BULL. AUSTRAL. MATH. SOC. 05B25, 11T35, 14G15, 15A15, 15A69, 51E20
VOL. 57 (1998) [479–492]

THE MODULAR COUNTERPARTS OF CAYLEY'S HYPERDETERMINANTS

DAVID G. GLYNN

Let H be a hypersurface of degree m in $PG(n,q)$, $q=p^h$, p prime.

(1) If $m<n+1$, H has 1 (mod p) points.

(2) If $m=n+1$, H has 1 (mod p) points $\iff$ H^{p-1} has no term $x_0^{p-1}\dots x_n^{p-1}$.

We show some applications, including the generalised Hasse invariant for hypersurfaces of degree $n+1$ in $PG(n,F)$, various properties of finite projective spaces, and in particular a p-modular invariant $\det_p$ of any $(n+1)^{r+2}=(n+1)\times\cdots\times(n+1)$ array or hypercube A over a field of prime characteristic p. This invariant is multiplicative in that $\det_p(AB)=\det_p(A)\det_p(B)$, whenever the product (or convolution) of the two arrays A and B is defined, and both arrays are not 1-dimensional vectors. (If A is $(n+1)^{r+2}$ and B is $(n+1)^{s+2}$, then AB is $(n+1)^{r+s+2}$.) The geometrical meaning of the invariant is that over finite fields of characteristic p the number of projections of A from $r+1$ points in any given $r+1$ directions of the array to a non-zero point in the final direction is 0 (mod p). Equivalently, the number of projections of A from r points in any given r directions to a non-singular $(n+1)^2$ matrix is 0 (mod p). Historical aspects of invariant theory and connections with Cayley's hyperdeterminant Det for characteristic 0 fields are mentioned.

1. FINITE PROJECTIVE GEOMETRIES AND FUNCTIONS

$PG(n,q)$ is the finite projective geometry of dimension n over the finite field $F=GF(q)$, where $q=p^h$, p prime. Here we denote the set of all points of $PG(n,q)$ by $[n]$.

$C=C(n,q)$ is the set of all functions $[n]\to F$. It is a vector space of dimension $v:=q^n+\cdots+q+1$ over F, with the usual addition of functions and multiplication of functions by a scalar in F. A function $f\in C$ can be thought of as a vector $(a_1,\dots,a_v)$ with $f(P_j)=a_j$, for some labelling P_j of the points. A function can be called a *word*, and the *weight* of a word is the number of non-zero values that the function takes. With these well-known assumptions any subspace of functions of C gives a linear code, of which the most important are probably the *geometric codes* investigated in, for example, [**1, 10**].

Received 5th November, 1997

Each point of $[n]$ is represented as usual by homogeneous coordinates $(x_0, \ldots, x_n)$, $x_i \in F$. Any function $f \in C$ is represented by a unique polynomial g in the $n+1$ variables x_i, every term of which is *reduced*; that is, a general term is $t = c\prod_{i=0}^{n} x_i^{a_i}$, where $c \in GF(q)$ and $0 \leqslant a_i \leqslant q-1$, such that $f(x) = g(x)$ for all x in $PG(n,q)$ with $g(0) = 0$. Every term of g has total degree $\sum_{i=0}^{n} a_i = e(q-1)$, where $e \in \mathbb{Z}$, $1 \leqslant e \leqslant n+1$; e need not be the same for different terms. A *term* has the form $cx_0^{a_0} \ldots x_n^{a_n}$ while the corresponding *monomial* is $x_0^{a_0} \ldots x_n^{a_n}$.

The *reduction mapping* $r : \mathbb{Z} \to \{0, \ldots, q-1\}$ is such that $r(0) = 0$, $r(n) \equiv n \pmod{q-1}$, and $r\big(e(q-1)\big) = q-1$, for all $e \in \mathbb{Z} \setminus \{0\}$. If $t = c\prod_{i=0}^{n} x_i^{a_i}$ is a term with $c \neq 0$, then the *reduced degree* of t is $\deg(t) := \sum_{i=0}^{n} r(a_i)$. We use this reduction to make certain that we always have valid terms with powers between 0 and $q-1$.

Note that we do not allow a constant in F to be a valid term of a homogeneous function. Thus all valid polynomials g representing functions f of C satisfy the natural condition $g(0) = 0$.

Indeed, every function in C can be written in the form (see [**10**])

$$f(x) = \sum_{\substack{0 \leqslant \lambda_0 \leqslant q-1 \\ \cdots \\ 0 \leqslant \lambda_n \leqslant q-1}} k_{\lambda_0 \ldots \lambda_n} x_0^{\lambda_0} \ldots x_n^{\lambda_n},$$

for unique $k_{\lambda_0 \ldots \lambda_n} \in F$. $k_{0 \ldots 0} = 0$ and so f has no constant term. Each term of $f(x)$ must have degree that is divisible by $q-1$ and is non-zero. The coefficients $k_{\lambda_0 \ldots \lambda_n}$ of the terms are uniquely determined by the function.

We say that a polynomial function f in C *has a term* t if the coefficient in $GF(q)$ of t in the unique representation of f, as a sum of terms, is non-zero.

