J. Fac. Sci. Univ. Tokyo
Sect. IA, Math.
34 (1987), 699-707.

On Dickson's theorem on invariants

To my friend Nagayoshi Iwahori

By Robert STEINBERG

1. Introduction.

The theorem in question, first proved in [3], is as follows.

THEOREM A. *Let $G=GL_n(k)$ $(k=\boldsymbol{F}_q)$ act on $k[X_1, X_2, \cdots, X_n]$, the algebra of formal polynomials, in the usual way. Then $k[X_1, X_2, \cdots, X_n]^G$, the algebra of invariants, is a polynomial algebra on the generators*

$$I_r=[01\cdots\hat{r}\cdots n]/[01\cdots n-1] \qquad (0\leqq r\leqq n-1).$$

Here, for any nonnegative integers, $[e_1e_2\cdots e_n]$ denotes the determinant of the matrix whose ij entry is the q^{e_j}th power of X_i. This is nonzero if the e's are distinct, since the main diagonal term is not cancelled by any other term. Also since it is reproduced by each T in $GL_n(k)$ with the scalar factor $\det(T)$, by an easy calculation using Fermat's theorem that $c^q=c$ for each c in k, it follows that every $[e_1\cdots e_n]$ is $SL_n(k)$-invariant and that the ratio of any two such is $GL_n(k)$-invariant. The I_r's are called the Dickson invariants.

Simultaneously with Theorem A we shall consider the following result which we believe to be new.

THEOREM B. *$k[X_1, X_2, \cdots, X_n]$ is free as a module over $k[X_1, X_2, \cdots, X_n]^G$ with a basis consisting of the monomials $X_1^{i_1}X_2^{i_2}\cdots X_n^{i_n}$ $(0\leqq i_r<q^n-q^{r-1}$ for all $r)$.*

Dickson's original proof of Theorem A was very complicated. Since then a number of other proofs have appeared (see Bourbaki [2, p. 137-8], Ore [7, p. 566-8] and Wilkerson [11]), some of them quite simple. In Section 2 we present a proof which, we believe, is simpler than any of these and has the additional advantage that it yields Theorems A and B together. It should be mentioned that the polynomial nature of $k[X]^G$ and the freeness of $k[X]$ over $k[X]^G$ are equivalent under quite general conditions, but the proof involves some serious commutative algebra (see [2]). The essential step of our proof is a relative version of the above results (Theorem E below) which at once yields Theorems A and B and analogous results for many of its subgroups.

In [3] Dickson also proved the following result.

THEOREM C. *If $G=SL_n(k)$ in Theorem* A *then the conclusion there holds with I_0 replaced by $I_0^{1/(q-1)}=[01\cdots n-1]$.*

The equality here follows from $I_0=[12\cdots n]/[01\cdots n-1]=[01\cdots n-1]^q/[01\cdots n-1]$. As is known Theorem C follows easily from Theorem A (and vice versa), and we have included a proof of this fact.

The invariants that depend on several formal vectors have not yet been determined, except in case $n=2$, q is a prime, and there are just two vectors, by Krathwohl [5]. In Section 3 we present some contributions to this problem which lead to yet another proof of Dickson's Theorem.

Finally we consider the situation over $\boldsymbol{Z}/p^r\boldsymbol{Z}$. Here the invariants have been obtained by Feldstein [4], following Turner [10] who did the case $r=2$. In Section 4 we obtain more general results in a simpler way which makes transparent the transition from mod p invariants to mod p^r invariants for $GL_n(\boldsymbol{Z}/p^r\boldsymbol{Z})$ and many of its subgroups.

2. Dickson's theorem.

In this section we shall prove Theorems A, B and C. Our approach is that of Artin [1, p. 39-41] who considered the invariants of the symmetric group S_n.

