm can be used to find multilinear polynomial identities up to a certain degree 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

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

Consider $f \in (\mathbb{F}S_n)^t$ and let r be the rank of the matrix $\mathrm{RCF}(\phi_\lambda(f))$.

Each of the r nonzero rows $g_1, \ldots, g_r$ generates an irreducible submodule.

The isotypic component for λ is the direct sum of these r submodules; in other words, r is the multiplicity of λ in the submodule generated by f.

Extending the first row lemma 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

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

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

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 $[U^\lambda_{1j}]_k$:

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

This step can be very time-consuming, since the number of terms in the elements $U^\lambda_{1j} \in \mathbb{F}S_n$ is roughly $n!$.

Combining the last two equations we obtain

$$h(a_1, \dots, a_n) = \sum_{i=1}^{d} \left[\sum_{k=1}^{t} \sum_{j=1}^{d_\lambda} c^i_{kj} x^k_{1j} \right] b_i.$$

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

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

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

This is an application of Bondari's method to loops and loop algebras.
Recall that a loop L is a set with a binary operation $*$ with 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.
Juriaans & Peresi showed that there are three RA2 loops of order 16 which are not RA loops.
Each 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 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.

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 $\mathbb{F}_p$.

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.

By the structure theory of the group algebra, we know that the correct rational coefficients have denominators which are divisors of $n!$, where n is the degree of the identities.

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.

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 mod 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, then $X \approx 0.688 \cdot 10^{-11}$, which for practical purposes is indistinguishable from 0.

Nonexistence of identities

Suppose that $f \equiv 0$ is an identity with rational coefficients satisfied by the algebra A with 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: 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.

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 $\mathrm{rank}(M) = r$ 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.

Section 3: Identities of Cayley-Dickson Algebras

- The standard reference for the theory of alternative algebras is:
  K. Zhevlakov, A. Slin'ko, I. Shestakov, A. Shirshov:
  Rings That Are Nearly Associative.
  Pure and Applied Mathematics, 104.
  Academic Press, Inc., New York-London, 1982.
- See also the unpublished book by K. McCrimmon, available at:
  `mysite.science.uottawa.ca/neher/papers/alternative/`

- The most important example is the division algebra $\mathbb{O}$ of real octonions, arising from the Cayley-Dickson doubling process: $\mathbb{R} \subset \mathbb{C} \subset \mathbb{H} \subset \mathbb{O}$.
- Cayley-Dickson algebras $C(\alpha, \beta, \gamma)$, or generalized octonion algebras, are 8-dimensional alternative algebras depending on $\alpha, \beta, \gamma \in \mathbb{F} \setminus \{0\}$.
- Cayley-Dickson algebras are quadratic algebras: 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:

  $t(x) = x + \overline{x} = 2a,$
  $n(x) = x\overline{x} = a^2 - \alpha a_1^2 - \beta a_2^2 + \alpha\beta a_3^2 - \gamma a_4^2 + \alpha\gamma a_5^2 + \beta\gamma a_6^2 - \alpha\beta\gamma a_7^2.$
- Kleinfeld classified simple alternative algebras that are not nilalgebras in terms of Cayley-Dickson algebras.
- The most general result: *Every 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 $\mathrm{char}\,\mathbb{F} \neq 2$ then there is a basis $1, e_1, \ldots, e_7$ of $C(\alpha, \beta, \gamma)$ so that its multiplication table is as follows:

	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	$-\alpha\gamma e_2$	$\beta\gamma e_1$	$\alpha\beta\gamma$

Main problem

Problem

Find a basis for the T-ideal of polynomial identities of a Cayley-Dickson algebra C. (Problem 1.55 in the Dniester Notebook.)

- Isaev found a finite basis of the T-ideal $T(C)$ when $\mathbb{F}$ is finite.
- Iltyakov proved that $T(C)$ is finitely generated when $\mathrm{char}\,\mathbb{F} = 0$, but did not give a set of generators.
- We now mention some other results obtained "by hand" (in other words, by theoretical insight) without any reliance on computer algebra.
- These proofs show what properties of the algebra produce the identity in question, and give more information about its structure.
- In contrast, computer algebra is very effective as an exploratory tool, but usually gives no understanding "why" the identity holds.

