Studies in Applied Mathematics, Vol. LIII, No. 3, September 1974

On the Foundations of Combinatorial Theory: IX Combinatorial Methods in Invariant Theory

Stud. Appl. Math. 53 (1974)

By Peter Doubilet, Gian-Carlo Rota, and Joel Stein

Dedicated
to
Beniamino Segre

Contents

1. Introduction

We develop an algebraic system designed for computation with subspaces of a finite-dimensional vector space over an arbitrary field, based upon two operations, which we call join and meet. The join is the same as the wedge product in exterior algebra, and the meet roughly corresponds to Grassmann's regressive product, with one important difference. Whereas Grassmann and all other authors up to and including Bourbaki defined the regressive product by means of the duality of vector spaces, we introduce a special device which enables us to define the meet directly. This device is the notion of *Cayley space*, namely, a vector space endowed with a non-degenerate alternating multilinear form, called the *bracket*. It seems astonishing that this notion should not have been previously singled out, as it is the basic tool—recognized or not—of classical invariant theory. A Cayley space

* Supported by NSF Contract #P36739

should be thought of as a natural companion to Hilbert space and symplectic space.

The present definition leads to the derivation of a complete set of identities holding between join and meet, an undertaking that in the past would have been notationally impossible to carry out. We call these identities the *alternative laws*. The body of this work consists in various applications of the alternative laws. We show that these laws easily yield the classical identities holding among minors of a matrix, as well as a systematic procedure for translation of universal theorems of synthetic projective geometry into identities. The main application we derive of the alternative laws is the *straightening formula*; this can be considered to be the end product and the definitive version of a train of thought which began with Clebsch, was developed by Gordan and Capelli, and later by Young and Turnbull. The straightening formula can be interpreted as giving the solution of a word problem. It is a central result in the characteristic-free theory of the projective group; in fact it holds over commutative rings.

As an application of the straightening formula we obtain a characteristic-free version of the classical theory of representations of the symmetric group, as well as two elementary proofs of the First Fundamental Theorem of invariant theory over arbitrary fields. The only previous work on this subject is Igusa's.

Various other applications, which we hope to further develop elsewhere, are sketched throughout the paper. These will include a thorough treatment of classical invariant theory over arbitrary fields, as well as of the symmetric group.

2. Cayley spaces

Throughout this work V will denote a vector space over an arbitrary field. A *bracket*, written

$$[x_1, \ldots, x_n], \quad \text{where } x_i \in V,$$

is a non-degenerate (that is, not identically zero) multilinear alternating form, taking values in the field.

A *Cayley space* is the pair consisting of the vector space V, together with a bracket.

A *standard Cayley space* is a Cayley space over a vector space V of dimension n, whose bracket has the additional property that for every vector x in V, there exist vectors $x_2, \ldots, x_n$ such that

$$[x, x_2, \ldots, x_n]$$

is not equal to zero. In a standard Cayley space the length of the bracket equals the dimension of the space, and conversely. Unless otherwise stated, all Cayley spaces occurring in this work will be standard.

The *exterior algebra* of a standard Cayley space is constructed by imposing an equivalence relation on sequences of vectors. Given two sequences of vectors of length k, we shall write

$$a_1 \ldots a_k \sim b_1 \ldots b_k$$

when for every choice of the vectors $x_{k+1}, \ldots, x_n$ we have

$$[a_1, \ldots, a_k, x_{k+1}, \ldots, x_n] = [b_1, \ldots, b_k, x_{k+1}, \ldots, x_n].$$

An equivalence class under this relation will be called an *extensor*, or *decomposable k-vector*, and will be written as

$$a_1 \vee a_2 \vee \cdots \vee a_k.$$

The operation $\vee$ is called the *join* (and is elsewhere written $\wedge$; our departure from customary notation is well motivated). Note that the join is non-zero if and only if $\{a_1, \ldots, a_k\}$ is a linearly independent set.

A non-zero extensor is of *step k* if it is the join of k linearly independent vectors. If it is of step zero it is called a *scalar*.

The extensors of V span a vector space of dimension 2^n, called $\bar{V}$, whose elements are called *antisymmetric tensors*. The algebra of $\bar{V}$ together with join is the *exterior algebra* of V. It is an antisymmetric associative algebra with identity (the scalar, one) with the usual properties which will not be recalled here.

The extensors of step n form a one dimensional sub-space of $\bar{V}$. Choosing a basis $\{a_1, \ldots, a_n\}$ of V, whose bracket $[a_1, \ldots, a_n]$ equals one, or a *unimodular basis*, we may construct a basis for this subspace, the element

$$E = a_1 \vee \cdots \vee a_n.$$

E is called the *integral*.

We shall frequently indicate the join of extensors by simple juxtaposition of symbols,

$$ab = a \vee b.$$

Also, if A and B denote two extensors the sum of whose steps is n, we shall write $[A, B] = [AB]$ for their bracket.

Every extensor A defines a unique subspace of the vector space V, namely

$$\bar{A} = \text{span}\{a_1, \ldots, a_k\}$$

where $\{a_1, \ldots, a_k\}$ is any set of vectors such that

$$a_1 \vee \cdots \vee a_k = A.$$

The subspace $\bar{A}$ is called the *support* of A. If A and B are extensors, then $A \vee B$ is non-zero if and only if $\bar{A} \cap \bar{B} = 0$, in which case the support of $A \vee B$ is the sub-space $\bar{A} \cup \bar{B}$ spanned by $\bar{A}$ and $\bar{B}$.

A linear transformation T of V into itself is said to be *unimodular* if it preserves the bracket.

Given an extensor A of step n-k in a standard Cayley space, we define the *bracket relative to A* by

$$[x_1 \ldots x_k]_A = [x_1 \ldots x_k A].$$

A relative bracket is an alternating k-linear form on a vector space of dimension n. Conversely, any alternating k-linear form on an n-dimensional vector space defines a unique relative bracket. The pair consisting of a vector space V with a relative bracket is a non-standard Cayley Space, called the *contraction* of the standard Cayley Space by the extensor A.

In a non-standard Cayley space on V, a vector in V is said to be of *rank zero* when for all choices of the vectors $x_1, \ldots, x_{k-1}$ in V

$$[x, x_1, \ldots, x_{k-1}] = 0.$$

Otherwise it is said to be of *rank one*.

3. Splits and shuffles

A *split* of the linearly ordered set, or sequence $A = a \ldots bc \ldots de \ldots f$ is a partition of A into blocks which are intervals of A, namely

$$B_1 = (a, \ldots, b), \qquad B_2 = (c, \ldots, d), \ldots, \qquad B_k = (e, \ldots, f).$$

If B_j contains i_j elements for each j we call the split the $(i_1, \ldots, i_k)$-*split* of A.
A *shuffle* of the $(i_1, \ldots, i_k)$-split of A is a permutation $\sigma : A \to \sigma(A)$ of the elements of A with the property that each block of the $(i_1, \ldots, i_k)$-split of $\sigma(A)$ is a subsequence of A. That is, the linear order of A is preserved in each block of $\sigma(A)$.

A *bracket product* is an expression of the form

$$[a_1 \ldots a_n][b_1 \ldots b_n] \ldots [c_1 \ldots c_n] d_1 \vee \ldots \vee d_p$$

for some arbitrary number of brackets.
Let

$$a_1 \ldots a_i \quad b_1 \ldots b_j \ldots c_1 \ldots c_k \quad d_1 \ldots d_m$$

denote a subsequence of the vectors in a bracket product. We define the *split-sum* of their $(i, j, \ldots, k, m)$-split as the expression

$$\sum_\sigma \operatorname{sgn}(\sigma)[\sigma(a_1) \ldots \sigma(a_i) a_{i+1} \ldots a_n][\sigma(b_1) \ldots \sigma(b_j) b_{j+1} \ldots b_n] \ldots$$
$$\times [\sigma(c_1) \ldots \sigma(c_k) c_{k+1} \ldots c_n] \sigma(d_1) \ldots \sigma(d_m) d_{m+1} \ldots d_p$$

where the sum ranges over all shuffles of the above split. Alternatively, we write this as

$$[a_1^\sigma \ldots a_i^\sigma \quad a_{i+1} \ldots a_n][b_1^\sigma \ldots b_j^\sigma \quad b_{j+1} \ldots b_n] \ldots$$
$$[c_1^\sigma \ldots c_k^\sigma \quad c_{k+1} \ldots c_n] \, d_1^\sigma \ldots d_m^\sigma \quad d_{m+1} \ldots d_p.$$

The split-sum is thus formed by applying to the sequence of variables marked by the superscript σ in a bracket product, the shuffles of the split whose blocks are determined by the brackets.

One can iterate a split-sum. When the sets are disjoint, iteration reduces to an interchangeable double summation. In the general case, split-sums are not commutative.

As an example,

$$[a^\theta b^\theta c^\sigma d^\sigma ef][g^\theta h^\sigma i^\sigma jkl]$$

denotes the split-sum of the (2, 1)-split of a, b, g either followed or preceded by the split-sum of the (2, 2)-split of c, d, h, i. However,

$$[a^{\theta\sigma} b^{\theta\sigma} c^\theta def][g^{\theta\sigma} h^\theta i^\theta jkl]$$

denotes the (non-commuting) split-sum of the (3, 3)-split of a, b, c, g, h, i followed by the split-sum of the (2, 1)-split of the sequence $\theta(a), \theta(b), \theta(g)$.

In a single split-sum, we often replace the superscripts by dots. Thus,

$$[a^\sigma b^\sigma cd][e^\sigma fgh] = [\dot{a}\dot{b}cd][\dot{e}fgh].$$

The use of dots to indicate split-sums will be called the *Scottish Convention* after H. W. Turnbull who used it informally.

4. Cayley algebras

We now define a second operation on a Cayley space, called the *meet*. Let $A = a_1 \dots a_k$ and $B = b_1 \dots b_p$ be extensors of indicated steps satisfying $k + p \geq n$. We define their *meet*

$$A \wedge B = a_1 \dots a_k \wedge b_1 \dots b_p$$

by the expression

$$A \wedge B = \sum_{\sigma} \operatorname{sgn}(\sigma)[a_{\sigma(1)}a_{\sigma(2)} \dots a_{\sigma(n-p)}b_1 \dots b_p]a_{\sigma(n-p+1)} \dots a_{\sigma(k)},$$

where σ ranges over all shuffles of the $(n - p, k - n + p)$-split of $a_1 \dots a_k$. Alternatively, we may write this as

$$A \wedge B = [\dot{a}_1 \dots \dot{a}_{(n-p)}b_1 \dots b_p]\dot{a}_{(n-p)+1} \dots \dot{a}_k,$$

where the dots indicate the split-sum of the $(n - p, k - (n - p))$-split of $a_1 \dots a_k$. If $k + p < n$ the meet is defined to be zero and in either case it is extended by linearity to all linear combinations of extensors.

PROPOSITION. *The meet satisfies the identity*

$$[\dot{a}_1 \dots \dot{a}_{(n-p)}b_1 \dots b_p]\dot{a}_{(n-p)+1} \dots \dot{a}_k = \dot{b}_1 \dots \dot{b}_{p-(n-k)}[a_1 \dots a_k\dot{b}_{p-(n-k)+1} \dots \dot{b}_p].$$

The verification is a simple consequence of the alternating property of the bracket.

THEOREM 1. *The meet is associative and anticommutative following the rule*

$$B \wedge A = (-1)^{(n-p)(n-k)}A \wedge B,$$

where A is an extensor of step k, and B of step p.

The verification is a straightforward computation.

The *Cayley algebra* of a Cayley space is the algebraic structure obtained by endowing the exterior algebra with the additional operations of bracket and meet. Thus, a Cayley algebra is the vector space $\bar{V}$ endowed with three operations in the sense of universal algebra: meet, join and bracket.

COROLLARY. *The integral E is an identity for meet in the Cayley algebra, that is,*

$$E \wedge A = A \wedge E = A$$

for all A.