THEOREM D. *For $0\leqq r\leqq n$ let $G(r)$ be the subgroup of $GL_n(k)$ consisting of the matrices that agree with the identity in the first r rows.*

(a) *$k[X]^{G(r)}$ is generated by $X_1,\cdots,X_r,I_r,\cdots,I_{n-1}$, hence is a polynomial algebra over k.*

(b) *For $r\geqq 1$ $k[X]^{G(r)}$ is free as a module over $k[X]^{G(r-1)}$ with $\{X_r^i \mid i<q^n-q^{r-1}\}$ as a basis.*

Observe that $G(n)=1$ and $G(0)=GL_n(k)$ and that the proposed generators in (a) all belong to $k[X]^{G(r)}$. Part (a) with $r=0$ yields Theorem A, while part (b) for the values of r from 1 to n combined yields Theorem B as well as an analogous theorem for each of the groups $G(r)$. We start our proof with two simple lemmas.

LEMMA 1. *The polynomial $T^{q^n}-I_{n-1}T^{q^{n-1}}+I_{n-2}T^{q^{n-2}}-\cdots+(-1)^nI_0T$ has as its roots the q^n distinct linear forms in $X_1,X_2,\cdots,X_n$.*

Let $[01\cdots n]$ be defined as $[01\cdots n-1]$ above but with an extra variable X_{n+1} and an extra exponent q^n. As a polynomial in X_{n+1} its degree is q^n since the highest coefficient $[01\cdots n-1]$ is, as noted above, nonzero. If X_{n+1} is replaced in $[01\cdots n]$ by any linear combination of the first n X's then the $(n+1)$th row is replaced by the same linear combination of the first n rows. It follows that the roots are just the q^n linear forms in the first $n\,X$'s. If we expand $[01\cdots n]$ along the $(n+1)$th row, divide through by $[01\cdots n-1]$ and then replace X_{n+1} by T, we get the polynomial of Lemma 1, as required.

An immediate consequence of this lemma is that the I's are all polynomials.

LEMMA 2. *For each r, X_r is a root of a monic polynomial of degree q^n-q^{r-1} over $k[X_1,\cdots,X_{r-1},I_{r-1},\cdots,I_{n-1}]$.*

If F_n is the polynomial of Lemma 1 and $F_{r\;1}$ is defined similarly so that its roots are the q^{r-1} linear forms in the first $r-1$ X's, then F_n/F_{r-1} is the required polynomial. First it is monic of degree q^n-q^{r-1} and has X_r as a root. Further its coefficients are in $k[X_1,\cdots,X_{r-1},I_1,\cdots,I_n]$ since those of F_{r-1} are in $k[X_1,\cdots,X_{r-1}]$ and those of F_n are in $k[I_1,\cdots,I_n]$, by Lemma 1. Finally, the I_s $(s<r-1)$ can be dropped since their degrees are all larger than that of F_n/F_{r-1}.

We come now to the heart of the proof, the deduction of Theorem D from Lemma 2. Since the argument is valid in a number of other cases of interest to us, we carry it out in a more general context which focuses attention on the essential features of the situation.

THEOREM E. *Let G and H be subgroups of $GL_n(k)$ with H contained in G. Let S be a subset of $k[X]^H$ and T a subset of $k[X]^G$ such that* (1) *S generates $k[X]^H$,* (2) *S contains just one element, Y, which is not in T, and* (3) *Y is a root of a polynomial over $k[T]$ which is monic and of degree at most $|G/H|$. Then T generates $k[X]^G$, and $k[X]^H$ is free as a module over $k[X]^G$ with $\{Y^i\,|\,0\leqq i<|G/H|\}$ as a basis.*

Here k can be any field as long as G is finite. Let A be the polynomial of (3) and F any element of $k[X]^H$. Then F is in $k[S]$ by (1), and by using the equation $A(Y)=0$ we can write F as a polynomial in Y of degree less than $|G/H|$, with coefficients in $k[T]$ because of the assumption (2). Thus the given Y^i's generate $k[X]^H$ as a module over $k[T]$, hence also $k[X]^H$ over $k[X]^G$, and $k(X)^H$ over $k(X)^G$, the last because for any $P(X)/Q(X)$ in $k(X)^H$ the denominator can always be converted to one in $k[X]^G$, namely,