Identities for various classes of alternative algebras

- In characteristic 0, Shestakov & Zhukavets found a basis of identities (one of degree 5, two of degree 6) for the skew-symmetric identities satisfied by the octonion algebra $\mathbb{O}$.
- In characteristic $\neq 2, 3, 5$ Shestakov 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 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 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 $\mathbb{O}$.

Review of identities of degree ≤ 6

The identities for C of degree ≤ 5 when $\mathrm{char}\,\mathbb{F} \neq 2, 3, 5$ are well-known. One of the simplest of these identities is this:

$$(1) \qquad [[x, y]^2, x] \equiv 0.$$

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

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

For every $a, b \in C$, we have

$$a \circ b - t(a)b - t(b)a + q(a, b)1 = 0, \quad \text{and} \quad t([a, b]) = 0.$$

It follows that $[x, y] \circ [z, t] \in \mathbb{F}1$ for every $x, y, z, t \in C$, and hence:

$$(3) \qquad [\,[x, y] \circ [z, t],\ w\,] \equiv 0.$$

A new identity of degree 6 satisfied by Cayley-Dickson algebras was found by Hentzel & Peresi using Bondari's algorithm.

Identities of degree $n \leq 6$ of Cayley-Dickson algebras

We summarize all of these results as follows.

Theorem

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*

$n = 3$: $(x, x, y) \equiv 0$ *and* $(x, y, y) \equiv 0$ *(alternative laws)*

$n = 4$: *no new identities*

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

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

where σ permutes x, y, z, t, w and $\epsilon \colon S_5 \to \{\pm 1\}$ is the sign.

(We list only identities which do not follow from those of lower degrees.)

Computational study: multilinear identities of degree ≤ 7

We apply these computational techniques to the multilinear polynomial identities satisfied by the algebra $\mathbb{O}$ of octonions.

We recover the known results 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 $1, e_1, \ldots, e_7$.

The structure constants depend on parameters $\alpha, \beta, \gamma \in \mathbb{F}$.

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.

Degree 3

Every multilinear identity of degree 3 satisfied by $\mathbb{O}$ follows from the linearizations of the alternative laws.

A previous example shows the alternative laws, their linearizations, and their consequences in degree 4.

Degree 4

Racine showed that every multilinear identity of degree 4 satisfied by $\mathbb{O}$ follows from the consequences of the alternative laws.

We will verify this result using our computational methods.

Partitions: 4, 31, 22, 211, 1111. Corresponding dimensions: 1, 3, 2, 3, 1.

Association types: $((**)*)*$, $(*(**))*$, $(**)(**)$, $*((**)*)$, $*(*(**))$.

We give details for $\lambda = 22$; the other cases are similar.

Standard tableaux: $\begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline\end{array}$ $\begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline\end{array}$

The elements $U^\lambda_{11}, U^\lambda_{12} \in \mathbb{Q}S_4$ corresponding to the first row matrix units:

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

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:

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

Any identity for $\mathbb{O}$ of type λ is 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.

We obtain the matrix whose rows represent identities of type λ for $\mathbb{O}$:

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

Using Clifton's algorithm we obtain the matrix for partition λ which represents the 10 consequences in degree 4 of the alternative laws.

This equals $M_\lambda = \phi_\lambda(f^1, \ldots, f^{10})$ from the degree 4 example, and the RCF of that matrix equals $\texttt{allmat}(\lambda)$.

Degree 5

The multilinear form of $V(t^2) - V(t) \circ t \equiv 0$ can be written as

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

where

- $s_3^+(x,y,z) = s_3(R_\circ(x), R_\circ(y), R_\circ(z))$, operator acting on the right,
- s_3 is the standard polynomial of degree 3,
- $R_\circ$ is the (right) multiplication operator using $\circ$: $xR_\circ(y) = x \circ y$.