Consider a monomial $t = {x_0}^{a_0} \ldots {x_n}^{a_n}$, where the *exponents* a_i satisfy $0 \leqslant a_i \leqslant q-1$, $(a_i \in \mathbb{Z})$. Also, the degree of t should satisfy $\deg(t) := \sum_{i=0}^{n} a_i = j(q-1)$, for some $1 \leqslant j \leqslant n+1$. Thus t is a *valid monomial* of $C(n,q)$. The *degree sequence* $S(t)$ of t is the *cycle* of integers $(s_0, s_1, \ldots, s_{h-1})$, where $s_k = \deg\big(t^{p^{-k}}\big)/(q-1)$, and the subscripts k are considered to be in the cyclic group modulo h. The degree sequences of valid monomials are characterised by the fact that they satisfy a certain set of inequalities (due to Hamada), that are very important for various properties of the geometric codes; see [**10**]. However, we shall not need these properties in the following.

2. Hypersurfaces and Functions

Let us recall that a *hypersurface* (or *primal*) of degree d of $PG(n,q)$ is defined by a single homogeneous algebraic equation of degree d in the $n+1$ variables

https://doi.org/10.1017/S0004972700031890 Published online by Cambridge University Press

$x_0, \ldots, x_n$. Clearly, if the equation is $H = 0$, then the properties of the function $f : x \mapsto H(x)^{q-1}, (x \in [n])$, are tied in with those of the hypersurface. In fact, f takes the value 0 at all points of H, and the value 1 at all other points. If $N(H) := \left|\{x \in [n] \mid H(x) = 0\}\right|$ then we have:

LEMMA 2.1.

$$N(H) \equiv 1 - \sum_{x \in [n]} h(x) \pmod{p},$$

where $h : x \mapsto H(x)^{q-1}, (x \in [n])$.

PROOF: This follows immediately from the two facts that $\left|[n]\right| = 1 + q + \cdots + q^n \equiv 1 \pmod{p}$, and $x^{q-1} = 1$ if $x \neq 0$ and $x \in F$. □

The following is a well-known generalisation of the fact that

$$\sum_{y \in F} y^i = \begin{cases} 0, & \text{if } 0 \leqslant i < q-1 \\ -1, & \text{if } i = q-1. \end{cases}$$

LEMMA 2.2. *If g is the algebraic function representing a function f of C, (that is $f : [n] \to F$), then $\sum_{x \in [n]} g(x) = (-1)^n k$, where k is the coefficient of $x_0^{q-1} \ldots x_n^{q-1}$ in g.*

Now let h be the algebraic function that represents $H(x)^{q-1}$, where H is a hypersurface of $PG(n, q)$. Applying Lemmas 2.1 and 2.2 we obtain:

LEMMA 2.3.

$$N(H) = 1 - (-1)^n k, \ (\in GF(p)),$$

where k is the coefficient of $x_0^{q-1} \ldots x_n^{q-1}$ in h.

Note that h can be calculated from H by expanding out H^{q-1} as a sum of monomials, and then using the reduction mapping of general terms; see Section 1.

Now the reduction mapping takes a term either to itself or to a term of lesser degree. Thus the only way to get the reduced term $t := x_0^{q-1} \ldots x_n^{q-1}$ from a term of H^{q-1} is if $\deg\left(H^{q-1}\right) \geqslant \deg(t) = (n+1)(q-1)$. Thus only if $\deg(H) \geqslant n+1$. Hence we have the following simple and useful result:

THEOREM 2.4. *A hypersurface H of degree $\leqslant n$ in $PG(n, q)$ has $N(H) \equiv 1 \pmod{p}$ points.*

PROOF: This uses Lemma 2.3, and the fact that $k = 0$ from the preceding degree argument. □

Now suppose that $\deg(H) = n + 1$. Let us calculate the coefficient of $t := x_0^{q-1} \ldots x_n^{q-1}$ in the reduced function representing H^{q-1}. Let $J := H^{p-1}$. Then

$$H^{q-1} = J.J^p.J^{p^2} \ldots J^{p^{h-1}},$$

https://doi.org/10.1017/S0004972700031890 Published online by Cambridge University Press

where $q = p^h$. Thus, to obtain the reduced term t there must be some product

$$t = t_0 . t_1^p . t_2^{p^2} \dots t_{h-1}^{p^{h-1}},$$

where each t_i is a term of J. Since $\deg(t_i) = (n+1)(p-1)$, then $\deg(t_i^{p^i}) = (n+1)p^i(p-1)$, and reduction only reduces this degree, then we see that the only way to obtain t is if there is actually no reduction in the above equation. That is, $t_i^{p^i}$ is already reduced, for all $0 \leqslant i < h$. This means that the above equation for t is exact, and not modulo reduction.

Now consider the power of x_j modulo p on both sides. Then on the left it is $q - 1 \equiv -1$, while on the right it is the power of x_j in t_0, since the other t_i's have powers divisible by p. Thus t_0 contains x_j at least to the power $p-1$. But $\deg(t_0) = (n+1)(p-1)$, and so $t_0 = x_0^{p-1} \dots x_n^{p-1}$.

Next, divide both sides of the equation for t by t_0, and then take p'th roots. Again we obtain an exact equation (not modulo reduction) and using a similar argument we see that also $t_1 = x_0^{p-1} \dots x_n^{p-1}$. Indeed, by this sequence of arguments (or using induction) we see that $t_i = x_0^{p-1} \dots x_n^{p-1}$, for all i with $0 \leqslant i < h$.