The meet of two extensors has an important geometric interpretation:

PROPOSITION. *If A and B are extensors of step k and p, supporting subspaces $\bar{A}$ and $\bar{B}$ of a standard Cayley space over V, and the span $\bar{A} \cup \bar{B}$ equals V, then the meet $A \wedge B$ supports the intersection $\bar{A} \cap \bar{B}$.*

Proof: Take a basis $e_1, \dots, e_n$ of V such that $e_1, \dots, e_r$ is a basis of $\bar{A} \cap \bar{B}$, $e_1, \dots, e_k$ a basis of $\bar{A}$ and $e_{k+1}, \dots, e_n, e_1, \dots, e_r$ a basis for $\bar{B}$. We may therefore write, for some scalars c and d,

$$A = c\, e_1 \dots e_k$$

$$B = d\, e_1 \dots e_r e_{k+1} \dots e_n.$$

Expanding $A \wedge B$, we get

$$A \wedge B = cd[e_1 \ldots e_n]e_1 \ldots e_r. \qquad \text{q.e.d.}$$

COROLLARY. *The meet of two or more extensors is an extensor.*

5. Duality

Let $A_1, \ldots, A_n$ be extensors of step $n-1$ in V, which we call *covectors*. We define a new alternating multilinear form on covectors in V, called the *double bracket*, by setting

$$[[A_1, \ldots, A_n]] = A_1 \wedge \cdots \wedge A_n.$$

We infer from the properties of meet that the double bracket is non-degenerate and of step zero (that is, a scalar). Thus, since the vector space spanned by covectors is of dimension n, the double bracket defines a Cayley space on covectors. The associated Cayley algebra is called the *dual Cayley algebra*. A Cayley algebra and its dual are isomorphic. The role of join and meet are interchanged under the canonical isomorphism.

A set of covectors $A_1, \ldots, A_n$ with non-zero double bracket constitutes a basis of covectors. In such a case, a corresponding basis of vectors $a_1, \ldots, a_n$ can always be found satisfying

$$A_i = a_1 \ldots \hat{a}_i \ldots a_n,$$

where $\hat{a}_i$ indicates that a_i is deleted. It is verified in Section 7 that

$$[[A_1, \ldots, A_n]] = [a_1, \ldots, a_n]^{n-1},$$

an identity known as Cauchy's theorem on the adjugate. By duality and Cauchy's theorem, we may construct from every identity between joins and meets, another identity where the roles of join and meet are interchanged, step k is replaced by step $n-k$, and suitable powers of the bracket appear as multipliers to restore homogeneity.

For example, if A is an extensor of step k and the b_i are covectors, the identity

$$A \vee (b_1 \wedge \cdots \wedge b_{p+k}) = (\dot{b}_1 \wedge \cdots \wedge \dot{b}_p) \vee (A \wedge \dot{b}_{p+1} \wedge \cdots \wedge \dot{b}_{p+k})$$

is immediate, as it is the dual of the identity

$$B \wedge (a_1 \vee \cdots \vee a_{p+k}) = (\dot{a}_1 \vee \cdots \vee \dot{a}_p) \wedge (B \vee \dot{a}_{p+1} \vee \cdots \vee \dot{a}_{p+k}),$$

where B is an extensor of step $n-k$ and the a_i are vectors.

The *principle of complementary minors* which associates with every identity holding among the minors of a matrix another identity holding among the complementary minors of the adjugate matrix, is a special case of the duality between joins and meets.

By introducing the analogue of the contraction of a bracket by an extensor A, in the dual Cayley algebra, we may construct in the given Cayley algebra the dual operation, called the *co-contraction* or *reduction* by A. Thus, if A is of step k and the x_i are covectors, write A as the meet of $n-k$ covectors and define the reduction by A as

$$[x_1, \ldots, x_k]^A = [[x_1, \ldots, x_k]]_A$$
$$= [[x_1, \ldots, x_k, A]].$$

The notions of contraction and reduction in the Cayley algebra correspond roughly to the meanings these terms have in combinatorial geometry.

6. Identities in the Cayley algebra

We present a sampling of identities which describe how joins and meets are distributed through each other, or *alternative laws.*

We begin with some notation. Juxtaposition of vectors denotes join and juxtaposition of covectors denotes meet. The *inner product* of a vector a and a covector x is defined as

$$\langle a|x\rangle = a \wedge x.$$

Similarly, if extensors $A = a_1 \dots a_k$ and $X = x_1 \wedge \cdots \wedge x_k$ are given, where the a_i are vectors and the x_i are covectors, we define their inner product of length k as

$$\begin{aligned}\langle A|X\rangle &= \langle a_1 \dots a_k|x_1 \dots x_k\rangle \\ &= (a_1 \dots a_k) \wedge (x_1 \dots x_k).\end{aligned}$$

THEOREM 6.1. *Let $a_1, \dots, a_k$ be vectors and $x_1, \dots, x_s$ be covectors. If $k \geq s$, then*

$$(a_1 \dots a_k) \wedge (x_1 \dots x_s) = \langle \dot{a}_1|x_1\rangle \dots \langle \dot{a}_s|x_s\rangle \dot{a}_{s+1} \dots \dot{a}_k;$$

If $k < s$, then

$$(a_k \dots a_1) \vee (x_s \dots x_1) = \dot{x}_s \dots \dot{x}_{k+1}\langle a_k|\dot{x}_k\rangle \dots \langle a_1|\dot{x}_1\rangle.$$

Proof: We verify the first identity. From the definition of meet,

$$\begin{aligned}(a_1 \dots a_k) \wedge (x_1 \dots x_s) &= \langle a_1^\sigma|x_1\rangle (a_2^\sigma \dots a_k^\sigma) \wedge (x_2 \dots x_s) \\ &= \langle a_1^\sigma|x_1\rangle \langle a_2^{\sigma\theta}|x_2\rangle (a_3^{\sigma\theta} \dots a_k^{\sigma\theta}) \wedge (x_3 \dots x_s) \\ &\ \ \vdots\end{aligned}$$

Here σ ranges over the split-sum of the $(1, k-1)$-split of $a_1 \dots a_k$, θ ranges over the split-sum of the $(1, k-2)$-split of $a_{\sigma(2)} \dots a_{\sigma(k)}$, and so forth. But by an elementary coset argument this is equal to the split-sum of the $(1, \dots, 1, k-s)$-split of $a_1 \dots a_k$.

THEOREM 6.2. *Let $a_1, \dots, a_k$ be vectors and $x_1, \dots, x_s$ be covectors. If $k \geq s$, then*

$$\begin{aligned}(a_1 \dots a_k) \wedge (x_1 \dots x_s) = \langle \dot{a}_1 \dots \dot{a}_i|x_1 \dots x_i\rangle \dots \\ \times \langle \dot{a}_{i+\cdots+j+1} \dots \dot{a}_s|x_{i+\cdots+j+1} \dots x_s\rangle \dot{a}_{s+1} \dots \dot{a}_k.\end{aligned}$$

If $k < s$, then

$$\begin{aligned}(a_k \dots a_1) \vee (x_s \dots x_1) = \dot{x}_s \dots \dot{x}_{k+1}\langle a_k \dots a_{i+\cdots+j+1}|\dot{x}_k \dots \dot{x}_{i+\cdots+j+1}\rangle \dots \\ \langle a_i \dots a_1|\dot{x}_i \dots \dot{x}_1\rangle\end{aligned}$$

Proof: By the associative law,

$$(a_1 \dots a_k) \wedge (x_1 \dots x_s) = (a_1 \dots a_k) \wedge (x_1 \dots x_i) \wedge \cdots \wedge (x_{i+\cdots+j+1} \dots x_s)$$

whence proceed as in Theorem 6.1. The second expression is derived similarly.

COROLLARY 1. *Let $C_1, \ldots, C_s$ be extensors of step $n - i, \ldots, n - j, n - l$ and let $k = i + \cdots + j + l$. Then*

$$a_1 \ldots a_k \wedge (C_1 \wedge \cdots \wedge C_s) = [\dot a_1 \ldots \dot a_i C_1] \ldots [\dot a_{i+\cdots+j+1} \ldots \dot a_k C_s]$$

If $A = a_1 \ldots a_k$ and $X = x_1 \ldots x_s$ are vector and covector decompositions of flats we shall sometimes employ the notation

$$\dot A = \dot a_1 \ldots \dot a_k \qquad \text{and} \qquad \dot X = \dot x_1 \ldots \dot x_s.$$

COROLLARY 2. *Let A_j and X_j be extensors of complementary step for each j. Then*

$$(A_1 \vee \cdots \vee A_{k+1}) \wedge (X_1 \wedge \cdots \wedge X_k) = \langle \dot A_1 | X_1 \rangle \ldots \langle \dot A_k | X_k \rangle \dot A_{k+1}$$

$$(A_k \vee \cdots \vee A_1) \vee (X_{k+1} \wedge \cdots \wedge X_1) = \dot X_{k+1} \langle A_k | \dot X_k \rangle \ldots \langle A_1 | \dot X_1 \rangle$$

THEOREM 6.3. *Let A_k, B_l, C_p, D_q, $X_{n-(k+l)}$, and $Y_{n-(p+q)}$ be extensors of indicated steps. Then*

$$(A \vee X) \wedge BC \wedge (D \vee Y) = \pm((A \vee \dot B) \wedge X) \vee ((\dot C \vee D) \wedge Y).$$

Proof:

$$\begin{aligned} (A \vee X) \wedge BC \wedge (D \vee Y) &= \pm[A\dot B X]\dot C \wedge (D \vee Y) \\ &= \pm((A \vee \dot B) \wedge X)\dot C \wedge (D \vee Y) \\ &= \pm((A \vee \dot B) \wedge X)[\dot C D Y] \\ &= \pm((A \vee \dot B) \wedge X) \vee ((\dot C \vee D) \wedge Y). \end{aligned}$$

THEOREM 6.4. *If $A \vee B$ is of step n, then*

$$A \vee B = (A \wedge B) \vee E$$

where E is the integral.

The proof is a simple verification.

We now present the main result of this section.

THEOREM 6.5. *Let $C^{(1)}, \ldots, C^{(r)}$ be extensors of step $n - q_1, \ldots, n - q_r$ and let $k + s = q_1 + \cdots + q_r$. Then*

$$(a_1 \ldots a_k b_1 \ldots b_s) \wedge C^{(1)} \wedge \cdots \wedge C^{(r)} = (\dot b_1 \ldots \dot b_s) \vee \sum_{i_1 + \cdots + i_r = s} (-)^{\iota(i_1, \ldots, i_r)}$$
$$\times \{\dot a_1 \ldots \dot a_{q_1 - i_1} C^{(1)}\} \wedge \cdots \wedge \{\dot a_{[q_1 - i_1 + \cdots + q_{r-1} - i_{r-1}]+1} \ldots \dot a_k C^{(r)}\}$$

where the integer $\iota(i_1, \ldots, i_r)$ is specified below.

Proof: For simplicity of notation take $s < q_r$. By Theorem 6.2, we have, calling the left side I,

$$I = [\dot a_1 \ldots \dot a_{q_1} C^{(1)}][\dot a_{q_1+1} \ldots \dot a_{q_1+q_2} C^{(2)}] \ldots [\dot a_{[q_1+\cdots+q_{r-1}]+1} \ldots \dot a_k \dot b_1 \ldots \dot b_s C^{(r)}]$$

The permutations acting in this equation may be separated into classes according to their effect on the b's. Thus, a given permutation positions, say, i_1 of the b's in the bracket containing $C^{(1)}, \ldots, i_r$ of the b's in the bracket containing $C^{(r)}$. Affixing the appropriate sign,

$$I = \sum_{i_1 + \cdots + i_r = s} (-)^{(i_1,\ldots,i_r)}[b_1^\theta \ldots b_{i_1}^\theta a_1^\sigma \ldots a_{q_1 - i_1}^\sigma C^{(1)}]$$
$$\times [b_{i_1+1}^\theta \ldots b_{[i_1+i_2]}^\theta a_{[q_1-i_1]+1}^\sigma \ldots a_{[q_1-i_1+q_2-i_2]}^\sigma C^{(2)}] \ldots$$
$$\times [b_{[i_1+\cdots+i_{r-1}]+1}^\theta \ldots b_s^\theta a_{[q_1-i_1+\cdots+q_{r-1}-i_{r-1}]+1}^\sigma \ldots a_k C^{(r)}].$$

Here θ ranges over the split-sum of the $(i_1, \ldots, i_r)$-split of $b_1 \ldots b_s$, and σ ranges over the split-sum of the $(q_1 - i_1, \ldots, q_r - i_r)$-split of $a_1 \ldots a_k$.