Identities (1) and (3) are satisfied by $\mathbb{O}$, but do not generate $\mathrm{New}(5)$.

The S_5-module $\mathrm{New}(5)$ is generated by (2) and (3).

Using our computational techniques, we obtained these results:

λ	d_λ	r_{all}	r_{old}	$r_{\text{old}+1+2}$	$r_{\text{old}+2+3}$	$r_{\text{old}+2}$	$r_{\text{old}+3}$
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	$\boxed{64}$	63	$\boxed{63}$	64	63	64
2111	4	$\boxed{48}$	46	$\boxed{47}$	48	47	47
11111	1	$\boxed{11}$	10	$\boxed{10}$	11	10	11

r_{all}: multiplicity of irreducible S_5-module $[\lambda]$ in module of all multilinear identities satisfied by $\mathbb{O}$

r_{old}: multiplicity of $[\lambda]$ in module of all consequences of alternative laws

$r_{\text{old}+1+2}$: multiplicity of $[\lambda]$ in module generated by consequences of alternative laws and identities (1) and (2)

Identities (1) and (2) are sufficient in the first four representations, but in the last three representations, the multiplicities are one less than required.

$r_{\mathsf{old}+2+3}$: multiplicity of $[\lambda]$ in module generated by consequences of alternative laws together with identities (2) and (3).

These values equal r_{all} for all λ, and the corresponding matrices are equal.

The last two columns $r_{\mathsf{old}+2}$ and $r_{\mathsf{old}+3}$ verify that, modulo consequences of alternative laws, neither (2) or (3) generates $\mathrm{New}(5)$ 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 to obtain new identities from the matrix units in the group algebra.

For partition $\lambda = 11111$ with dimension $d_\lambda = 1$, we obtain the following matrices $\mathrm{allmat}(\lambda)$ and $\mathrm{oldmat}(\lambda)$, with ranks 11 and 10 respectively.

Matrices of new octonion identities for $\lambda = 11111$:

$$
\mathrm{allmat}(\lambda) \;=\; \begin{bmatrix}
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -2 & 1 & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -2 & \cdot & 1 \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -3 & 1 & 1 \\
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \boxed{1} & \cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & -2 & \cdot & 1 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & -1 & -1
\end{bmatrix}
$$

$$
\mathrm{oldmat}(\lambda) = \begin{bmatrix}
1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -3 & \cdot & \cdot & 1 & 1 & \cdot \\
\cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & \cdot & -1 & \cdot & 1 \\
\cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & -2 & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & -2 & \cdot & \cdot & -1 & 1 & 1 \\
\cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & -1 & \cdot & \cdot & 1 & -1 & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & -2 & \cdot & \cdot & 1 & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & -2 & \cdot & 1 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & -1 & -1
\end{bmatrix}
$$

The row space of $\mathrm{oldmat}(\lambda)$ is a subspace of the row space of $\mathrm{allmat}(\lambda)$.

Row 9 of $\mathrm{allmat}(\lambda)$ has a leading 1 in column 9.

However, $\mathrm{oldmat}(\lambda)$ has no leading 1 in this column.

Therefore row 9 of $\mathrm{allmat}(\lambda)$ represents an identity satisfied by $\mathbb{O}$ which is not a consequence of the identities of lower degree.

In terms of matrix units, this row is $E_{9,9} - E_{9,12}$.

It therefore represents the following identity:

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

Degree 6

Hentzel & Peresi discovered a central polynomial of degree 5 for $\mathbb{O}$, which produces the following identity where S_5 permutes x, y, z, t, w:

$$(4) \quad \Big[\sum_{\sigma \in S_5} \epsilon(\sigma) \Big(24x(y(z(tw))) + 8x((y,z,t)w) - 11(x,y,(z,t,w)) \Big), u \Big] \equiv 0.$$

Shestakov and Zhukavets found a simpler central polynomial:

$$(5) \qquad \Big[\sum_{\sigma \in S_5} \epsilon(\sigma)\big(12([x,y][z,t])w - [[[[x,y],z],t],w]\big),\ u \Big] \equiv 0.$$

Using our computational techniques, we obtained the following results:

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	$\boxed{40}$	36	$\boxed{39}$

r_{all}: multiplicity of irreducible S_6-module $[\lambda]$ in module of all multilinear identities satisfied by $\mathbb{O}$

r_{alt}: multiplicity of $[\lambda]$ in module of all consequences of alternative laws

r_{old}: multiplicity of $[\lambda]$ in module generated by consequences of alternative laws and identities (2) and (3).