The coefficient of t in the reduced function of H^{q-1} is thus $k = c^{1+p+p^2+\dots+p^{h-1}}$, where c is the coefficient of $s := x_0^{p-1} \dots x_n^{p-1}$ in $J := H^{p-1}$. Using Lemma 2.3 we obtain:

THEOREM 2.5. *A hypersurface H of degree $n+1$ in $PG(n,q)$ has*

$$N(H) = 1 - (-1)^n c^{1+p+p^2+\dots+p^{h-1}}, \ (\in GF(p))$$

points, where c is the coefficient of $x_0^{p-1} \dots x_n^{p-1}$ in H^{p-1}.

From Theorems 2.4 and 2.5 this follows immediately:

COROLLARY 2.6. *Let H be a hypersurface of degree m in $PG(n,q)$, $q = p^h$, p prime.*

(1) *If $m < n+1$, H has 1 (mod p) points.*
(2) *If $m = n+1$, H has 1 (mod p) points $\iff$ H^{p-1} has no term $x_0^{p-1} \dots x_n^{p-1}$.*

In the next section we give numerous useful consequences of the above result.

3. Examples and Applications

EXAMPLE 3.1. Hyperplanes have 1 (mod p) points. Indeed, a hyperplane of $PG(n,q)$ has a linear equation (of degree 1) and so if $n \geqslant 1$ it has 1 (mod p) points. In fact, the number of points is $1 + q + \dots + q^{n-1}$.

https://doi.org/10.1017/S0004972700031890 Published online by Cambridge University Press

EXAMPLE 3.2. The Discriminant of a Quadratic Form in Two Variables. Put $n = 1$ and let H be of degree 2 in 2 variables. Then H has 1 (mod p) solutions if and only if it has 1 solution if and only if the discriminant D is 0. One can indeed check that the coefficient of $x_0^{p-1}x_1^{p-1}$ in H^{p-1} is $D^{(p-1)/2}$ if p is an odd prime. See [7].

EXAMPLE 3.3. Plane Quadrics have 1 (mod p) points. Thus every plane quadric of $PG(2,q)$ has at least one *real* point.

From this it follows immediately that every irreducible conic in $PG(2,q)$ has precisely $q+1$ points, since through a *real* point of such a conic there is precisely one tangent and all the other q lines are chords.

EXAMPLE 3.4. The Hasse Invariant of Plane Cubic Curves. A cubic curve H in $PG(2,q)$ has 1 (mod p) points if and only if H^{p-1} has no term $(x_0x_1x_2)^{p-1}$. Clearly this coefficient is an algebraic invariant under the group $PGL(3,q)$ of homographies of $PG(2,q)$. It is known in algebraic geometry by the name Hasse invariant, and is not just an invariant for finite fields, but for all fields of characteristic p; see [12]. The invariant there actually is for plane elliptic cubic curves: if it is zero the curves are called *super-singular*. A word that is used elsewhere is *equianharmonic*. For $p = 2$ see a detailed analysis of the invariants, syzygies, and related number-theoretic properties of plane cubic curves in characteristic two in [9].

EXAMPLE 3.5. Quadrics of $PG(n,q)$ have 1 (mod p) points if $n \geqslant 2$. This follows since quadrics have degree 2. In fact, one has a formula for all types of quadrics, reducible or not; see [13].

EXAMPLE 3.6. Intersections of Quadrics in $PG(3,q)$. By taking the product of two quadrics $Q_1 = 0$ and $Q_2 = 0$ one obtains a (reducible) quartic hypersurface $Q_1Q_2 = 0$. Using the invariant that is the coefficient of $x_0^{p-1}x_1^{p-1}x_2^{p-1}x_3^{p-1}$ in $(Q_1Q_2)^{p-1}$ one can calculate the size i of the intersection between the two quadrics modulo p. For each quadric alone has 1 (mod p) points, and so their union has $2 - i$ (mod p) points. In particular, when $p = 2$ this gives a simple criterion about the intersection modulo 2.

EXAMPLE 3.7. Semifield Theory. A semifield of dimension $n+1$ over $GF(q)$ gives a hypersurface H of degree $n+1$ with no point in $PG(n,q)$. It is the so-called *norm form*. Given a basis of the semifield over $GF(q)$, $b_0, \ldots, b_n$, one writes the product of any two basis elements in the semifield as $b_i * b_j = \sum_{ijk} a_{ijk}b_k$, obtaining a so-called *non-singular cube* A over $GF(q)$. The norm form is then the hypersurface corresponding to setting the determinant of a general sum of slices of A in one direction equal to zero. Thus, with the array multiplication of the next section, the form could be defined as $H : \det(xA) = 0$.

Since H can have no points in x in $PG(n,q)$, H^{p-1} has a certain coefficient of the

diagonal term to the power $p-1$. In fact, this invariant coefficient $c = \det_p(A)$ should satisfy $c^{1+p+p^2+\cdots+p^{h-1}} = (-1)^n$ in $GF(p)$; see Theorem 2.5. For example, consider the case $q = 3$, $n = 1$, $h = 1$, $p = 3$, and the field $GF(9)$ over $GF(3)$. Using a basis $1, \alpha$, where $\alpha^2 = -1$ we have $1 * 1 = 1$, $1 * \alpha = \alpha * 1 = \alpha$, $\alpha * \alpha = -1$. One can construct a $2 \times 2 \times 2$ non-singular cube A with 2-dimensional slices

$$A_0 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \text{ and } A_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Thus $\det_3(A)$ is the coefficient of $(x_0x_1)^2$ in $\det(x_0A_0 + x_1A_1)^2 = (x_0^2 + x_1^2)^2$. Thus $\det_3(A) = 2 \equiv -1 \pmod 3$ as we have stated.