We first evaluate $(i_1, \ldots, i_r)$:

$$\begin{aligned}(i_1, \ldots, i_r) &= i_1(q_1 + \cdots + q_r - s) + i_2(q_2 + \cdots + q_r - s + i_1) + \cdots + i_r(q_r - s + i_1 \ldots + i_{r-1}) \\ &= q_1(i_1) + q_2(i_1 + i_2) + \cdots + q_r(i_1 + \cdots + i_r) - s(i_1 + \cdots + i_r) \\ &\quad + i_2(i_1) + i_3(i_2 + i_1) + \cdots + i_r(i_{r-1} + \cdots + i_1) \\ &= q_1(i_1) + q_2(i_1 + i_2) + \cdots + q_r(i_1 + \cdots + i_r) + h_2(i_1 \ldots i_r)\end{aligned}$$

where $h_2(i_1 \ldots i_r)$ is the homogeneous symmetric function of degree two on $i_1, \ldots, i_r$.

We now factor out the b's using Theorem 6.2. This gives the desired identity.

We conclude this Section with two examples which illustrate the correspondence of theorems of projective geometry with identities in Cayley algebras.

Desargues's Theorem. *The corresponding sides of two collinear triangles intersect in collinear points if and only if the joins of corresponding vertices are concurrent.*

Proof: Let a, b, c be vectors and x, y, z be covectors in a Cayley space of three dimensions. Juxtaposition of vectors denotes join and juxtaposition of covectors denotes meet. The identity

$$abc \wedge [(a \vee yz) \wedge (b \vee zx) \wedge (c \vee xy)] = xyz \wedge [(bc \wedge x) \vee (ca \wedge y) \vee (ab \wedge z)]$$

is easily verified. Now let $x = b'c'$, $y = c'a'$, $z = a'b'$ so that $xyz = [a'b'c']^2$. This gives

$$[(bc \wedge b'c') \vee (ca \wedge c'a') \vee (ab \wedge a'b')] = [(aa') \wedge (bb') \wedge (cc')][abc][a'b'c'].$$

Desargues's theorem for triangles whose vertices are a, b, c and a', b', c' is then the statement that one side of this identity is zero if and only if the other side is zero.

Pappus' Theorem. *If a, b, c are collinear, and a', b', c' are collinear and if all six points are distinct, then $ab' \wedge a'b$, $bc' \wedge b'c$, and $ca' \wedge c'a$ are also collinear.*

Proof: The theorem is a restatement of the identity

$$\begin{aligned}&(bc' \wedge b'c) \vee (ca' \wedge c'a) \vee (ab' \wedge a'b) \\ &\quad = [aa'b'][bb'c'][cc'a'][abc] - [abb'][bcc'][caa'][a'b'c'].\end{aligned}$$

Note that the algebraic version of each of these theorems is the stronger one, as it includes the geometric result as well as degeneracies.

7. Determinant identities

Identities between minors of matrices find elegant verification in the language of Cayley algebras. We illustrate with some examples.

Let $\{e_1, \ldots, e_n\}$ be a unimodular basis of vectors. With it we associate a basis of covectors $\{1, \ldots, n\}$ by setting $j = x_j = e_1 \ldots \hat{e}_j \ldots e_n$.* Thus any extensor A of step k may be uniquely expressed as a linear combination of monomials of the form $i_1 \ldots i_{n-k}$, where $i_1 < \cdots < i_{n-k} \in I$ and juxtaposition indicates meet. It is easily verified that $\{1, \ldots, n\}$ is also unimodular, that is, that $1 \ldots n$ is equal to unity.

Given an extensor $A = a_1 \ldots a_n$ of step n we may re-express its determinant $[a_1 \ldots a_n]$ in coordinate form by applying the alternative laws to $A \wedge 1 \ldots n$:

$$A \wedge 1 \ldots n = [a_1 \ldots a_n] = \langle a_1 | i \rangle \ldots \langle a_n | n \rangle$$

where $\langle a_i | j \rangle = a_i \wedge j$ is the j-th coordinate of a_i relative to $e_1, e_2, \ldots, e_n$.

A similar procedure may be used to coordinatize a flat of any step. Thus, if A is of step k we may write

$$A = A \wedge e_1 \ldots e_n = \dot{e}_1 \ldots \dot{e}_k (A \wedge \dot{e}_{k+1} \ldots \dot{e}_n)$$

or

$$A = A \vee \dot{1} \ldots n = \dot{x}_1 \dot{x}_2 \ldots \dot{x}_{n-k} (A \wedge \dot{x}_{n-k+1} \ldots \dot{x}_n).$$

The first expansion represents a covariant coordinatization while the second represents the associated contravariant coordinatization. The numerical coefficients occurring in these expansions are the well known Plücker coordinates of the flat relative to the indicated basis.

Given a determinant $\Delta = [a_1, a_2, \ldots, a_n]$, the *adjugate* of Δ is the determinant

$$\Delta^* = \bar{a}_n \wedge \bar{a}_{n-1} \wedge \cdots \wedge \bar{a}_1$$

where $\bar{a}_i = a_1 \ldots \hat{a}_i \ldots a_n$. The adjugate is thus the determinant of $(n-1) \times (n-1)$ minors of Δ. Many determinant identities describe the relationships between these two determinants.

We begin with the expansion of Δ due to Laplace.

(1) *The Laplace expansion*. This describes how to expand Δ in terms of the set of minors of Δ in a given subset of $\{a_1, \ldots, a_n\}$. Thus by Theorem 6.2,

$$\begin{aligned} \Delta &= a_1 \ldots a_n \wedge 1 \ldots n \\ &= (a_1 \ldots a_n) \wedge (1 \ldots k) \wedge (k+1 \ldots n) \\ &= \langle a_1 \ldots a_k | 1 \ldots k \rangle \langle a_{k+1} \ldots a_n | k+1 \ldots n \rangle. \end{aligned}$$

The Laplace expansion is thus a consequence of one of the alternative laws.

(2) *Cauchy's Theorem on the adjugate*: The adjugate is the $(n-1)$th power of the original determinant. By the associative law for meet,

$$\begin{aligned} \Delta^* &= (-)^{n(n-1)/2} \bar{a}_1 \wedge \cdots \wedge \bar{a}_n \\ &= (-)^{n(n-1)/2} (a_2 \ldots a_n) \wedge (a_1 a_3 \ldots a_n) \wedge \cdots \wedge (a_1 \ldots a_{n-1}) \\ &= -(-)^{n(n-1)/2} [a_1 \ldots a_n] (a_3 \ldots a_n) \wedge (a_1 a_2 a_4 \ldots a_n) \wedge \cdots \wedge (a_1 \ldots a_{n-1}) \\ &\;\;\vdots \\ &= \Delta^{n-1}. \end{aligned}$$

* Note our unconventional usage of integers as variables.

(3) *Jacobi's Theorem on the adjugate*: A minor of order r of the adjugate is equal to the complementary minor in the original determinant multiplied by the $(n-r-1)$th power of Δ.

We illustrate with the case $r = 2$. Consider the identity

$$\begin{aligned}\bar a_3 \wedge \cdots \wedge \bar a_n &= a_1a_2\hat a_3\ldots a_n \wedge a_1a_2a_3\hat a_4\ldots a_n \wedge \cdots \wedge a_1a_2\ldots \hat a_n\\ &= (-)^{(n-2)(n-3)/2}a_1a_2[a_1a_2\hat a_3\ldots a_n]\ldots[a_1\ldots a_{n-1}\hat a_n]\\ &= (-)^{(n-2)(n-3)/2}a_1a_2\Delta^{n-3}.\end{aligned}$$

Now meet both sides with

$$ij = e_1\ldots\hat e_i\ldots\hat e_j\ldots e_n.$$

This gives

$$\langle \bar a_n\ldots\bar a_3|e_1\ldots\hat e_i\ldots\hat e_j\ldots e_n\rangle = \Delta^{n-3}\langle a_1a_2|ij\rangle,$$

which is the desired result.

(4) *The Bazin–Reiss–Picquet Identity*: Starting with Cauchy's theorem on the adjugate, meet both sides with $a\ \ b\ldots c\ \ a_{k+1}\ldots a_n$. This gives

$$\begin{aligned}&[a\ \ b\ldots c\ \ a_{k+1}\ldots a_n][a_1\ldots a_n]^{n-1}\\ &\quad= [\dot a\ \ a_2\ldots a_n][a_1\dot b\ \ a_3\ldots a_n][a_1a_2\ldots\dot c\ \ a_{k+1}\ldots a_n][a_1\ldots\dot a_{k+1}\ldots a_n]\ldots[a_1\ldots\dot a_n]\end{aligned}$$

so that

$$[a\ \ b\ldots c\ \ a_{k+1}\ldots a_n]\,\Delta^{k-1} = [\dot a a_2\ldots a_n][a_1\dot b\ldots a_n]\ldots[a_1\ldots\dot c\ldots a_n],$$

as desired.

(5) *Sylvester's Theorem on Compound Determinants*: Form the set of monomials $a_{i_1}\ldots a_{i_k}$ where $i_1 < \cdots < i_k$ from the sequence $\{a_1,\ldots,a_n\}$ and order them lexicographically as $\{A_1,\ldots,A_{\binom{n}{k}}\}$. Also, let the set $\{X_1,\ldots,X_{\binom{n}{k}}\}$ be formed from the set $\{1,\ldots,n\}$ of covectors, in the same way. The determinant

$$\Delta_k = \langle \dot A_1|\dot X_1\rangle\ldots\langle A_{\binom{n}{k}}|\dot X_{\binom{n}{k}}\rangle$$

is called the k-th compound of Δ. Sylvester's theorem states that $\Delta_k = \Delta^{\binom{n-1}{k-1}}$.

We illustrate the method for the case $n = 4$ and $k = 2$, so that $\binom{n-1}{k-1} = 3$. By Cauchy's theorem,

$$\begin{aligned}[abcd]^3 &= (abc) \wedge (abd) \wedge (acd) \wedge (bcd)\\ &= (ab\ \ [acbd]) \wedge ([adbc]\ \ cd)\\ &= (ab \vee (ac \wedge bd)) \wedge ((ad \wedge bc) \vee cd).\end{aligned}$$

Similarly,

$$[1234]^3 = (12 \vee (13 \wedge 24)) \wedge ((14 \wedge 23) \vee 34).$$

Now substitute for $[abcd]^3$ and $[1234]^3$ on the left hand side of

$$[abcd]^3 \vee [1234]^3 = [abcd]^3$$

and expand the resulting expression by the alternative laws. This gives the result.

Sylvester's identity shows how to construct a Cayley space on the extensors of step k.

8. The Straightening Formula

We now derive the basic result of the theory of Cayley algebras. In its simplest form, it can be viewed as stating that a set of vectors is a basis of a certain vector space. It can also be interpreted as the solution to a word problem in the Cayley algebra, (see Section 12).

Our main application of the Straightening Formula is a characteristic-free proof of the First Fundamental Theorem of invariant theory. We also sketch applications to the classification of identities in associative algebras and to the theory of symmetric functions.

Some of the results below can be extended to spaces of arbitrary dimensions, but we have preferred to preserve the more elegant approach by Cayley algebras. The finite-dimensional case proved here is actually the stronger.

Let K be a field of arbitrary characteristic and let R_K be the polynomial ring over K obtained by adjoining mn transcendentals $(a_i|x_j)$ where $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$.