We see that $r_{\text{old}} = r_{\text{all}}$ except for $\lambda = 111111$ where difference is 1; hence there is a new identity which alternates in all 6 variables.

We checked that the multiplicities for the alternative laws and (2) and (3), together with either (4) or (5), are equal to r_{all} for all λ.

Hence either (4) or (5) can be taken as the new generator in degree 6.

We also found 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 $[f(v,w,x,y,z),u] \equiv 0$:

$$(6) \qquad \sum_{\sigma \in S_6} \epsilon(\sigma)\Big(5\, x_1(x_2((x_3x_4)(x_5x_6))) - x_1(x_2(x_3(x_4(x_5x_6))))\Big) \equiv 0.$$

We can therefore use this identity as the new generator in degree 6.

Degree 7

Our computations showed that degree 7 has no new identities.

Theorem

Every multilinear polynomial identity of degree ≤ 7 satisfied by the octonion algebra $\mathbb{O}$ over a field of characteristic 0 is implied by:

- *the consequences of the alternative laws,*
- *the identities (2) and (3), and*
- *either identity (4) or (5) or (6).*

We therefore conclude with the following conjecture.

Conjecture

The alternative laws, together with the two identities (2) and (3), and one of the identities (4) or (5) or (6), generate the T-ideal of polynomial identities satisfied by the octonion algebra $\mathbb{O}$ in characteristic 0.

Acknowledgements

- We thank the 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 I

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

F. Benanti, J. Demmel, V. Drensky, P. Koev: Computational approach to polynomial identities of matrices — a survey. *Polynomial Identities and Combinatorial Methods* (Pantelleria, 2001), pages 141–178. Dekker, New York, 2003.

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

H. Boerner: *Representations of Groups. With Special Consideration for the Needs of Modern Physics*. Second English edition. North-Holland Publishing Co., Amsterdam-London, 1970.

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

References II

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

M. Bremner: *Lattice Basis Reduction: An Introduction to the LLL Algorithm and Its Applications*. CRC Press, Boca Raton, 2012.

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

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

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

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

References III

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

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

J. Clifton: *Complete Sets of Orthogonal Tableaux*. Ph.D. Thesis, Iowa State University, 1980.
`lib.dr.iastate.edu/cgi/viewcontent.cgi?article=7688&context=rtd`

J. Clifton: 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.

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

References IV

- V. Drensky: 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.
- V. Drensky, A. Kasparian: Polynomial identities of eighth degree for 3×3 matrices. *Annuaire Univ. Sofia Fac. Math. Méc.* 77 (1983), no. 1, 175–195.
- V. Filippov, V. Kharchenko, I. Shestakov (editors): *Dniester Notebook: Unsolved Problems in the Theory of Rings and Modules.* Nonassociative Algebra and its Applications, pages 461–516. Chapman & Hall/CRC, Boca Raton, 2006. `math.usask.ca/~bremner/research/publications/dniester.pdf`
- F. Henry: Some graded identities of the Cayley-Dickson algebra. `arxiv.org/abs/1205.5057` (submitted on 22 May 2012).

References V

- I. Hentzel: 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.
- I. Hentzel: Applying group representation to nonassociative algebras. *Ring Theory* (Proc. Conf., Ohio Univ., Athens, Ohio, 1976), pages 133–141. Dekker, New York, 1977.
- I. Hentzel, S. Juriaans, L. Peresi: Polynomial identities of RA and RA2 loop algebras. *Comm. Algebra* 35 (2007), no. 2, 589–595.
- I. Hentzel, L. Peresi: Identities of Cayley-Dickson algebras. *J. Algebra* 188 (1997) no. 1, 292–309.
- A. Iltyakov: The Specht property of ideals of identities of certain simple nonassociative algebras. *Algebra Logic* 24 (1985), no. 3, 210–228.