$\Pi g\cdot Q(X)$ (g in G). However the number of Y^i's is at most $|G/H|$, which by the first theorem of Galois theory equals the dimension of $k(X)^H$ as a vector space over $k(X)^G$ (see [1, Th. 14]). It follows that the Y^i's form a basis for this space and hence a free generating set for $k[X]^H$ over $k[X]^G$. It remains to show that each F in $k[X]^G$ belongs to $k[T]$. Write $F=\Sigma C_i Y^i$ ($i<|G/H|$) as a linear combination over $k[T]$, and $F=FY^0$, over $k[X]^G$. Since the Y^i's are free over $k[X]^G$, we get $F=C_0$, an element of $k[T]$, as required.

Labelling the two parts of Theorem D (a_r) and (b_r), we shall now show that for $r\geqq 1$ (a_r) implies (b_r) and (a_{r-1}). Then since (a_n) is obviously true, the theorem will follow. The hypotheses of Theorem E are satisfied with $G=G(r-1)$, $H=G(r)$, $S=\{X_1,\cdots,X_r,I_r,\cdots,I_{n-1}\}$, $T=\{X_1,\cdots,X_{r-1},I_{r-1},\cdots,I_{n-1}\}$ and $Y=X_r$, in view of Lemma 2 and the assumption (a_r). We conclude that (b_r) and (a_{r-1}) hold, except for the last point of (a_{r-1}). Since $k[X]$ is algebraic over $k[X]^{G(r-1)}$, the transcendence degree of the latter over k equals that of the former, which is n, so that that point also holds and the proof of Theorem D is complete.

REMARKS. (a) If X_1 is replaced by 1 the group G_1 above becomes the affine group on the remaining coordinates. It follows that for this group also the invariants form a polynomial algebra, with generators obtained from the I_r's ($r\neq 0$) by the same replacement. (b) The above method, as embodied in Theorem E, also works for all parabolic subgroups of $GL_n(k)$ and their unipotent radicals. To indicate the results obtained we state them in a simple, but typical, case. Let $n=4$ and let P be the parabolic subgroup in which the 2 by 2 block in the upper right hand corner is required to be 0. Then $k[X]^P$ is generated by I_0', I_1', I_2, I_3, with the first two I's calculated on the space of the first two coordinates, and a basis for $k[X]/k[X]^P$ consists of the monomials with $0\leqq i_r<q^2-q^{r-1}$ for $r=1$ and 2 and $0\leqq i_r<q^4-q^{r-1}$ for $r=3$ and 4.

We come now to the proof of Theorem C. For this we need another (well-known) lemma.

LEMMA 3. $[01\cdots n-1]$ *equals the product of all of the nonzero linear forms in the X's for which the last nonzero coefficient is* 1, *and it divides every $SL_n(k)$-invariant which X_1 divides.*

Any such invariant is divisible by all of the transforms of X_1 under $SL_n(k)$, hence by all of the linear forms as above and hence also by their product. But this product and $[01\cdots n-1]$ have the same lowest terms (lexicographically). Thus they are equal and the lemma follows.

It also follows that $[01\cdots n-1]$ divides every $[e_1e_2\cdots e_n]$, which provides another proof that the I_r's are polynomials.

Now let F be any invariant for $SL_n(k)$. We must show that it is expressible in terms of the n invariants given by Theorem C. The elements $T_c=\text{diag}(c, 1, 1, \cdots, 1)$ (c in k^*) form a system of representatives for the cosets of $GL_n(k)$ over $SL_n(k)$. If we write $F=-\sum T_cF+\sum(T_cF-F)=F_1+F_2$, say, it follows that F_1 is an invariant for $GL_n(k)$ and hence by Theorem A is expressible as required by Theorem C. From the form of F_2 it is divisible by X_1 and hence also by $[01\cdots n-1]$ by Lemma 3. We can now finish by induction since the degree of $F_2/[01\cdots n-1]$ is less than that of F.