We can show that this $\det_p$ is an invariant of cubical arrays that is independent of the $3! = 6$ symmetries (permuting the three directions of the cube). In the theory of semifields these symmetries are called the 'isotopes'. A similar definition could be applied to algebras over a field of characteristic p. Thus the 'isotopes' of such an algebra all have the same invariant $\det_p$. A generalisation to hypercubes is investigated in the next section.

4. The Invariant $\det_p$

First let us recall a slight generalisation of matrix multiplication:

Definition 4.1. *Let A be an $m_0 \times \cdots \times m_{r+1}$ array ($r+2$-dimensional over F), and B an $n_0 \times \cdots \times n_{s+1}$ array ($s+2$-dimensional over F), then the product $C := AB$ is defined if $m_{r+1} = n_0$, and is the $m_0 \times \cdots \times m_r \times n_1 \times \cdots \times n_{s+1}$ array ($(r+s+2)$-dimensional over F), where if $A = (a_{i_0 \ldots i_{r+1}})$ and $B = (b_{j_0 \ldots j_{s+1}})$, then $C = (c_{k_0 \ldots k_{r+s+1}})$, where $c_{k_0 \ldots k_{r+s+1}} := \sum_{l=0}^{m_{r+1}} a_{k_0 \ldots k_r l}\, b_{l k_{r+1} \ldots k_{r+s+1}}$.*

Note that it is possible to multiply two higher-dimensional arrays down any common direction of the same length. For example, an array of dimension $r+2$ can be multiplied by a vector in $r+2$ ways. In this section, we often multiply without specifying the particular direction. In that case, the multiplication must be done only if it makes sense.

A 'hypercube' A is a higher-dimensional array in which all the directions have the same length. Thus it generalises a square matrix of dimension 2. We denote the dimensions of a hypercube by $(n+1)^{r+2}$ if it has $r+2$ dimensions of length $n+1$; thus it is $n+1 \times \cdots \times n+1$, $r+2$ times, over F. Let V be the vector space F^{n+1}: we consider every element of V to be a 1-dimensional array over F. The multiplication of a hypercube by a sequence of vectors from V is called a 'projection'. Clearly, the order of multiplication by vectors is irrelevent: it is the directions that are important. Also,

let us note the important property of multiplication of higher-dimensional arrays: it is associative if the various multiplications are defined.

Let p be a fixed prime, and as before, let $q = p^h$, and $F = GF(q)$. Let A be an $(n+1)^{r+2}$ hypercube over any field G of characteristic p. There are $r+2$ subscripts or directions to A, and $r \geqslant -2$.

We shall define $\det_p(A)$ in an inductive manner, for $r = -1$, and so on.

Firstly for 1-dimensional matrices a (that is, vectors),

$$\det_p(a_0, \ldots, a_n) := (a_0 \ldots a_n)^{p-1}.$$

Note that $\det_p$ is not an invariant for vectors, but it is interesting that we can start the definition from here.

For matrices with $r \geqslant 0$ define $\det_p(A)$ to be the coefficient of $(x_0 \ldots x_n)^{p-1}$ in $\det_p(Ax)$, where Ax is the product of A with the vector x in any of the $r+2$ directions of A. Naturally we must show that this is well-defined. Let us proceed by looking at the case $r = 0$.

THEOREM 4.1. *If A is an $(n+1)^2$ (square $(n+1) \times (n+1)$) matrix, then*

$$det_p(A) = det(A)^{p-1}.$$

PROOF: Consider the coefficient of $(x_0 \ldots x_n)^{p-1}$ in $\det_p(Ax)$. Let the base field G be $F = GF(q)$. $(\det_p(Ax))^{1/(p-1)}$ is the product of $n+1$ hyperplanes in $PG(n,q)$ which is a degenerate hypersurface H of degree $n+1$ in $PG(n,q)$.

From Theorem 2.5 H has 1 (mod p) points if and only if $\det_p(A) = 0$. Now H consists of $n+1$ dependent hyperplanes if and only if H has 1 (mod p) points, for we can use a non-singular linear transformation to assume in the independent case that $H : x_0 \ldots x_n = 0$, and then $\det_p(Ax) = 1$: that is, $n+1$ independent hyperplanes of $PG(n,q)$ cover $1 + (-1)^{n+1}$ (mod p) points. In the dependent case, we can assume that the hyperplanes pass through the point $(1, 0, \ldots, 0)$: that is, that there is no variable x_0 in any of them. Then the product of the hyperplanes cannot contain any variable x_0, and so $\det_p(Ax) = 0$. In this case the dependent hyperplanes cover 1 (mod p) points. Hence the coefficient of $(x_0 \ldots x_n)^{p-1}$ in $\det_p(Ax)$ is zero if and only if $\det(A) = 0$. Since the degree of $\det_p(A)$ in the coefficients of A is $(n+1)(p-1)$ we have $\det_p(A) = \det(A)^{p-1}$, and since this identity is valid for all finite fields $GF(p^h)$ it is valid for all fields of characteristic p. (A polynomial that is zero for an infinite number of values is identically zero.) □

Note that the above identity is valid if we put A^t instead of A. Thus the direction that we multiply A by to form Ax is irrelevent. We can show that this is the general situation as follows.