Let $\alpha = (\alpha_1, \ldots, \alpha_m)$ and $\beta = (\beta_1, \ldots, \beta_n)$ be sequences of non-negative integers. We define

$$V_{\alpha,\beta}$$

to be the vector space over K spanned by all monomials in the $(a_i|x_j)$ which contain α_i occurences of a_i and β_j occurrences of x_j, or all monomials of *content* (α, β) for short.

A *double tableau* of content (α, β) is denoted by the double matrix

$$T = \left(\begin{array}{ccc|ccc} a_{11} & \cdots & a_{1\lambda_1} & x_{11} & \cdots & x_{1\lambda_1} \\ \vdots & & & \vdots & & \\ a_{s1} & \cdots & a_{s\lambda_s} & x_{s1} & \cdots & x_{s\lambda_s} \end{array}\right)$$

where $n \geq \lambda_1 \geq \cdots \geq \lambda_s$ and where the elements a_{ij} of the left tableau are chosen from $\{a_1, \ldots, a_m\}$ and the elements x_{ij} of the right tableau are chosen from $\{x_1, \ldots, x_n\}$, such that each a_i occurs with multiplicity α_i and each x_j occurs with multiplicity β_j.

The tableau T is defined to be the expression

$$T = (a_{11} \ldots a_{1\lambda_1}|x_{11} \ldots x_{1\lambda_1}) \ldots (a_{s1} \ldots a_{s\lambda_s}|x_{s1} \ldots x_{s\lambda_s}),$$

where we set

$$(a_{j1} \ldots a_{j\lambda_j}|x_{j1} \ldots x_{j\lambda_j}) = \sum_{\sigma} \operatorname{sgn}(\sigma)(a_{j1}|x_{j\sigma(1)}) \ldots (a_{j\lambda_j}|x_{j\sigma(\lambda_j)}),$$

the above sum extending over all permutations σ of the sequence $1, \ldots, \lambda_j$.

Assign to the a_i and x_j the linear orderings

$$a_1 < \cdots < a_m \quad \text{and} \quad x_1 < \cdots < x_n.$$

Relative to these orderings, a double tableau is said to be *standard* when in each tableau the entries in each row are increasing from left to right and the entries in each column are non-decreasing downward.

The *shape* of a double tableau T is the row length vector

$$\lambda[T] = (\lambda_1, \dots, \lambda_s).$$

Shapes of tableaux are ordered lexicographically by $\lambda > \mu$ when $\lambda_i > \mu_i$ and $\lambda_j = \mu_j$ for $j < i$.

Using this ordering on shapes we now linearly order all tableaux. Associate with T the sequence

$$\pi[T] = a_{11} \dots a_{1\lambda_1} a_{21} \dots a_{s\lambda_s} x_{11} \dots x_{s\lambda_s},$$

and order the set of these sequences lexicographically.

If S denotes another double tableau then set $T > S$ if $\lambda[T] > \lambda[S]$, or if $\lambda[T] = \lambda[S]$ and $\pi[T] < \pi[S]$.

Remark: Identities in a Cayley algebra between inner products may be interpreted in R_K. To do this, substitute for each inner product $\langle a_{i_1} \dots a_{i_k} | x_{j_1} \dots x_{j_k} \rangle$ the double tableau $(a_{i_1} \dots a_{i_k} | x_{j_1} \dots x_{j_k})$. Conversely, *any* identity in R_K may be interpreted in a Cayley algebra over the integral domain R_K, and we shall use the two notations interchangeably.

LEMMA 1. *Let $k \geq l$ and*

$$B = b_1 \dots b_{j-1} \qquad Y = y_1 \dots y_k$$
$$C = c_{j+1} \dots c_l \qquad Z = z_1 \dots z_l$$

where the b_i and c_i are vectors taken from the set $\{a_1 \dots a_m\}$ and the y_i and z_i are covectors from $\{x_1 \dots x_n\}$. Then the expression

$$I = \langle B\dot{b}_j \dots \dot{b}_k | Y \rangle \langle \dot{c}_1 \dots \dot{c}_j C | Z \rangle$$

is equal to a sum of products of pairs of inner products, each pair containing one inner product of length at least $k + 1$.

Proof: By Theorem 6.3 we have

$$I = (\pm)(B \vee Y) \wedge (b_j \dots b_k \quad c_1 \dots c_j) \wedge (C \vee Z).$$

Setting $b_j \dots b_k \quad c_1 \dots c_j = D$, we now use Theorem 6.5 to distribute B through the other factors. This gives

$$I = (\pm) Y \wedge \sum_{s=0}^{j-1} (-)^{(s)} \{(b_1^\sigma \dots b_s^\sigma) \vee D\} \wedge \{(b_{s+1}^\sigma \dots b_{j-1}^\sigma) \vee (C \vee Z)\}$$

Distributing Z by the dual of Corollary 1 to Theorem 6.2, this becomes

$$I = (\pm) \sum_{s=0}^{j-1} (-)^{(s)}$$

$$\{(b_1^\sigma \dots b_s^\sigma D) \wedge Y \wedge (\dot{z}_1 \dots \dot{z}_{s+1})\} \vee \{(b_{s+1}^\sigma \dots b_{j-1}^\sigma C) \wedge (\dot{z}_{s+2} \dots \dot{z}_l)\},$$

or

$$I = (\pm) \sum_{s=0}^{j-1} (-)^{(s)} \begin{pmatrix} b_1^\sigma \dots b_s^\sigma D & Y \dot{z}_1 \dots \dot{z}_{s+1} \\ b_{s+1}^\sigma \dots b_{j-1}^\sigma C & \dot{z}_{s+2} \dots \dot{z}_l \end{pmatrix},$$

which concludes the proof.

THEOREM 1. *(Straightening Formula) The double standard tableaux of content (α, β) span $V_{\alpha\beta}$.*

Proof: Any monomial of step zero equals a linear combination of monomials of the form

$$\langle a|x\rangle\langle b|y\rangle\ldots\langle c|z\rangle = \begin{pmatrix} a & x \\ b & y \\ \vdots & \vdots \\ \vdots & \vdots \\ c & z \end{pmatrix}$$

We show that any double tableau equals a linear combination of double standard tableux. We proceed by induction on the linear ordering of tableaux, and show that every non-standard double tableau T of content (α, β) equals a linear combination of greater tableaux of content (α, β). Since there are only finitely many double tableaux of content (α, β) iteration of this argument must then eventually express T as a linear combination of double standard tableaux.

If two entries in T satisfy $t_{ij} \geq t_{i,j+1}$ or $t_{ij} > t_{i+1,j}$ call this a *violation* of standard form in T

Assume a violation occurs in the left tableau. If it is a row violation, $a_{ij} > a_{i,j+1}$ then set $T \to -S$ where S is obtained by reversing the positions of a_{ij} and $a_{i,j+1}$ in T. Note that $\pi[T] > \pi[S]$ so that $S > T$.

Now assume a column violation $a_{ij} > a_{i+1,j}$ occurs.

Let T_1 denote the first $i-1$ rows of T, T_2 denote the next two rows of T, and T_3 denote the remaining rows. We are primarily concerned with T_2, which we display as

$$T_2 = \begin{pmatrix} B & b_j \ldots b_k & Y \\ c_1 \ldots c_j & C & Z \end{pmatrix},$$

where

$$B = b_1 \ldots b_{j-1} \quad Y = y_1 \ldots y_k$$

$$C = c_{j+1} \ldots c_l \quad Z = z_1 \ldots z_l.$$

Consider the expression

$$I = \begin{pmatrix} B & \dot{b}_j \ldots \dot{b}_k & Y \\ \dot{c}_1 \ldots \dot{c}_j & C & Z \end{pmatrix}.$$

Since any indicated permutation σ, except the identity, exchanges elements from the first row of I with elements from the second row, and since

$$c_1 < \cdots < c_j < b_j < \cdots < b_k,$$

it must be true that $\sigma(c_j) > \sigma(b_j)$. Thus we have that

$$I = T_2 + \sum_{S > T_2} c(S)S$$

where $c(S)$ are integers. By Lemma 1 we also have

$$I = \sum_{Q > T_2} c(Q)Q$$

Combining these results gives

$$T_2 = \sum_{Q > T_2} c(Q)Q + \sum_{S > T_2} c(S)S$$

which expresses T_2 as a linear combination of greater tableaux. Appending this expression for T_2 to T yields an expression for T as a linear combination of greater tableaux. Similarly, if violations occur in the right tableau of T, they may be straightened by an analogous procedure.

This completes the proof.

In the course of the proof the following result has been implicitly established:

COROLLARY. *Let P and Q be elements of $V_{\alpha\beta}$, and let*

$$P = (a_{i_1} \cdots a_{i_s} | x_{j_1} \cdots x_{j_s})Q.$$

Then P equals a linear combination with integer coefficients of double standard tableaux, whose first rows are of length s or greater.

Theorem 1 has an interpretation in a Cayley algebra over K.

THEOREM 2 (*Straightening Formula for Cayley Algebras*). *Any monomial of content (α, β) of step zero in the vectors a_i and the covectors x_j, built out of joins and meets in the Cayley algebra of a vector space of dimension d equals a linear combination with integer coefficients of double standard tableaux of content (α, β), whose rows are of length at most d.*

We next establish the linear independence of the double standard tableaux, using a new kind of polarization. We begin with some definitions.

The *set-polarization operator*

$$D^k(b, a) = D^k_{ba}$$

acts on a monomial in $V_{\alpha\beta}$ by replacing it by the sum of the monomials obtained by replacing in turn every subset of k entries equal to a by a subset of k entries equal to b. If the given monomial has p occurrences of the symbol a, then the result of applying the operator D^k_{ba} is the sum of $\binom{p}{k}$ terms. If the monomial has fewer than k occurrences of the symbol a, the result is 0. For $k = 1$ the operator $D^1(b, a)$ is the classical polarization operator.

The *substitution operator*

$$S(b, a) = S_{ba}$$

acts on monomials in $V_{\alpha\beta}$ by replacing each occurrence of the symbol a by an occurrence of the symbol b.

Now extend set-polarization and substitution to all of $V_{\alpha,\beta}$ by linearity.

The following combinatorial lemma is easily proven by the pigeonhole principle:

LEMMA 1. *Let S and T be single tableaux of the same content with $\lambda[S] \leq \lambda[T]$. If in each tableau the entries in each row are strictly increasing, then one of two alternatives occurs:*

(1) *S and T are of the same shape, and the entries in each column of T are obtained by permuting the entries in the corresponding column of S, or*

(2) *Some row of T contains at least two entries which appear in the same column of S.*

We are now ready to prove the linear independence of the double standard tableaux in R_K.

THEOREM 3. *The double standard tableaux of content* (α, β) *form a basis for* $V_{\alpha\beta}$.

Proof: It suffices to produce for any double standard tableau $\{T_1|T_2\}$ a linear transformation $P(T_1|T_2)$ from $V_{\alpha\beta}$ to some vector space satisfying

$$(*) \qquad P(T_1|T_2)\{T_1|T_2\} = w$$
$$P(T_1|T_2)\{D_1|D_2\} = 0$$

for $w \neq 0$ and where $\{D_1|D_2\}$ is any other double standard tableau $\{D_1|D_2\}$ of shape $\geq \lambda$, where $\lambda =$ shape of $\{T_1|T_2\}$. For then, if the double standard tableaux were not independent, there would be a non trivial linear combination $\mathscr{L}$ of double standard tableaux equalling zero, and if we were to take a tableau $\{T_1|T_2\}$ of least shape with non zero coefficient in $\mathscr{L}$ (say the coefficient of $\{T_1|T_2\}$ is d), then applying $P(T_1|T_2)$ to $\mathscr{L}$ would yield $d \cdot w = 0$ which is impossible since $d \neq 0$ and $w \neq 0$. Hence the double standard tableaux would have to be independent.

Let M_K be the polynomial ring over K obtained by adjoining transcendentals $(s_{ij}|t_{kl})$ and $(b_p|y_q)$ where indices range over finite sets of sufficient size to perform the following constructions. Let W denote the vector space with the $(s_{ij}|t_{kl})$ and $(b_p|y_q)$ as a basis.