3. Several vectors.

In this section we consider $GL_n(k)$ acting on several vectors. The results are fragmentary, but the ideas introduced may be of further use.

THEOREM F. *If the group $G=GL_n(k)$ acts on a set of $m\geqq n$ independent formal vectors whose coordinates are viewed as functions on $\bar{k}^{mn}$ and written in matrix form as $v_1v_2\cdots v_m=XY$ with X consisting of the first n v's and Y of the others, then on the set where* $\det X=[X]\neq 0$, *or more generally on any locally closed G-invariant subset of this set, the map $f: XY\to(X^{-1}X^{(q)})(X^{-1}Y)$ defines a quotient for the action.*

Here $X^{(q)}$ denotes the matrix of qth powers of the entries of X. The Lang-Speiser Theorem states that the map $X\to X^{-1}X^{(q)}$ on $GL_n(\bar{k})$ is surjective. In [9] a refinement is proved, that this map is a finite morphism. From this the theorem readily follows.

THEOREM G. *In addition to the above notation let $[X_{ij}]$ denote the quantity obtained from $[X]$ by replacing v_i by v_j^q if $j\leqq n$, by v_j if $j>n$. Then on the set in $\bar{k}^{mn}$ where $[X]\neq 0$ the algebra of polynomials invariant under G is generated by the mn elements $[X_{ij}]/[X]$ together with $[X]^{1-q}$.*

By Theorem F we need only work out the coordinates of $X^{-1}X^{(q)}$, $X^{-1}Y$ and $[X^{-1}X^{(q)}]^{-1}$ in terms of those of X and Y. By Cramer's rule for X^{-1} the results are as stated.

This result also holds for $SL_n(k)$ provided that $[X]^{1-q}$ is replaced by $[X]^{-1}$.

COROLLARY. *If $m\geqq n$, then $k(v_1v_2\cdots v_m)^G$ is purely transcendental over k with the mn $[X_{ij}]/[X]$'s as a generating set.*

This follows at once from Theorem G.

For $m<n$ analogous results may be obtained as follows. We first expand $v_1 \cdots v_m$ to a square matrix X by adjoining the q^jth powers of v_m for $j=1, \cdots, n-m$. (We could also adjoin powers of several v's.) With this X and the resulting n^2+1 functions of Theorem G which can be brought down to $mn+1$ since those with $m \leqq j<n$ are constant, the result there holds. Finally, dropping the function $[X]^{1-q}$, we see that the corollary also holds, all for $m<n$.

To solve our main problem, the determination of $k[v_1 \cdots v_m]^G$, we would, according to the present approach, have to remove the condition $[X] \neq 0$, that is, determine the polynomials in the $[X_{ij}]$'s that are divisible by $[X]$, and we can not do this. We can, however, do this in case $m=1$ and thus obtain another proof of Dickson's Theorem, with which we close this section.

Let v be any formal vector. We expand it to a square matrix X by adjoining $n-1$ of its powers as in the preceding proof. The $n+1$ nonconstant invariants that result from Theorem G in this case are just the Dickson invariants $I_0, \cdots, I_{n-1}$ together with I_0^{-1}, in terms of the coordinates $X_1, \cdots, X_n$ of v. These therefore generate the invariant polynomials on the open set in $\bar{k}^n$ where $I_0 \neq 0$. Let F be any element of $k[X]^G$. Then F is a polynomial in the I's and I_0^{-1}, so that for some $m \geqq 0$, $I_0^m F$ is a polynomial in the I's and is thus expressible as $\sum H_r(I_1, \cdots, I_{n-1}) I_0^r$ with each H_r a polynomial. Assuming m to be minimal we must show that $m=0$. Suppose not. Then $H_0 \neq 0$. Since X_n divides I_0, H_0 vanishes at $X_n=0$. However, as easily follows from the definitions, the substitution $X_n=0$ converts the I_r's $(1 \leqq r<n-1)$ into the qth powers of the $(n-1)$-dimensional Dickson invariants, and these are algebraically independent over k, by what has been said above (or else by a direct proof by induction). Thus $H_0=0$, a contradiction. This shows that $r=0$ and thus that F is a polynomial in the I's, as required.