We shall consider our invariant of a general $(n+1)^{r+2}$ matrix A over the finite field F in a certain way. Suppose $\det_p(A)$ is zero. Then it means that the number of solutions of $\det(Axy\dots z) \neq 0$ is $0 \pmod p$, where $x, y, \dots, z$ are points in r $PG(n,q)$'s corresponding to r selected directions of A. Note that this projection is to ordinary $(n+1)^2$ matrices and so we can replace $\det_p$ by $\det$. Also, the number of r-tuples $(x, \dots, z)$ is $(1+q+\dots+q^n)^r \equiv 1 \pmod p$, and this is why the number of solutions above is $0 \pmod p$. In addition, all the projections to non-singular matrices are obtained by points of the $PG(n,q)$'s, as the projection from any zero vector is a zero matrix. Counting points $xy\dots z$ in $PG\big(r(n+1)-1, q\big)$ amounts to the same as counting r-tuples of points in $PG(n,q)$ up to the multiplication by ± 1, since $(q-1)^i \equiv \pm 1 \pmod p$. A similar comment holds when we project one further as in the next paragraph.

Next, $\det_p(A) = 0$ means that the number of projections $Axy\dots zw$, not being a zero vector of length $n+1$, is also $0 \pmod p$, where $x, y, \dots, z, w$ are points in $r+1$ $PG(n,q)$'s corresponding to $r+1$ selected directions of A. Given the first r directions of $x, y, \dots, z$ there are just two choices for the direction of w. However, these give the same result $\pmod p$ when we count the number of projections to a zero vector. This is because if $\det(Axy\dots z) \neq 0$ then every w gives a non-zero projection, and so the number of these w's is $1+q+\dots+q^n \equiv 1 \pmod p$. On the other hand, if $\det(Axy\dots z) = 0$ we get equal numbers of non-zero projections in either direction, and these numbers are congruent to $0 \pmod p$, since the number of points in a non-empty subspace of dimension d of $PG(n,q)$ (assuming $Axy\dots z$ has rank $n-d$, $d \geqslant 0$) is $1+q+\dots+q^d \equiv 1 \pmod p$.

A projection from $r+2$ directions down to two directions, and then to one direction, in either of the two ways, can be reversed, but on going up from one direction to two there are $r+1$ possibilities. If $\mathrm{Inv}(ij)$ denotes the invariant $\det_p(A)$ obtained by projecting down to the ith and jth directions, and $\mathrm{Inv}(i)$ and $\mathrm{Inv}(j)$ denotes those by going further to directions i and j, then we have shown above that $\mathrm{Inv}(ij) = \mathrm{Inv}(i) = \mathrm{Inv}(j)$. If we want to prove that $\mathrm{Inv}(ij) = \mathrm{Inv}(kl)$ then $\mathrm{Inv}(ij) = \mathrm{Inv}(i) = \mathrm{Inv}(ik) = \mathrm{Inv}(k) = \mathrm{Inv}(kl)$ gives the result. It follows that the definition of $\det_p(A)$ is independent of the directions that are chosen.

Now we are in the position to prove the multiplicative property.

THEOREM 4.2. *If A is an $(n+1)^{r+2}$ matrix, and B is an $(n+1)^{s+2}$ matrix over G, and if $r, s \geqslant 0$, then*

$$det_p(AB) = det_p(A) det_p(B).$$

PROOF: Let us use induction on the dimensions of the matrices (r,s). Firstly, it is true for $(r,s) = (0,0)$, as follows from ordinary 2-dimensional matrices, since

https://doi.org/10.1017/S0004972700031890 Published online by Cambridge University Press

$\det_p(X) = \det(X)^{p-1}$. Next, assume it is true for (r,s) for some $r,s \geqslant 0$. Let A have size $(n+1)^{r+2}$, and B have size $(n+1)^{s+3}$. Then

$$\begin{aligned}
\det_p(AB) &= \text{ coefficient of } (x_0 \dots x_n)^{p-1} \text{ in } \det_p\big((AB)x\big) \\
&= \text{ coefficient of } (x_0 \dots x_n)^{p-1} \text{ in } \det_p\big(A(Bx)\big) \\
&= \text{ coefficient of } (x_0 \dots x_n)^{p-1} \text{ in } \det_p(A)\det_p(Bx), \text{ (by assumption)} \\
&= \det_p(A).\ \text{coefficient of } (x_0 \dots x_n)^{p-1} \text{ in } \det_p(Bx) \\
&= \det_p(A)\det_p(B).
\end{aligned}$$

Note that AB is multiplied by the vector x in any direction of B that is not the one common to A. Of course, once the induction has been reduced from (r,s) to $(r,0)$, we can multiply on the left by x to reduce similarly to the case $(0,0)$ of square matrices. □

5. Brief Historical Comments

When Leibniz died some calculations which amounted to determinants were found in his papers. These were unpublished, but the date is around 1693. Determinants of ordinary square matrices were certainly in use in the 18th century and were known by Euler and especially Vandermonde. However matrix theory as we know it today was not developed until early in the 19th century. For example, we see no mention of it in the major work of Gauß [**8**].

In 1845 Cayley published in the Cambridge Math. Journal a generalisation of the $n \times n$ matrix determinants to higher-dimensional matrices. As mentioned by Cayley in note 92 at the end of [**5**], and in particular in paper [**4**], the composition of two quadratic forms in two variables to give another (as defined by Gauß; see [**8**]) is related to 2^3 matrices and in particular to Cayley's Det invariant.