References VI

- A. Iltyakov: 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.
- A. Iltyakov: On finite basis of identities of Lie algebra representations. *Nova J. Algebra Geom.* 1 (1992), no. 3, 207–259.
- I. Isaev: Identities of a finite Cayley-Dickson algebra. *Algebra Logic* 23 (1984), no. 4, 407–418.
- N. Jacobson: Structure theory for algebraic algebras of bounded degree. *Ann. of Math.* (2) 46 (1945) 695–707.
- S. Juriaans, L. Peresi: Polynomial identities of RA2 loop algebras. *J. Algebra* 213 (1999), no. 2, 557–566.
- I. Kaplansky: Rings with a polynomial identity. *Bull. Amer. Math. Soc.* 54 (1948) 575–580.

References VII

A. KEMER: Finite basability of identities of associative algebras. *Algebra Logic* 26 (1987), no. 5, 362–397.

A. KEMER: *Ideals of Identities of Associative Algebras*. American Mathematical Society, Providence, 1991.

E. KLEINFELD: Simple alternative rings. *Ann. of Math.* (2) 58 (1953) 544–547.

D. KNUTH: *The Art of Computer Programming, Vol. 3: Sorting and Searching*. Second edition. Addison-Wesley, Reading, 1998.

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

U. LERON: Multilinear identities of the matrix ring. *Trans. Amer. Math. Soc.* 183 (1973) 175–202.

References VIII

A. Malcev: On algebras defined by identities. *Mat. Sbornik N. S.* 26(68) (1950) 19–33.

J. Novelli, I. Pak, A. Stoyanovskii: A direct bijective proof of the hook-length formula. *Discrete Math. Theor. Comput. Sci.* 1 (1997) no. 1, 53–67.

M. Racine: Minimal identities for Jordan algebras of degree 2. *Comm. Algebra* 13 (1985), no. 12, 2493–2506.

M. Racine: Minimal identities of octonion algebras. *J. Algebra* 115 (1988), no. 1, 251–260.

Y. Razmyslov: *Identities of Algebras and Their Representations*. American Mathematical Society, Providence, 1994.

A. Regev: The representations of S_n and explicit identities for P.I. algebras. *J. Algebra* 51 (1978) no. 1, 25–40.

References IX

- A. Regev: On the codimensions of matrix algebras. *Algebra — Some Current Trends* (Varna, 1986), pages 162–172. Lecture Notes in Math., 1352. Springer, Berlin, 1988.
- D. Rutherford: *Substitutional Analysis*. Edinburgh, at the University Press, 1948.
- I. Shestakov: Associative identities of octonions. *Algebra Logic* 49 (2011), no. 6, 561–565.
- I. Shestakov, N. Zhukavets: Skew-symmetric identities of octonions. *J. Pure Appl. Algebra* 213 (2009), no. 4, 479–492.
- W. Specht: Gesetze in Ringen, I. *Math. Z.* 52 (1950) 557–589.
- A. Vaĭs, E. Zel'manov: Kemer's theorem for finitely generated Jordan algebras. *Soviet Math. (Iz. VUZ)* 33 (1989), no. 6, 38–47.

References X

- B. van der Waerden. *Algebra, Vol. II*. Based in part on lectures by E. Artin and E. Noether. Springer, New York, 1991.
- W. Wagner: Über die Grundlagen der projektiven Geometrie und allgemeine Zahlensysteme. *Math. Ann.* 113 (1937), no. 1, 528–567.
- A. Young: *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, No. 21. University of Toronto Press, Toronto, Ont., Buffalo, N. Y., 1977.
- K. Zhevlakov, A. Slin'ko, I. Shestakov, A. Shirshov: *Rings That Are Nearly Associative*. Academic Press, Inc., New York-London, 1982.