In the double tableau $\{T_1|T_2\}$ let α_{ij} be the number of entries equal to a_i in column j of T_1 and let β_{ij} be the number of entries equal to x_i in column j of T_2. Set

$$D(T_1|T_2) = \prod_{i,j} D^{\alpha_{ij}}(s_{ij}, a_i) \prod_{i,j} D^{\beta_{ij}}(t_{ij}, x_i)$$

Now let

$$S(T_1|T_2) = \prod_{ij} S(b_j, s_{ij}) \prod_{ij} S(y_j, t_{ij}).$$

By the above definitions, the operator

$$P(T_1|T_2) = S(T_1|T_2)D(T_1|T_2)$$

is a linear operator which maps $V_{\alpha,\beta}$ into W.

To see that $P(T_1|T_2)$ satisfies (*), we begin by computing $D(T_1|T_2)\{T_1|T_2\}$. This is a sum of the form

$$D(T_1|T_2)\{T_1|T_2\} = \{T_1'|T_2'\} + \sum \{V_1|V_2\}$$

where $\{T_1'|T_2'\}$ is obtained by replacing the α_{ij} entries in the j-th column of T_1 which are equal to a_i by s_{ij} and simultaneously replacing the β_{ij} entries in the j-th column of T_2 which are equal to x_i by t_{ij}. Each term $\{V_1|V_2\}$ has the property that it may not be obtained from $\{T_1'|T_2'\}$ by permuting the elements within a column. We claim that

$$P(T_1|T_2)\{T_1|T_2\} = \{T_1''|T_2''\} \neq 0.$$

where all entries in the jth column of T_1'' or T_2'' equals b_j or y_j, respectively. Clearly $\{T_1''|T_2''\}$ is one term in $P(T_1|T_2)\{T_1|T_2\}$ since it is the image of $\{T_1'|T_2'\}$ under $S(T_1|T_2)$. But by the above property of the other terms $\{V_1|V_2\}$, and since $D(T_1|T_2)$ preserves the shape of a double tableau, we have by the preceding lemma that $S(T_1|T_2)\{V_1|V_2\} = 0$.

Now consider any other double standard tableau $\{G_1|G_2\}$ of shape $\geq \lambda$. $D(T_1|T_2)\{G_1|G_2\}$ is a sum of terms $\{Y_1|Y_2\}$ of shape $\geq \lambda$ which are not equal to $\{T_1'|T_2'\}$ or obtained from it by rearranging elements within a column. Hence by the Lemma $P(T_1|T_2)\{G_1|G_2\} = 0$. This completes the proof.

The Straightening Formula for R_K states that $V_{\alpha\beta}$ has two bases, the monomials of content (α, β) and the double standard tableaux of content (α, β). This result can be related to an identity in the theory of the symmetric group.

Let $M(\alpha, \beta)$ be the dimension of $V_{\alpha,\beta}$, and note that this number equals the number of matrices with non-negative integer entries and with row sums $(\alpha_1, \alpha_2, \ldots)$ and column sums $(\beta_1, \beta_2, \ldots)$. Let $K(\alpha, \lambda)$ be the number of *single* standard tableaux of content α and shape λ. Then the above yields the identity

$$M(\alpha, \beta) = \sum_{\lambda} K(\alpha, \lambda) K(\beta, \lambda),$$

as λ ranges over all partitions of the integer n.

We now extend the linear independence of the standard tableaux to a more general ring. We begin by motivating our construction with imprecise but, we hope, suggestive language. In a vector space of dimension d, monomials in the inner products $\langle a_i | x_j \rangle$ are not always linearly independent. This leads to constructing a homomorphic image of R_K which is isomorphic with the ring of inner products of vectors and covectors in dimension d.

Consider the ideal J_d in R_K generated by the elements

$$\det_{\substack{i \in I \\ k \in K}} (a_i | x_k)$$

as I and K range over all subsets of $d+1$ elements of $\{1, \ldots, m\}$ and $\{1, \ldots, n\}$, where d is a given integer.

The ideal J_d is invariant under permutation of the variables a_i and x_j. Furthermore, every double tableau having one row longer than d belongs to J_d. By Theorem 3, these double standard tableaux are independent, and by the Corollary to Theorem 2 every element of J_d equals a linear combination of double standard tableaux each of which has a row longer than d. Concluding we have proved the

LEMMA. *The ideal J_d has a basis consisting of all double standard tableaux in the entries a_i and x_j having at least one row of length greater than d.*

We can now state the main result of this Section:

THEOREM 4. *In the quotient ring $G_d(K)$ the double standard tableaux whose rows are of length at most d form an integral basis.*

Proof: By the preceding lemma, taking the quotient by the ideal J_d amounts to setting to zero all double standard tableaux having one row longer than d, and only these. Hence, the conclusion follows from Theorem 3.

Finally we note the remarkable fact that by Theorem 4, even though monomials in the $(a_i | x_j)$ are not independent, nevertheless the double standard tableaux are.

9. The First Fundamental Theorem

We now apply the Straightening Formula to derive the main results on vector invariants over arbitrary fields. The technique is simpler than the ones classically used, which apply only to fields of characteristic zero.

Let V be an n-dimensional vector space over a field K, and let

$$F(x_1, \ldots, x_N) = F(x_1, \ldots, x_N : e_1, \ldots, e_n)$$

be a polynomial function of the coordinates of the vectors $x_1, \ldots, x_N$ relative to the basis of covectors $e_1, \ldots, e_n$. Since the jth coordinate of the vector x_i may be written as

$$\pm x_{ij} = x_i \wedge e_j = \langle x_i | e_j \rangle = \langle x_i | j \rangle,$$

the function $F(x_1, \ldots, x_N)$ equals a linear combination of double tableaux in the vectors x_i and the covectors e_j.

A polynomial is *invariant* when for every non-singular linear transformation T on V,

$$F(Tx_1, \ldots, Tx_N; e_i, \ldots, e_n) = \lambda(T) F(x_1, \ldots, x_N; e_1, \ldots, e_n)$$

where $\lambda(T)$ is some scalar function.

Since T induces through its adjoint T^*, a non-singular linear transformation on covectors satisfying

$$\langle Tx_i | e_j \rangle = \langle x_i | T^* e_j \rangle,$$

and since F depends only on the $\langle x_i | e_j \rangle$, we may alternately define an invariant as a polynomial which satisfies

$$F(x_1, \ldots, x_N; T^*e_1, \ldots, T^*e_n) = \mu(T^*) F(x_1, \ldots, x_N; e_1, \ldots, e_n)$$

for all non-singular linear transformations T^* acting on covectors.

We also define a *formal invariant* as a polynomial $F(x_1, \ldots, x_N)$ which is an invariant when considered over the extension field $K(x_{11}, \ldots, x_{Nn})$, where K is the ground field of V and the coordinates x_{ij} are transcendentals.

We shall prove the following result over an arbitrary field.

THEOREM 1. *Every invariant (or formal invariant when the field K is finite) in the vectors $x_1, \ldots, x_N$ is expressible as a linear combination of products of brackets in the x_i, where each summand has the same number of bracket factors. In other words, every invariant is a word in the Cayley algebra, built out of joins and meets of $x_1, \ldots, x_N$ alone with no explicit reference to $e_1, \ldots, e_n$, in which every summand is of the same total degree.*

Proof. As noted F may be written as a linear combination of double tableaux, and thus, by the Straightening Formula, as a linear combination $\mathscr{L} = \sum_i \lambda_i (C_i | D_i)$ of double standard tableaux. We must therefore show that the right tableau of each summand in $\mathscr{L}$ is given by writing j in place of e_j

$$D = \begin{pmatrix} 12 & \ldots & n \\ & \vdots & \\ 12 & \ldots & n \end{pmatrix}$$

where D has (say) g rows.

We begin by showing that each right tableau in $\mathscr{L}$ contains each variable $e_1, e_2, \ldots, e_n$ the same number of times. From the definition of an invariant, by considering the linear transformation

$$T^*e_i = ce_i$$
$$T^*e_j = e_j \qquad j \neq i,$$

for some scalar c, we may conclude that each integer i occurs the same number of times, say g_i, in each right tableau in $\mathscr{L}$. Now by considering the linear transformation

$$T^*e_i = e_j$$
$$T^*e_j = e_i$$
$$T^*e_k = e_k \qquad k \neq i, j,$$

we conclude that $g_i = g_j$ for all i and j, and call the common value g.

Let us now analyze the possible order of the entries in a right tableaux D_i, more particularly in the rows. If in each row of every D_i, every integer j is immediately followed by $j + 1$, then the proof is concluded. We may therefore assume that there is a *smallest* integer j and a *first* row, say the $(k + 1)$th, such that j is not followed by $j + 1$ in this row. The rows with this property will be adjacent and below the kth. Say there are Q such rows, $R_{k+1}, \ldots, R_{k+Q}$. Then there are Q entries equal to $j + 1$ out of position. They cannot be in any of the rows preceding R_{k+1}, because these rows already contain an entry equal to $j + 1$. Hence they must lie in the rows following R_{k+Q}. Let R be one such row containing an entry equal to $j + 1$. Then $j + 1$ must be at the left of this row. For it cannot be to the right of the jth place, otherwise the tableau would not be standard in the corresponding column, and it cannot be between the first and the jth place, otherwise the minimality of j would fail.

Hence, following row R_{k+Q} there are Q further rows $R_{k+Q+1}, \ldots, R_{k+2Q}$ for each of which the left entry is $j + 1$.

Thus, the tableau must be of the following form:

$$\begin{array}{l}
k \text{ rows}\left\{\begin{array}{llllll}
1 & 2 \;\ldots & j & j+1 & * & \cdots \\
1 & 2 \;\ldots & j & j+1 & * & \cdots \\
\vdots & & & & & \\
1 & 2 \;\ldots & j & j+1 & * & \cdots
\end{array}\right. \\
Q \text{ rows}\left\{\begin{array}{lllll}
1 & 2 \;\ldots & j & * & \cdots \\
\vdots & & & & \\
1 & 2 \;\ldots & j & * & \cdots
\end{array}\right. \\
Q \text{ rows}\left\{\begin{array}{lll}
j+1 & * & \cdots \\
\vdots & & \\
j+1 & * & \cdots
\end{array}\right. \\
\qquad\quad\begin{array}{lll}
* & \cdot & \cdots \\
* & \cdot & \cdots \\
* & \cdot & \cdots
\end{array}
\end{array}$$

where the stars stand for entries greater than $j + 1$.

Since this analysis accounts for all $(j + 1)$'s out of position, we must have $Q = g - k$. Thus, since k was chosen to be minimal Q is the maximal number of j's not followed by $j + 1$ in any D_i. Say this occurs in the tableaux $\{C_1|D_1\}, \ldots,$

$\{C_\alpha|D_\alpha\}$ of $\mathscr{L}$, so that

$$F = \sum_{i=1}^{\alpha} \lambda_i\{C_i|D_i\} + \text{other terms.}$$

Consider the linear transformation

$$T^*e_j = e_j + e_{j+1}$$
$$T^*e_i = e_i \qquad i \neq j$$

Under T^*, each tableaux $\{C_i|D_i\}$ is sent into the sum of the tableaux obtained by replacing in turn every subset of Q or fewer entries equal to j by $j+1$. Of course the resulting tableaux may not be standard or may even equal zero.

Let us see what happens to the first α tableaux by this substitution. Replacing the Q entries equal to j in rows $R_{k+1}, \ldots, R_{k+Q}$ by $j+1$ we obtain standard tableaux with Q fewer entries equal to j. These standard tableaux have fewer j's than necessary, and must be cancelled out by tableaux obtained from other substitutions. By the maximality of Q and the linear independence of the standard tableaux this is impossible. We have thus reached a contradiction which concludes the proof.

We now give an alternative version of the First Fundamental Theorem valid for all fields.