In fact hyperdeterminants can be constructed for a much larger class of higher dimensional matrices. This has been explained in a book by Gelfand, Kapranov and Zelevinsky; see [**12**, Chapter 14]. Basically their hyperdeterminants work for $(k_0+1)\times \cdots \times (k_{r+1}+1)$ matrices which satisfy $k_i \leqslant k_0 + \cdots + k_{i-1} + k_{i+1} + \cdots + k_{r+1}$, for all $0 \leqslant i \leqslant r+1$. The reason for these perhaps surprising constraints upon the dimensions is the following. In order to construct the invariant there must be in general projections to a zero vector from any $r+1$ directions. Considering the Segre product of the points of the projective spaces of dimensions k_j corresponding to these directions we obtain k_i+1 conditions on the points of the Segre variety of dimension $k_0+\cdots+k_{i-1}+k_{i+1}+\cdots+k_{r+1}$. There is no projection to zero in general if the condition is not met. Their (and Cayley's) hyperdeterminant is zero if and only if there is a collection of $r+2$ points, in each of the $r+2$ directions, such that taking the projections corresponding to any $r+1$, we obtain the zero vector in the final direction.

In 1852 Schläfli [**14**] published an alternative method to obtain hyperdeterminants. As explained in [**12**, Section 14.4] usually this method gives a larger polynomial invariant of which Cayley's Det is a factor. Sometimes, however, it gives precisely Det. This happens at least in the cases $2 \times m \times m$, $3 \times m \times m$, and $2 \times 2 \times 2 \times 2$. This is also related to our recent work on counting points on hypersurfaces over $GF(q)$. See below.

Later on (according to Coolidge [**6**]) the word *hyperdeterminant* had taken on a more general meaning unrelated to higher dimensional matrices, and more closely related to the term *invariant*. In fact, by about 1860 Sylvester had disposed of the former terminology.

Our p-modular hyperdeterminant $\det_p$ has degree $(n+1)(p-1)$ in the coefficients of the $(n+1) \times \cdots \times (n+1)$ matrix. For smaller values of the prime p this is much better than Cayley's. For example, Cayley's hyperdeterminant for $2 \times 2 \times 2$ matrices has degree 4, the p-modular one has degree $2(p-1)$. In this case the two invariants are in fact just powers of one-another. For $3 \times 3 \times 3$ Cayley's has degree 24: ours has degree 3 for $p = 2$, degree 6 for $p = 3$, degree 12 for $p = 5$, and so on. The p-modular hyperdeterminant has a simple method of calculation. Best of all is the fact that it is preserved by products, something that is impossible for Cayley's hyperdeterminant.

Here is Cayley's determinant for the matrix $A := (a_{ijk})$, $(i, j, k = 0, 1)$:

$$\begin{aligned}\operatorname{Det}(A) = {} & a_{000}^2 a_{111}^2 + a_{001}^2 a_{110}^2 + a_{010}^2 a_{101}^2 + a_{011}^2 a_{100}^2 \\ & - 2\,(a_{000}a_{001}a_{110}a_{111} + a_{000}a_{010}a_{101}a_{111} + a_{000}a_{011}a_{100}a_{111} \\ & \qquad + a_{001}a_{010}a_{101}a_{110} + a_{001}a_{011}a_{110}a_{100} + a_{010}a_{011}a_{101}a_{100}) \\ & + 4(a_{000}a_{011}a_{101}a_{110} + a_{001}a_{010}a_{100}a_{111}).\end{aligned}$$

Notice that in this case $\operatorname{Det}(A)$ is the square of the permanent of A modulo 2, which is $\det_2(A)^2$.

6. Additional Notes

For 2^3 matrices we can show that Cayley's hyperdeterminant and $\det_p$ are essentially the same by proving that

$$\det_p(A) \equiv \operatorname{Det}(A)^{(p-1)/2} \pmod{p}, \quad (p \geqslant 3).$$

This can be done as follows. From [**11**] $\operatorname{Det}(A)$ is the discriminant of the quadratic form $H := \det(Ax)$ in two variables $x = (x_0, x_1)$. Thus Schläfli's method amounts to Cayley's in this case. However, we noted in Example 3.2 that the coefficient of $(x_0x_1)^{p-1}$ in H^{p-1} is $D^{(p-1)/2}$ if p is an odd prime.

It is also possible to look at the above in a geometrical way. If $n = 1$, the three projective spaces of an $(n+1)^3$ matrix corresponding to the three directions are all

lines. Our $\det_p$ invariant is zero if and only if the number of projections to a non-singular square matrix from any of the three directions is 0 (mod p) (taking the field to be $GF(q)$). It is not hard to see that there are three different possibilities for the numbers in the three directions in this case: either there are $\{0,0,0\}$, $\{0,0,q\}$, or $\{q,q,q\}$. In the latter situation of a cube with Det $= 0$, there is a unique triple of points, in the three directions, such that taking the projection corresponding to any two we get the zero vector in the third. This is just the criterion of Cayley. Otherwise, if the cube has Det $\neq 0$ the numbers are $\{q-1,q-1,q-1\}$ or $\{q+1,q+1,q+1\}$. The latter case corresponds to a so-called 'non-singular cube' over F, and then the cube can be constructed as the set of multiplication constants of a basis of the division algebra or semifield that is $GF(q^2)$ over $GF(q)$; see Example 3.7. Over a quadratic extension field in both these cases there will be six points, say $r_0, r_1 \in [n]_0$, $s_0, s_1 \in [n]_1$, $t_0, t_1 \in [n]_2$, where $[n]_i$ are the projective spaces in the three directions of A, such that $Ar_0s_0 = As_0t_0 = At_0r_1 = Ar_1s_1 = As_1t_1 = Ar_0t_1 = 0$.