4. Invariants mod p^r.

One of our main theorems in this area, extended by the remarks that follow its proof, is the following.

THEOREM H. *Let H be the subgroup of $GL_n(\boldsymbol{Z}/p^r\boldsymbol{Z})$ $(n \geqq 2, r \geqq 1)$ consisting of the matrices that are unimodular and congruent to the unit matrix* mod p. *Then F in $\boldsymbol{Z}/p^r\boldsymbol{Z}[X_1, X_2, \cdots, X_n]$ is invariant under H if and only if it is expressible in one of the following forms.*

(a) *$F=\sum p^i F_i$ with F_i a polynomial in the p^{r-i-1}th powers of*

$X_1, X_2, \cdots, X_n$ for each $i=0,1,2,\cdots,r-1$.

(b) *$F=\sum c_I X^I$ with p^{r-1} dividing $c_I I$ for every multi-index $I=(i_1,\cdots,i_n)$.*

Since the sufficiency in (a) and (b) are easily verified, we turn to the proof of necessity, the one in (b), since that yields the one in (a) at once. Let F be an invariant for H written as the sum of its homogeneous parts relative to the total degree in X_1 and X_2. If we apply $T_{12}:(X_1\rightarrow X_1+pX_2, X_j\rightarrow X_j$ if $j\neq 1)$ to F then each of these parts is kept fixed. If $F_m=\sum A_i X_1^i X_2^{m-i}$ (with each A_i a polynomial in $X_3, X_4, \cdots$) is the part of degree m then $T_{12}F_m=F_m$ yields, when like terms are grouped together, $\sum_{j=0}^{m-1} X_1^j X_2^{m-j} \sum_{0<k\leqq m-j} p^k \binom{j+k}{j} A_{j+k}=0$. Here each of the inner sums must be 0. The first term in the jth sum is $p(j+1)A_{j+1}$ and the later terms may be written as $(p^{k-1}/k)\binom{j+k-1}{j}p(j+k)A_{j+k}$. Here each p^{k-1}/k is an integer $\bmod p$ since $p^{k-1}\geqq k$ so that p can divide k at most $k-1$ times. It follows by downward induction that p^r divides piA_i, that is, p^{r-1} divides iA_i, for all i. In terms of F written as in (b) this means that p^{r-1} divides every $c_I i_1$. Similarly this holds with i_1 replaced by $i_2, i_3, \cdots$, which proves the necessity in (b) and hence the theorem.

REMARK. The group H in its entirety is not needed for the above proof of necessity. If T_{ij} $(i\neq j)$ is defined as above then one could replace H by the subgroup generated by any set of T_{ij}'s such that the index i takes on all values from 1 to n, for example, T_{12} and the set T_{i1} $(2\leqq i\leqq n)$.

REMARK. A similar theorem holds for the group of unimodular matrices that are congruent to the unit matrix $\bmod p^s$ $(0<s<r)$. It is only necessary to replace p^{r-1} by p^{r-s} in (b) and make an analogous change in (a). The proof is essentially the same, with the previous remark still applying.