The following lemma is a simple consequence of the multinomial expansion:

LEMMA 1. *Let $F(x, \ldots, z)$ be a homogeneous polynomial function, of degree g, of the coordinates of the vectors $x, \ldots, z$. Then for any scalars $\lambda_i, \ldots, \mu_i$ and vectors $x_i, \ldots, z_i$ we have*

$$F\left(\sum_i \lambda_i x_i, \ldots, \sum_i \mu_i z_i\right) = \sum_{i_1, i_2 \ldots} \cdots \sum_{k_1, k_2 \ldots} \lambda_1^{i_1}\lambda_2^{i_2} \cdots \mu_1^{k_1}\mu_2^{k_2}$$
$$F_{i_1 i_2 \ldots, k_1 k_2 \ldots}(x_1, x_2, \ldots, z_1, z_2, \ldots)$$

where the sum ranges over all $i_1, \ldots, k_1, \ldots$ such that

$$\sum_j i_j + \cdots + k_j = g$$

and the $F_{i_1 i_2 \ldots, k_1 k_2 \ldots}$ are homogeneous of degree g.

The proof is omitted, as the result is well-known.

LEMMA 2. *In a Cayley space of dimension n, let $F(x_1, x_2, \ldots, x_n)$ be a scalar valued function of vectors $x_1, \ldots, x_n$ which is invariant under all non-singular linear transformations T, that is, such that for some scalar function $\lambda(T)$,*

$$F(Tx_1, Tx_2, \ldots, Tx_n) = \lambda(T)F(x_1, \ldots, x_n)$$

Then

$$F(x_1, x_2, \ldots, x_n) = c[x_1, x_2, \ldots, x_n]^g$$

for some constant c and integer g.

The proof is omitted, as the result is well known to hold over an arbitrary field, and an easy consequence of the fact that the determinant is an irreducible polynomial.

LEMMA 3. *Let $F(x_1, \ldots, x_N)$ be a homogeneous invariant of degree g. Then the polynomial*

$$[x_1, \ldots, x_n]^g F(x_1, \ldots, x_N)$$

equals a polynomial in the brackets $[x_{i_1}, x_{i_2}, \ldots, x_{i_n}]$.

Proof: Since the function F is homogeneous of degree g,

$$[x_1, \ldots, x_n]^g F(x_1, \ldots, x_N)$$
$$= F([x_1, \ldots, x_n]x_1, \ldots, [x_1, \ldots, x_n]x_n, [x_1, \ldots, x_n]x_{n+1}, \ldots, [x_1, \ldots, x_n]x_N).$$

Using the identity

$$[x_1, \ldots, x_n]x_j = \sum_{k=1}^{n} [x_1, \ldots, x_{k-1}, x_j, x_{k+1}, \ldots, x_n]x_k$$

and expanding as in Lemma 1, we find that

$$(*) \qquad [x_1, \ldots, x_n]^g F(x_1, \ldots, x_N) = \sum_m c_m F_m(x_1, x_2, \ldots, x_n)$$

where the subscript m ranges over a set of multi-indices, and the coefficients c_m are products of brackets of the form

$$b_j = [x_1, x_2, \ldots, x_{k-1}, x_j, x_{k+1}, \ldots, x_n].$$

Note that for $j > n$ the b_j are algebraically independent (in the case of finite fields of p elements, after making the reduction $x^p = x$). This follows from the algebraic independence of the $\langle x_i | e_j \rangle$.

Because of Lemma 2, the proof will be concluded if we can show that each of the $F_m(x_1, \ldots, x_n)$ is an invariant. Since multiplying an invariant by a product of brackets preserves invariance, we may conclude that

$$[x_1 \ldots x_n]^g F(x_1 \ldots x_N)$$

is an invariant. Thus

$$[Tx_1 \ldots Tx_n]^g F(Tx_1 \ldots Tx_N) = \lambda(T)[x_1 \ldots x_n]^g F(x_1 \ldots x_N)$$

Substituting in (*) we get, since the c_m are also invariants,

$$\sum_m c_m(Tx_1 \ldots Tx_N)F_m(Tx_1 \ldots Tx_n) = \mu(T) \sum_m c_m(x_1, \ldots, x_N)F_m(x_1 \ldots x_n)$$

Since both sides are polynomials in the b_j, and since the b_j are algebraically independent, their coefficients must coincide. This gives

$$F_m(Tx_1 \ldots Tx_n) = \nu(T)F_m(x_1 \ldots x_n)$$

which concludes the proof.

THEOREM 2. (*First Fundamental Theorem of Invariant Theory*). *Every homogeneous invariant in the vectors $x_1, \ldots, x_N$ is expressible as a word in the Cayley algebra, built out of joins and meets alone.*

Proof: By Lemma 3, there is an integer g such that

$$[x_1, \ldots, x_n]^g F(x_1, \ldots, x_N)$$

is a polynomial in the brackets, that is, a linear combination of double tableaux of the form

$$\sum_i \{T_i | D\}$$

where

$$D = \begin{pmatrix} 12 & \ldots & n \\ & \vdots & \\ 12 & \ldots & n \end{pmatrix}.$$

We wish to show that it is possible to cancel $[x_1, \ldots, x_n]^g$ while retaining the rectangular form of the right tableaux. By the Straightening Formula, F may be written as

$$F = \sum \{U_i | V_i\},$$

a linear combination of double standard tableaux. Let

$$U_i' = \begin{pmatrix} x_1 & \ldots & x_n \\ & \vdots & \\ x_1 & \ldots & x_n \\ & U_i & \end{pmatrix} \qquad V_i' = \begin{pmatrix} 1 & \ldots & n \\ & \vdots & \\ 1 & \ldots & n \\ & V_i & \end{pmatrix},$$

where vertical dots indicate that a total of g rows have been placed above each of U_i and V_i as shown. U_i' and V_i' are clearly standard. Now note that

$$[x_1, \ldots, x_n]^g F(x_1, \ldots, x_n) = \sum_i \{U_i' | V_i'\}.$$

We have thus written $[x_1, \ldots, x_n]^g F(x_1, \ldots, x_n)$ as a linear combination of double standard tableaux in two different ways. By the linear independence of the double standard tableaux these must agree, giving

$$V_i' = D.$$

It follows from this that V_i is also rectangular with rows equal to $1 \ldots n$, which concludes the proof.

10. Time-ordering (*sketch*)

We consider here the space $V_{\alpha,\beta}$ introduced in the statement of the Straightening Formula, and now assume that the entries of the vector β are all equal to zero or one; that is, that there are no repeated covectors in any monomial in $V_{\alpha,\beta}$. We now treat $V_{\alpha,\beta}$ as a module over the group-ring of the symmetric group acting on the set of covectors. The proof of the Straightening Formula, considered in this context, says that every submodule of $V_{\alpha,\beta}$ which is invariant under permutations of vectors is spanned by linear combinations of double standard tableaux.

We shall begin by determining the structure of minimal submodules. In characteristic zero, these give an irreducible representation of the symmetric group; but these representations make sense over any field, although they may not be irreducible.

A submodule M of $V_{\alpha\beta}$ which is spanned by inner products of the form $\langle x_i|X_j\rangle$ is minimal if and only if the set of double standard tableaux in M is the set of all possible right tableaux of some fixed shape λ, adjoined to one left tableau L of shape λ with the property that the vectors in row $i+1$ of L are a subset of the vectors in row i for all i.

Proof: We need to show (a) that a submodule of $V_{\alpha,\beta}$ which has as a basis any proper subset of S is no longer invariant under the given permutation group, and (b) that if the covectors in the right tableau of any double standard tableau in S are permuted, then the resulting double tableau may be written as a linear combination of tableaux in S.

Part (a) is true since the set M is transitive under the given permutation group. Part (b) is a consequence of the straightening algorithm, since upon straightening, any tableaux of higher shape which occur will have repeated elements in some row of the left tableau.

An example of minimal invariant module is associated with shape λ as follows. One takes the set S to be the set of all tableaux whose first column on the left side has all entries equal to x_1, whose second column has all entries equal to x_2, etc. These tableaux give explicitly the matrix units of a representation of the symmetric group which in characteristic zero is always irreducible: it can be shown that one obtains in this way all the irreducible representations of the symmetric group.

By extending the above reasoning one can classify all submodules of $V_{\alpha\beta}$ which are spanned by double standard tableaux. A submodule A of $V_{\alpha,\beta}$ spanned by double standard tableaux is spanned by the set of all standard tableaux obtained from a given set S of standard tableaux by iterating the straightening algorithm until no further standard tableaux may be obtained.

In characteristic zero, one obtains in this way the complete reducibility of invariant submodules. However, the algorithm gives an analog of complete reducibility for arbitrary fields.

The preceding idea can be applied to the study of submodules of free associative algebras which are invariant under arbitrary permutations of the variables, by the device of entangling and disentangling, which we now describe.

Let π be a partition of the integer n which we write as $n = \pi_1 + \cdots + \pi_k$ where $\pi_1 \leq \cdots \leq \pi_k$, and let W_π be the submodule of the free associative algebra in the variables $x_1, \ldots, x_n$ spanned by all monomials whose content is the vector $\alpha_\sigma = (\pi_{\sigma_1}, \ldots, \pi_{\sigma_k})$ for some permutation σ of $\{1, 2, \ldots, k\}$.

Such a monomial is of the form

$$x_{i_1} \ldots x_{i_n}$$

where the multiplicities of the x_{i_j} are the integers $\pi_1, \ldots, \pi_k$ in some order.

Associate with this monomial the product

$$\langle x_{i_1}|1\rangle \ldots \langle x_{i_n}|n\rangle$$

in the *commutative* variables $\langle x_{i_1}|1\rangle, \ldots, \langle x_{i_n}|n\rangle$. This association extends to a linear operator F, the *entangling operator*, from W_π to the vector space

$$V_\pi = \sum_{\alpha_\sigma} \sum_{\beta} V_{\alpha\beta},$$

where we sum over all β such that β has n ones and all other entries zero.

Conversely, given an element of V_π, we can recover an element of W_π by applying the *disentangling operator* F^{-1}. For example, from

$$\langle x_1|1\rangle\langle x_2|2\rangle - \langle x_1|2\rangle\langle x_2|1\rangle$$

we obtain, by disentangling, the element

$$x_1x_2 - x_2x_1$$

of W_π. In other words, the Roman numerals in the brackets of V_π indicate the positions of the variables x_i in W_π.

Now, any set of commutative symbols $\langle x_i|j\rangle$ can be interpreted as inner products of vectors x_i and covectors j. We can therefore apply the Straightening Formula, and by the entangling and disentangling operators express every element of W_π in a canonical way as a linear combination of the polynomials obtained in this way from the double standard tableaux.

In this way, the classification of identities in associative algebras is reduced under suitable homogeneity assumptions to the classification of the identities defined by double standard tableaux. Consider an associative algebra A in the variables $x_1, \ldots, x_N$. An identity holding in A is an expression of the form

$$\sum a_{i_1\ldots i_n} x_{i_1} \cdots x_{i_n} = 0,$$

where the $a_{i_1\cdots i_n}$ are elements of the field F which are invariant under any permutation of the variables $x_1, \ldots, x_N$. This identity is associated with the submodule generated by the monomials

$$\sum_{i_1,\ldots,i_n \in \{1\ldots N\}} a_{i_1\ldots i_n} x_{\sigma(i_1)} \cdots x_{\sigma(i_n)}$$

as σ ranges over all permutations. Upon applying the entangling operator, this submodule is mapped into a subspace V_π. The Straightening Formula now yields a basis of double standard tableaux. The image of this basis under the disentangling operator F^{-1} yields a canonical set of monomials in A which generate the submodule. For example, the tableau

$$\langle x_1 \ldots x_n|12\ldots n\rangle$$

gives after disentangling the standard identity

$$\sum_\sigma (\operatorname{sign} \sigma) x_{\sigma 1} x_{\sigma 2} \cdots x_{\sigma n}.$$

An interpretation of the First Fundamental Theorem in this context gives some pertinent information.

11. Symmetric functions (*sketch*)

The classical identities between symmetric functions can be obtained from identities in a Cayley algebra.