The situation for 2^4 hypercubes is quite interesting in that they are connected to quartic curves (in general, elliptic) contained in the intersection of two hyperbolic quadrics in $PG(3,F)$. For suppose A is the hypercube. Let x and y be points in two fixed directions of A. Multiplying A by x and y gives us the condition that $\det_p(A) = 0$ if and only if the number of pairs of points (x,y) having $\det(xAy)^{p-1} = 0$ ($\iff \det(xAy) = 0$) is 1 (mod p). Expanding $\det(xAy)$ out we can write it in the form

$$\begin{vmatrix} xBy & xCy \\ xDy & xEy \end{vmatrix} = 0,$$

where B, C, D, and E are the various slices of A, all being 2×2 matrices. Considering the corresponding algebraic variety (xBy, xCy, xDy, xEy) of $PG(3,F)$, where x and y vary over all points of $PG(1,F)$, we see that in general this is a hyperbolic quadric, since fixing y and varying x we get one set of lines of a regulus, while fixing x and varying y we get the other regulus. Another way to see that the points lie automatically on a quadratic surface is to note first that it is a two-dimensional variety (with x and y giving the two dimensions). Secondly, in the 10 different quadratic terms of the type $(xBy)^2$, $(xBy)(xCy)$, and so on, there occur precisely 9 monomials of type $x_0^2y_0^2$, $x_0^2y_0y_1$, $x_0x_1y_0y_1$, and so on. In general there will be $10-9=1$ non-trivial quadratic condition on the points (xBy, xCy, xDy, xEy), which will give the hyperbolic quadric. In addition, the determinant condition above is another hyperbolic quadric upon which the points should lie. Hence the condition $\det_p(A) = 0$ means that the number of points on the quartic curve, that is the intersection of the pair of hyperbolic quadrics, is 1 (mod p); see Example 3.6.

For 2^{r+2} matrices ($r \geqslant 2$), $\det_p$ and Det might be conjectured to be related invariants, as they are for $r = 0, 1$. Thus they could be powers of each other modulo

p. However, this can't be. Suppose we consider the matrix A with $a_{0\ldots0} = a_{1\ldots1} = 1$, and all other elements zero. Then it is not hard to see that $\mathrm{Det}(A) = 0$ for $r \geqslant 2$, whereas $\det_p(A) = 1$. It is left as a problem to construct a matrix A with $\mathrm{Det}(A) = 1$ and $\det_p(A) = 0$.

For 3^{r+2}, $r \geqslant 1$, and for larger side lengths, however, it should not hard to show that they are inequivalent, for example, because the Hasse invariant for cubic curves is not equivalent to the discriminant of the plane cubic curve. There exist non-singular cubic curves with Hasse invariant zero, and singular cubics with Hasse invariant non-zero.

There is a direct way to calculate $\det_p$ as follows. This provides independent algebraic verification that $\det_p$ is invariant under the symmetric group of $(r+2)!$ symmetries of the hypercube.

THEOREM 6.1. *Let A be an $(n+1)^{r+2}$ matrix over a field of characteristic p. Denote the subscripts of A by $i = (i_0, \ldots, i_{r+1})$. Thus $A = (a_i)$, $i \in I$, where I is the set of all possible $(n+1)^{r+2}$ subscripts. Then $\det_p(A)$ is the polynomial of degree $(n+1)(p-1)$:*

$$det_p(A) = (-1)^{n+1} \sum_e \prod_{i \in I} a_i^{e(i)}/e(i)!,$$

where the sum is over 'exponent' functions $e : I \to \{0, 1, \ldots, p-1\}$, and where for all 'slices' $j = 0, \ldots, r+1$, and $k = 0, \ldots, n$, we have $\sum_{i \in I, i_j = k} e(i) = p-1$. Thus, the coefficient of the monomial corresponding to e is the non-zero element of $GF(p)$ given by

$$\alpha_e := (-1)^{n+1} \Big/ \prod_i e(i)!.$$

(Naturally $0! := 1$.)

PROOF: We show this by induction on the dimension r. Firstly, for $r = -1$ the formula holds since $\det_p(A)$, where A is the vector $(a_0, \ldots, a_n)$, equals

$$a_0^{p-1} \ldots a_n^{p-1} = (-1)^{n+1} \frac{a_0^{p-1}}{(p-1)!} \cdots \frac{a_n^{p-1}}{(p-1)!},$$

since $(p-1)! \equiv -1 \pmod{p}$ by Wilson's theorem. The only exponent function satisfying the condition that each slice of the vector has weight $p-1$ is $e(i) := p-1$, $i = 0, \ldots, n$.