THEOREM I. *Let G be any subgroup of $GL_n(\boldsymbol{Z}/p^r\boldsymbol{Z})$ $(n\geqq 2, r\geqq 1)$ which contains the subgroup H of Theorem H, or, more generally, any subgroup of H for which the conclusion there holds (see the first remark above). Let $\{J_1,\cdots,J_m\}$ be a generating set for the algebra of polynomial invariants in the induced action of G on $(\boldsymbol{Z}/p\boldsymbol{Z})^n$. Then F is an invariant for G in the original action if and only if it can be written $F=\sum p^i F_i$ with F_i a polynomial in the p^{r-i-1}th powers of the J's for each $i=0,1,\cdots,r-1$.*

Again the sufficiency is easily verified; thus we turn to the proof of necessity. Let F be a G-invariant polynomial. Then by Theorem H it can be written, mod p, as a polynomial in the p^{r-1}th powers of the X's. Now G acts, mod p, on these powers of the X's and on the X's themselves by exactly the same formulas. It follows that, mod p, F is a polynomial in the J's evaluated at the p^{r-1}th powers of the X's, which, mod p, is the same polynomial in the p^{r-1}th powers of the J's. In other words, $F=F_0+pF'$ with F_0 as in the conclusion of the theorem and F' a polynomial. Further by the sufficiency, which has already been established, F_0 is an invariant and hence pF' is also. Thus F' is an invariant mod p^{r-1}, and the proof may be completed by induction.

THEOREM J. *A polynomial is invariant under $GL_n(\boldsymbol{Z}/p^r\boldsymbol{Z})$ if and only if it can be written $\sum p^i F_i$ with each F_i a polynomial in the I's, as in Theorem* A *but with q replaced by p, and similarly for SL_n with I_0 replaced by* $[01 \cdots n-1]$.

This result follows at once from Theorems A, C and I.

As they are stated above, Theorems H and I also apply to most parabolics and, in case n is even, to the symplectic group. With minor modifications they apply to all parabolics and their unipotent radicals and the groups G_r of Section 2.

Using the same methods, we have obtained a version of the above results for a class of local rings which includes $\boldsymbol{F}_q$ and $\boldsymbol{Z}/p^r\boldsymbol{Z}$ as special cases and thus a common generalization of Theorems A and J. Many other cases that we have worked out lead us to the conjecture that such results should hold for arbitrary finite local rings.

References

[1] Artin, E., Galois theory, Notre Dame Math. Lectures 2, Second Edition, NAPCO Inc., 1959.

[2] Bourbaki, N., Groupes et Algèbres de Lie, Ch. 4, 5 et 6, Hermann, Paris, 1968.

[3] Dickson, L. E., A fundamental system of invariants of the general modular linear group with a solution of the form problem, Trans. Amer. Math. Soc. **12** (1911), 75-98.

[4] Feldstein, M. M., Invariants of the linear group modulo p^k, Trans. Amer. Math. Soc. **25** (1923), 223-238.

[5] Krathwohl, W. C., Modular invariants of two pairs of cogredient variables, Amer. J. Math. **36** (1914), 449-460.

[6] Mui, H., Modular invariant theory and the cohomology algebras of symmetric groups, J. Fac. Sci. Univ. Tokyo Sect. IA Math. **22** (1975), 319-369.

[7] Ore, O., On a special class of polynomials, Trans. Amer. Math. Soc. **35** (1933),

559-584.

[8] Serre, J.-P., Groupes finis d'automorphismes d'anneaux locaux reguliers, Colloque d'Algèbre, ENSJF (1967), Exp. 8.

[9] Steinberg, R., On theorems of Lie-Kolchin, Borel and Lang, Contributions to Algebra (dedicated to Ellis Kolchin), Academic Press, New York, 1977.

[10] Turner, J.S., A fundamental system of invariants of a modular group of transformations, Trans. Amer. Math. Soc. **24** (1922), 129-134.

[11] Wilkerson, C., A primer on Dickson invariants, Amer. Math. Soc. Contemp. Math. **19** (1983), 421-434.

(Received May 1, 1987)

MSRI
1000 Centennial Drive
Berkeley, CA 94720
U.S.A.

and

Department of Mathematics
University of California
Los Angeles, CA 90024
U.S.A.