Let the field F be obtained from a base field K by adjoining as many transcendentals (variables) as will be needed in the sequel. Choose a doubly infinite sequence of vectors

$$x^{(1)}, x^{(2)}, x^{(3)}, \ldots$$

and covectors

$$U^{(1)}, U^{(2)}, U^{(3)}, \dots$$

in an n-dimensional vector space V over F, and assume that all coordinates, relative to a coordinate system which will remain fixed from now on, are independent transcendentals $x_j^{(i)}$ and $U_j^{(i)}$.

Let K_λ be the field obtained from K by adjoining n transcendentals λ_j and let L be the linear map of the field F into K_λ defined as follows

$$L(x_i^{(k)}U_j^{(k)}) = \delta_{ij}\lambda_j,$$

$$L(x_{i_1}^{(1)}U_{i_1}^{(1)}x_{i_2}^{(2)}U_{i_2}^{(2)}\dots x_{i_k}^{(k)}U_{i_k}^{(k)}) = L(x_{i_1}^{(1)}U_{i_1}^{(1)})L(x_{i_2}^{(2)}U_{i_2}^{(2)})\dots L(x_{i_k}^{(k)}U_{i_k}^{(k)})$$

and so forth, where the indices are not necessarily distinct. Other values of L on monomials are set equal to zero. Note that

$$L(x^{(i)} \wedge U^{(j)}) = \sum_k \lambda_k \delta_{ij}$$

The polynomial

$$L(\langle x^{(1)}\dots x^{(k)} | U^{(1)}\dots U^{(k)}\rangle)$$

equals $k!a_k$, the kth elementary symmetric function in the variables λ_j.

We shall carry out the proof only for the case $k = 2$, the general case being similar. Thus, in terms of the given basis $e_1 \dots e_n$, and dual basis $E_1 \dots E_n$,

$$x^{(1)} \vee x^{(2)} = \sum_{i<j} (x_i^{(1)}x_j^{(2)} - x_i^{(2)}x_j^{(1)})e_i \vee e_j$$

$$U^{(1)} \wedge U^{(2)} = \sum_{i<j} (U_i^{(1)}U_j^{(2)} - U_i^{(2)}U_j^{(1)})E_i \wedge E_j$$

so that the induced inner product becomes

$$\langle x^{(1)}x^{(2)} | U^{(1)}U^{(2)}\rangle = \sum_{i<j} (x_i^{(1)}x_j^{(2)} - x_i^{(2)}x_j^{(1)})(U_i^{(1)}U_j^{(2)} - U_i^{(2)}U_j^{(1)}).$$

Applying the linear functional L, this becomes

$$\sum_{i<j} L(x_i^{(1)}x_j^{(2)}U_i^{(1)}U_j^{(2)} + x_i^{(2)}x_j^{(1)}U_i^{(2)}U_j^{(1)}),$$

as the other two terms vanish when L is applied. But it is seen from the definition of L that the above equal $2!a_2$, as desired.

The polynomial

$$L((x^{(1)} \wedge U^{(2)}) \vee (x^{(2)} \wedge U^{(3)}) \vee \cdots \vee (x^{(k)} \wedge U^{(1)}))$$

equal s_k, the power-sum symmetric function in the λ_j.

Again we carry out the proof for $k = 2$, where we find, upon expanding,

$$L((x^{(1)} \wedge U^{(2)}) \vee (x^{(2)} \wedge U^{(1)})) = L\left(\sum_{i,j} x_i^{(1)}U_i^{(2)}x_j^{(2)}U_j^{(1)}\right)$$

all terms with $i \neq j$ vanish, by the definition of L, and this reduces to

$$L\left(\sum_i x_i^{(1)}U_i^{(1)}x_i^{(2)}U_i^{(2)}\right) = \sum_i \lambda_i^2,$$

as desired.

Every polynomial in the inner products $\langle x^{(i)}|U^{(j)}\rangle$ which contains as many occurrences of the vector variables $x^{(i)}$ as of the covector variable $U^{(i)}$ for each i, equals a symmetric function of the λ_k.

Indeed, every such polynomial can be written as a sum of products of disjoint cycles as in (*), and each such cycle equals a symmetric function s_k.

Identities for symmetric functions may have analogs in the Cayley algebra. The analog of Newton's formula, expressing the a_k in terms of the s_k, is obtained as follows. Expanding the inner product defining a_k, we find

$$\langle x_1 \dots x_k|U_1 \dots U_k\rangle = \langle x_1|U_1\rangle\langle x_2 \dots x_k|U_2 \dots U_k\rangle + \sum_{i>1} \pm \langle x_i|U_1\rangle\langle x_1 \dots \hat{x}_i \dots x_k|U_2 \dots U_k\rangle$$

The second term on the right is further expanded, giving $k-1$ summands of the form

(*) $\qquad (x_i \wedge U_1) \vee (x_1 \wedge U_i) \vee \langle \hat{x}_1 x_2 \dots \hat{x}_i \dots x_k|\hat{U}_1 U_2 \dots \hat{U}_i \dots U_k\rangle$

as well as other terms. The remaining terms are further expanded, giving terms of the form

(**) $\qquad (x_j \wedge U_1) \vee (x_i \wedge U_j) \vee (x_1 \wedge U_i) \vee$ (Inner Product)

as well as other terms. Clearly terms of the form (*) correspond to $s_2 a_{k-2}$, and terms of the form (**) to products $s_3 a_{k-3}$, etc.

Waring's formula, expressing the a_k in terms of the s_k, is even easier. It reduces to the remark that the determinant

$$\langle x_1 \dots x_k|U_1 \dots U_k\rangle = \det\langle x_i|U_j\rangle$$

is a sum of terms, each of which splits into disjoint cycles of a permutation of the indices.

We can define the Schur functions e_μ corresponding to a tableau of shape μ to be L applied to the symmetrized tableau (v. below) of shape μ in the variables x_i and U_j. It is then not difficult to derive the determinant expression for the Schur functions in terms of the elementary symmetric functions a_k. Various results on characters of the symmetric group can be derived and extended by the present approach.

12. Further work

We sketch some lines of work indicated by the present investigations. Some are intended to display applications of the present technique; others are topics which might be further pursued.

(1) *The Gordan–Capelli formula*

The Gordan–Capelli formula is a consequence of the Straightening Formula; we state it without proof—and in greater generality than is found in previous work—avoiding the use of polarization operators which distract from the combinatorial simplicity of the result.

By changing the linear ordering of the variable vectors in all possible ways, and adding the corresponding expressions, one obtains an expansion which is independent of the choice of a linear order, and in some ways simpler. The drawback of such an expansion is that it holds in general only in characteristic zero, unlike the Straightening Formula.

Define a symmetrized tableau $\sigma(T_1|T_2)$ as the sum of all the double tableaux obtained by permuting all the elements of each row of T_1 in turn and independently, repetitions allowed. Thus if a row has k entries, these will be $k!$ terms, even if the row contains repeated entries.

One can show that in characteristic zero the symmetrized tableaux form a basis for $V_{\gamma,\beta}$; this is, in the case of distinct variables, the Gordan–Capelli expansion.

(2) *Strength of identities*

The Birkhoff–Witt theorem can be read as stating that, in an associative algebra, the product xy can be *recovered* from the bracket $xy - yx$; in other words, the bracket is sufficiently strong to give back the product. On the other hand, it is known that the Jordan product $xy + yx$ is in general not strong enough to give back the product. The question can be posed more generally when a given non-commutative polynomial is strong enough to yield another. We hazard the conjecture that these questions can be attacked by the time-ordering device, where $xy - yx$ becomes $(xy|12)$, together with the Straightening Formula.

(3) *Syzygies*

The Cayley algebra analog of the Second Fundamental Theorem of invariant theory is the problem of finding a set of identities on joins and meets which, in a suitable sense, form a basis for the set of all identities.

More important is the problem of the identities between identities, or syzygies of the second order. Little work has been done on this difficult subject.

(4) *Other groups*

There are analogs of the Straightening Formula for the orthogonal and the symplectic groups, which could not be included here. For the orthogonal group it is closely related to identities for spherical harmonics and Hermite polynomials. For the symplectic group, the result is similar to the Straightening Formula, except that determinants are replaced by Pfaffians. One obtains a systematic way of deriving and proving identities for Pfaffians, as well as an explanation of the oft-noted analogy between the two.

(5) *Invariants*

The age-old problem of the computation of projective invariants for sets of linear varieties can be attacked by the present techniques, and we shall limit ourselves to a remark here. Plethysm can be reinterpreted in the Cayley algebra as the relationship between the induced Cayley algebra built on extensors of step k endowed with the bracket obtained from Sylvester's identity, and the given Cayley algebra.

(6) *Word problems and invariant theory*

The version of the Straightening Formula given above is not the most general; we have chosen it because the proof requires fewer notational artifices. A more general version is concerned with words in the Cayley algebra built out of vectors and covectors, and not necessarily of step zero. The result is similar, except that one requires double standard tableaux where the left and right side are not necessarily of the same shape. In this more general version, the Straightening Formula can be viewed as the solution of the word problem in the Cayley algebra for words containing at most vectors and covectors. Several generalizations are suggested by this viewpoint. One may ask in which cases other word problems in the Cayley algebra are solvable, for words containing symbols for extensors of all steps in prescribed numbers. This problem seems not to have ever been previously treated. While it is possible that all such word problems may be solvable, there is one subclass which lends itself to a more straight-forward treatment. This is the word problem for sets of extensors whose supports generate a semimodular lattice of flats in projective space.

(7) *Hopf algebras*

We have neglected the coalgebra structure of the exterior algebra. However, the Hopf algebra structure is indispensable for a better understanding of some of the problems mentioned here especially for syzygies of higher order. The symbolic method of invariant theory is a Hopf algebra technique in disguise.

(8) *Matching Theory*

We have stated elsewhere that matching theory can be systematized by the methods of linear algebra. In support of this contention we sketch a proof of Philip Hall's Marriage Theorem. Thus, given a bipartite graph G on $A \times B$ with the property that every subset of k vertices in A connected to at least k vertices in B, we must show that there exists an injective function $f: A \to B$ such that for every $a \in A$, $(a, f(a))$ is an edge of the graph. The function f is called a matching of A to B.

We define a ring $F(G)$, called the *free ring* of the graph G, following an idea that goes back to Frobenius. Let K' be the free extension of the rational field K obtained by adjoining independent transcendentals $(a_i|x_j)$ as a_i ranges over the set A and x_j over the set B and let $F(G)$ be the homomorphic image of K' obtained by setting $(a_i|x_j) = 0$ whenever the pair (a_i, x_j) is not an edge of the bipartite graph G. We can find a vector space V and in it vectors a_i and covectors x_j, such that $\langle a_i, x_j \rangle = (a|x_j)$.

The Marriage Theorem states (assuming for simplicity that there are as many vectors as there are covectors) that the matrix of the $(a_i|x_j)$ is non-singular under the stated hypotheses, or equivalently, that the vectors a_i as well as the covectors x_j form a basis under the stated hypothesis.

Proof: Suppose the conclusion fails. Then we can find a minimal dependent set of vectors $a_1, \ldots, a_j$, say, such that $a_1 a_2 \ldots a_j = 0$. Let X be an extensor of step $j-1$. Expanding

$$X \wedge (a_1 a_2 \ldots a_j) = 0$$

by the alternative law, we find that

$$\sum_{i=1}^{j} \pm \langle a_1 \ldots \hat{a}_i \ldots a_j | X \rangle a_i = 0. \qquad (*)$$

Since $a_1 \ldots a_{j-1}$ are independent, we can find covectors $x_1, \ldots, x_{j-1}$, say, such that $\langle a_1 \dots a_{j-1} | x_1 \dots x_{j-1} \rangle \neq 0$. Since $a_1, \dots, a_j$ is a minimal dependent set, it follows that $\langle a_1 \dots \hat{a}_i \dots a_j | x_1 \dots x_{j-1} \rangle \neq 0$ for all i. Expanding in the field $F(G)$. we find

$$\langle a_1 \dots \hat{a}_i \dots a_j | x_1 \dots x_{j-1} \rangle = \det(a_k | x_p) = c_i$$

where $1 \leq k \leq j$ with $k \neq i$, and $1 \leq p \leq j - 1$. Thus

$$\sum_{i=1}^{j} c_i (a_i | x_q) = 0.$$

where $c_i \neq 0$ for all i.