Next, assume that the formula is true for some $r \geqslant -1$. We consider the formula for an $(n+1)^{r+3}$ hypercube $A = (a_{im})$, where the first $r+2$ directions of A are indexed by $i \in I$, where $|I| = (n+1)^{r+2}$, and the last direction is indexed by $m \in N := \{0, \ldots, n\}$.

https://doi.org/10.1017/S0004972700031890 Published online by Cambridge University Press

Then $\det_p(A)$ is the coefficient of $(x_0 \dots x_n)^{p-1}$ in $\det_p(Ax)$. Using the formula for Ax, which is an $(n+1)^{r+2}$ hypercube, we have

$$\det_p(Ax) = (-1)^{n+1} \sum_e \prod_{i \in I} \Big(\sum_{m=0}^{n} a_{im} x_m \Big)^{e(i)} \Big/ e(i)!,$$

where the e are exponent functions satisfying the conditions of the theorem. This can be expanded using the multinomial theorem to give

$$(-1)^{n+1} \sum_e \prod_{i \in I} \sum_{\lambda_{i0}, \dots, \lambda_{in}} \binom{e(i)}{\lambda_{i0}, \dots, \lambda_{in}} \prod_{m=0}^{n} (a_{im} x_m)^{\lambda_{im}} / e(i)!,$$

where $\sum_{m=0}^{n} \lambda_{im} = e(i)$, and $\lambda_{im} \geqslant 0$. Using $\binom{e(i)}{\lambda_{i0}, \dots, \lambda_{in}} = e(i)!/(\lambda_{i0}! \dots \lambda_{in}!)$ and cancelling $e(i)!$ we obtain

$$(-1)^{n+1} \sum_e \prod_{i \in I} \sum_{\lambda_{i0}, \dots, \lambda_{in}} \prod_{m=0}^{n} \frac{a_{im}^{\lambda_{im}}}{\lambda_{im}!} x_m^{\lambda_{im}}.$$

Evaluating the coefficient of $(x_0 \dots x_n)^{p-1}$ in this we obtain

$$(-1)^{n+1} \sum_e \sum_{\lambda} \prod_{i \in I, m \in N} \frac{a_{im}^{\lambda_{im}}}{\lambda_{im}!},$$

where for each e, the $\lambda := (\lambda_{i0}, \dots, \lambda_{in})$ are in a set of exponents that satisfy $\sum_i \lambda_{im} = p-1$, for all $m = 0, \dots, n$, and $\sum_{m=0}^{n} \lambda_{im} = e(i)$, for all $i \in I$. Also $\sum_{i \in I, i_j = k, m \in N} \lambda_{im} = \sum_{i \in I, i_j = k} \sum_{m \in N} \lambda_{im} = \sum_{i \in I, i_j = k} e(i) = p-1$. Thus the coefficient equals

$$(-1)^{n+1} \sum_f \prod_{i' \in I'} \frac{a_{i'}^{f(i')}}{f(i')!},$$

where the sum is over 'exponent' functions $f : I' := I \times N \to \{0, 1, \dots, p-1\}$, and where for all 'slices' $j = 0, \dots, r+2$, and $k = 0, \dots, n$, we have $\sum_{i' \in I', i'_j = k} f(i') = p-1$, where $i' := im$ and $f(i') := \lambda_{im}$. This is the theorem in the case of an $(r+3)$-dimensional hypercube A. □

https://doi.org/10.1017/S0004972700031890 Published online by Cambridge University Press

References

[1] E. Assmus and J.D. Key, *Designs and their Codes* (Cambridge University Press, Cambridge, 1993).

[2] A. Cayley, 'On the theory of linear transformations', *Cambr. Math. J.* **4** (1845), 193–209.

[3] A. Cayley, 'On linear transformations', *Cambr. and Dublin Math. J.* **1** (1846), 104–122.

[4] A. Cayley, 'Note sur un système de certaines formules', *J. Reine Angew. Math.* **39** (1850), 14–15.

[5] A. Cayley, *Collected Mathematical Papers* **1** (Cambridge University Press, Cambridge, 1889).

[6] J.L. Coolidge, *A history of geometrical methods* (Oxford University Press, Oxford, 1940).

[7] L.E. Dickson, *History of the theory of numbers*, **III**, Chapter 19 (Chelsea Publishing Co., New York, 1966). (See also Vol I, pp.231–233, with a report on A. Hurwitz' work).

[8] C.F. Gauß, *Disquisitiones arithmeticae*, 1801.

[9] D.G. Glynn, 'On cubic curves in projective planes of characteristic two', *Australas. J. Combin.* **17** (1998), 1–20.

[10] D.G. Glynn and J.W.P. Hirschfeld, 'On the classification of geometric codes by polynomial functions', *Des. Codes Cryptogr.* **6** (1995), 189–204.

[11] I.M. Gelfand, M.M. Kapranov and A.V. Zelevinsky, *Discriminants, resultants and multidimensional determinants* (Birkhäuser, Boston, Basel, Berlin, 1994).

[12] R. Hartshorne, *Algebraic geometry* (Springer, Berlin, Heidelberg, New York, 1983), **pp.** 332–340.

[13] J.W.P. Hirschfeld, *Projective geometry over a finite field* (Oxford University Press, Oxford, 1979).

[14] L. Schläfli, 'Über die Resultante eines Systemes mehrerer algebraischen Gleichungen', *Denkschr. der Kaiserlicher Akad. der Wiss., math-naturwiss. Klasse* **4** (1852).

[15] L. Schläfli, *Gesammelte Abhandlungen* **2** (Birkhäuser Verlag, Basel, 1953).

Te Tari Tatau
Te Whare Wānanga o Waitaha
P.B. 4800
Ōtautahi
Aotearoa (New Zealand)
e-mail: D.Glynn@math.canterbury.ac.nz