If $q \geq j$ and $(a_i | x_q) \neq 0$ then $(a_i | x_q)$ is transcendental over the field obtained by adjoining the c_i to K. Hence the above equation can hold only if $(a_i | x_q) = 0$ for $1 \leq i \leq j$ and $j \leq q < n$, where n is the dimension of the space. We conclude that the set $a_1 \dots a_j$ of vertices of A is related at most to the $j - 1$ vertices $x_1, \dots, x_{j-1}$ of B. contradicting the hypothesis and ending the proof.

(9) *Translating Geometry into Algebra*

The identities developed in Section 6 indicate that the formalism of Cayley algebra should yield a technique for verifying geometric statements by algebraic methods. Such a hope was indeed the moving force behind much of the work on invariant theory carried out during the Nineteenth Century. Strangely, however, this hope remained unfulfilled, and treatises on invariant theory written at the time limit themselves to a few generalities, such as Gram's theorem. This paradoxical situation, which contributed in some measure to the downfall of classical invariant theory, is partly due to the lack of a clearly developed system of first-order logic in which to express geometric statements.

We confine the discussion to joins and meets of subspaces. If $\bar{A}$ and $\bar{B}$ are subspaces of a projective space S then we write $\bar{A} \cap \bar{B}$ for their intersection, and $\bar{A} \cup \bar{B}$ for their *sum*, that is, for the smallest subspace spanned by $\bar{A}$ and $\bar{B}$, at times also called the *join*.

The problem of translating an assertion of projective geometry into an equivalent assertion in the Cayley algebra can be subdivided into two headings:

(1) Develop an algorithm for verifying whether an identity involving intersections and sums (that is, a word in $\cup$ and $\cap$) of subspaces of projective space holds.

(2) Develop a decision procedure for the first-order theory of projective geometry.

Let $L(V)$ be the lattice of subspaces of the vector space V, where lattice-joins and lattice-meets are written $\cup$ and $\cap$. We shall be concerned with translating, and, insofar as possible, verifying a first-order logic statement in the algebra of lattice-joins and meets, into the language of Cayley algebras. We only consider universal sentences. These are sentences constructed from identities in the lattice of subspaces using the logical connectives "and", "not" and "implies", which we shall call *propositions*.

(a) Let the variables $a, b, \dots, c, x, y, \dots, z$ denote generic vectors; in other words, any identity in these variables states that the identity holds no matter what values are given to the variables. It follows from the Straightening Formula that the ring of brackets whose entries are generic vectors is an integral domain: it follows further that the *word problem* for any conjunction of identities in the algebra of brackets is *solvable*. Indeed, the proof of the Straightening Formula gives an explicit algorithm for the solution of the word problem (see remarks under *Word Problems*). Thus, if a given proposition can be shown to be equivalent to an identity in the algebra of brackets, then the truth of the proposition can be decided.

(b) It has been shown by Scarpellini and Whiteley that every true proposition in an integral domain is equivalent to the conjunction of equalities and inequalities.

This result is a logical equivalent of Hilbert's Nullstellensatz. It is not known whether, in the special case of the algebra of brackets, the equivalence can be obtained from an explicit algorithm.

An identity involving sums and intersections can be shown to be equivalent to a conjunction of identities and inequalities in the algebra of brackets by the following steps.

(c) An identity of the form

$$\bar{A} \geq \bar{B}$$

in the lattice $L(V)$ can be "translated" into an identity in brackets as follows. Let A and $B = b_1 b_2 \ldots b_k$ be extensors supporting $\bar{A}$ and $\bar{B}$. The above identity is equivalent to the conjunction of the k identities

$$b_i \vee A = 0, \qquad i \leq i \leq k.$$

Completing to brackets if necessary, we see that this is equivalent to a conjunction of bracket identities.

(d) An identity of the form

$$\bar{A} = \bar{B} \cup \bar{C}; \qquad \bar{A}, \bar{B}, \bar{C} \in L(V),$$

can be translated into an identity in brackets as follows. The above is equivalent to the proposition:

or every $\bar{X}$,

(*) $$\bar{X} \geq \bar{B} \quad \text{and} \quad \bar{X} \geq \bar{C} \quad \text{if and only if } \bar{X} \geq \bar{A}.$$

Each of the containment relations is constructively equivalent to a conjunction of bracket identities by (c): further, by (b) the implication is equivalent to a bracket identity.

(e) An identity of the form

(**) $$\bar{A} = \bar{B} \cap \bar{C}$$

is translated similarly.

(f) A lattice-identity (or inequality) is decomposed into a succession of identities of the form (c), (d), and (e), by introducing extra variables if necessary.

(g) An alternative approach to steps (d) and (e) is the following. In the special case when $\bar{B} \cup \bar{C} = \bar{V}$ then the verification of (**) becomes trivial, as it reduces to checking that $A = B \wedge C$. This can be done constructively, by (c), verifying $\bar{A} \geq \bar{B} \cap \bar{C}$ and $\bar{A} \leq \bar{B} \cup \bar{C}$ in turn. If $\bar{B} \cup \bar{C} \neq V$ then we can use the *reduced bracket* modulo a generic extensor X. Then $\bar{A} = \bar{B} \cap \bar{C}$ if and only if A is equivalent to $B \wedge C$ modulo every extensor X. The definition of $B \wedge C$ depends on the choice of X.

The verification can be cut down to a *finite* number of extensors X by a process that can be considered as the Cayley algebra analog of Herbrand's theorem. In fact, a reduced bracket can be considered as the Cayley algebra analog of a quantifier. Just as in Herbrand's theorem, the reduction to a finite number of X does not yield a decision algorithm.

(h) If a proof of a lattice proposition is available which uses ordinary projective coordinates, then this proof can be translated step by step into the algebra of brackets, and be made to yield constructively a conjunction of identities and inequalities which is equivalent to the lattice proposition. This idea was partially exploited by Whiteley, but can be made very simple in the language of Cayley algebras.

13. Acknowledgments

The idea of a standard tableau made its first appearance with Clebsch, who gave ingenious applications to geometry. With him appeared also the device of polarization, further developed and sharpened by Capelli in the celebrated expansion bearing his name. However, Capelli did not recognize the importance of Clebsch's basic idea. Alfred Young, after careful study of the ideas of Clebsch and Capelli, introduced in 1901 the tableau expansion that bears his name. However, it was not until Young's third paper, published in 1927, that standard tableaux made their reappearance. In this paper one finds the first version of what—suitably generalized—we have called the straightening algorithm, which has been used since in several circumstances.

It seems that Young may have had an inkling of the Straightening Formula. To be sure, double tableaux were used by him for the representations of the octohedral groups, but are nowhere else mentioned in his work. Turnbull, in the short appendix added to his book for the second edition, sketches Young's ideas. Our work grew largely out of trying to understand some of Turnbull's ideas, which are often purely heuristic. The machinery of Cayley algebras was developed under this stimulus. Our statement and proof of the Straightening Formula is, to the best of our knowledge, the first correct and complete one.

The definition of Cayley algebra is new, as is, to the best of our knowledge, the definition of meet. The Scottish convention is inspired by Turnbull, who used it informally. Of the alternative laws, several special cases were known, but we have not found the general case (Theorem 6.5) in the literature.

The brief treatment of symmetric functions was also inspired by some work of Turnbull and Wallace, combined with a linear functional device introduced by Rota.

The time-ordering device was introduced by R. P. Feynman in another context; the present treatment was motivated by the work of P. M. Cohn.

The proof of the Marriage Theorem was arrived at by analyzing some work of Edmonds.

The proofs of the First Fundamental Theorem and of independence of standard tableaux are new.

14. Bibliography

AITKEN, A. C. "A letter to an Edinburgh colleague", *Proc. Edinburgh Math. Soc.* **16** (1968), 166–172.

BOURBAKI, N., *Eléments de Mathématique, Algèbre*, Ch. 3, Hermann, Paris, 1970.

Cohn, P. M., "A theorem on the structure of tensor spaces", *Ann. Math.*, **56** (1952), 254–268.
De Pillis, John, Grassman algebras as Hilbert space. Journal of Algebra, Vol. **10** (1968), 485–500.
Edmonds, J., "Systems of distinct representatives and linear algebra", *J. Res. Nat. Bur. Standards, Sect. B.* **71B** (1967), 241–245.
Ericksen, W. E., "The intersection of subspaces", *Am. Math. Monthly*, **81** (1974), 154–160.
Feynmann, R. P., "An operator calculus having applications in quantum electrodynamics", *Phys. Rev.*, Vol. **84** (1951), 108–128.
Forder, H. G., *The calculus of extension*, Chelsea, New York, 1960.
Harper, L. and Rota, G.-C., "Matching theory, an introduction", in *Advances in Probability*, Vol. **1**, edited by Peter Ney, Dekker, New York, 1970.
Higman, G., "Representations of general linear groups and varieties of p-groups", *Proc. Int. Conf. Theory of Groups*, Australian Nat. Univ. Canberra, Aug. 1965, 167–173, New York, Gordon & Breach, 1967.
Hodge, W. D. and Pedoe, D., *Methods of Algebraic Geometry*, Vol. II, Cambridge Univ. Press, 1968.
Igusa, J., "On the arithmetic normality of the Grassman variety", *Proc. Nat. Acad. Sci.*, **40** (1954), 309–313.
Kähler, E., "Der innere Differentialkalkul", *Rediconti di Matematica*, **21** (1962), 425–523.
Knuth, Donald E., "Permutations, matrices and generalized Young tableaux", *Pacific J. M.*, **34** (1970), 709–727.
Littlewood, D. E., *The Theory of Group Characters and Matrix Representations of Groups*, Second edition, Oxford, Clarendon Press, 1950.
Meyer, F., "Bericht über den gegenwärtigen Stand der Invariantentheorie", *Jahresbericht der deutschen Math. Vereinigung*, **1** (1892), 82–292.
Muir, T. and Metzler, W. H., *A Treatise on the Theory of Determinants*, Dover, New York, 1960.
Rota, G.-C., "The number of partitions of a set", *Am. Math. Monthly*, Vol. **71**, 1964, 498–504.
Scarpellini, Bruno, "On the metamathematics of rings and integral domains", *TAMS*, **138** (1969), 71–96.
Semple, J. G. and Roth, L., *Introduction to Algebraic Geometry*, Oxford, Clarendon Press, 1949.
Stanley, Richard P., "Theory and application of plane partitions, part 1", *Studies in Appl. Math.*, **L** (1971), 167–188.
Turnbull, H. W., "The general symbolic notation for the principle of duality, and its application to determinants", *Proc. London Math. Soc.* 2, **22** (1923), 495–507.
Turnbull, H. W., "A synthetic application of the symbolic invariant theory to geometry", *The Journal of the Indian Math. Soc.*, **XVII** (1927), 1–20.
Turnbull, H. W., "The invariant theory of a general bilinear form", *Proc. London Math. Soc.* 2, **33** (1931), 1–21.
Turnbull, H. W., "The projective invariants of four medials", *Proc. Ed. Math. Soc.* (1942), 55–72.
Turnbull, H. W., and Wallace, A. H., "Clebsch–Aronhold symbols and the theory of symmetric functions", *Proc. Royal Soc. Edinburgh*, **55** (1951), 155–173.
Turnbull, H. W., *The Theory of Determinants, Matrices and Invariants*, Third Ed., Dover, New York, 1960.
Weyl, Hermann, *The classical groups*, Princeton U.P., 1946.
Whiteley, W., *Logic and invariant theory*, Ph. D. thesis, MIT, 1970.
Young, A., "On quantitative substitutional analysis (third paper)", *Proc. London Math. Soc.* 2, **28**, (1927), 255–292.

Columbia University, Massachusetts Institute of Technology, Harvard University

(Received May 15, 1974)