Enumerating algebras over a finite field

Michael Vaughan-Lee

November 2012

1 Introduction

Graham Higman wrote two immensely important and influential papers on enumerating p-groups in the late 1950s. The papers were entitled *Enumerating p-groups* I and II, and were published in the Proceedings of the London Mathematical Society in 1960 (see [2] and [3]). In the first of these papers Higman proves that if we let $f(p^n)$ be the number of p-groups of order p^n, then

$$p^{\frac{2}{27}n^2(n-6)} \leq f(p^n) \leq p^{(\frac{2}{15}+\varepsilon_n)n^3},$$

where ε_n tends to zero as n tends to infinity. In the second of the two papers Higman formulated his famous PORC conjecture concerning the form of the function $f(p^n)$. He conjectured that for each n there is an integer N (depending on n) such that for p in a fixed residue class modulo N the function $f(p^n)$ is a polynomial in p. For example, for $p \geq 5$ the number of groups of order p^6 is

$$3p^2 + 39p + 344 + 24\gcd(p-1,3) + 11\gcd(p-1,4) + 2\gcd(p-1,5).$$

(See [5].) So for $p \geq 5$, $f(p^6)$ is one of 8 polynomials in p, with the choice of polynomial depending on the residue class of p modulo 60. The number of groups of order p^6 is **P**olynomial **O**n **R**esidue **C**lasses. In the same paper Higman proved that, for any given n, the function enumerating the number of p-class 2 groups of order p^n is a PORC function of p. He obtained this result as a corollary to a very general theorem about vector spaces acted on by the general linear group. As another corollary to this general theorem, he also proved that for any given n the function enumerating the number of algebras of dimension n over the field of q elements is a PORC function of q. In this note we give the PORC formulae for the numbers of algebras of dimensions 2, 3 and 4 over GF(q), and we give an outline of how Higman's methods can be used to compute these formulae. I have written a survey article [7] entitled *Graham Higman's PORC conjecture*, and this article contains a detailed calculation of the formulae for the number of algebras of dimension 2 over GF(q). It must be said that the formulae themselves tell you very little apart from the fact that they can be very complicated. In particular the formulae for the number of algebras of dimension 4 might be considered to be some sort of weird joke in rather

poor taste. But nevertheless I had good fun obtaining them. A complete description of the algebras of dimension 2 over a finite field is given by Petersson and Scherer [6]. They also give formulae for the number of algebras of dimension 2 over GF(q), which agree with the formulae given here.

We give the formulae for the numbers of algebras of dimensions 2, 3 and 4 in Section 2. In Section 3 we give a broad outline of how the formulae can be obtained and in Section 4 we draw some quite precise conclusions about the asymptotic form of the formulae for general n.

2 The numbers of algebras of dimensions 2, 3, 4

By an algebra over GF(q) we mean a vector space over GF(q) with a bilinear product. There is no requirement for the product to satisfy any further conditions such as associativity or commutativity. Such general algebras are often called *non-associative algebras*.

Let $g(n,q)$ denote the number of algebras of dimension n over GF(q) . For $n=2$, $g(n,q)$ is given by one of three polynomials in q of degree 4. If q is a power of 2 then

$$g(2,q) = q^4 + q^3 + 4q^2 + 3q + 6.$$

If q is a power of 3 then

$$g(2,q) = q^4 + q^3 + 4q^2 + 4q + 6.$$

And if q is a power of p with $p > 3$ then

$$g(2,q) = q^4 + q^3 + 4q^2 + 4q + 7.$$

For $n=3$, $g(n,q)$ is again given by one of three formulae depending on whether q is a power of 2, or a power of 3, or a power of p for some $p > 3$. But now the formulae are not straightforward polynomials. If q is a power of 2 then

$$\begin{aligned} g(3,q) &= q^{18} + q^{17} + 2q^{16} + 3q^{15} + 4q^{14} + 5q^{13} + 7q^{12} + 8q^{11} + 10q^{10} + 13q^9 \\ &\quad +17q^8 + 19q^7 + 24q^6 + 26q^5 + 31q^4 + 28q^3 + 24q^2 + 20q + 20 \\ &\quad +(q^5 + 2q^4 + 2q^3 + 2q^2 + 3q + 3)\gcd(q-1,3) \\ &\quad +(q+1)\gcd(q-1,5). \end{aligned}$$

If q is a power of 3 then

$$\begin{aligned} g(3,q) &= q^{18} + q^{17} + 2q^{16} + 3q^{15} + 4q^{14} + 5q^{13} + 7q^{12} + 8q^{11} + 10q^{10} + 13q^9 \\ &\quad +19q^8 + 23q^7 + 31q^6 + 37q^5 + 43q^4 + 44q^3 + 38q^2 + 31q + 25 \\ &\quad +(q^4 + q^3 + 2q^2 + 3q + 2)\gcd(q-1,4) \\ &\quad +(q+1)\gcd(q-1,5) + (\gcd(q^3-1,7) + 2)/3. \end{aligned}$$

And if q is a power of p with $p > 3$ then

$$\begin{aligned} g(3,q) &= q^{18} + q^{17} + 2q^{16} + 3q^{15} + 4q^{14} + 5q^{13} + 7q^{12} + 8q^{11} + 10q^{10} + 13q^9 \\ &\quad +19q^8 + 23q^7 + 31q^6 + 36q^5 + 42q^4 + 43q^3 + 37q^2 + 29q + 23 \\ &\quad +(q^5 + 2q^4 + 2q^3 + 2q^2 + 3q + 4)\gcd(q-1,3) \\ &\quad +(q^4 + q^3 + 2q^2 + 3q + 2)(q-1,4) \\ &\quad +(q+1)\gcd(q-1,5) + (\gcd(q^3-1,7)+2)/3. \end{aligned}$$

For $n = 4$, $g(n,q)$ is given by one of four formulae, depending on whether q is a power of 2 or 3 or 5, or a power of p with $p > 5$. As mentioned above, the formulae are horrendous! If q is a power of 2 then $g(4,q)$ equals

$$\begin{aligned} &q^{48} + q^{47} + 2q^{46} + 3q^{45} + 5q^{44} + 6q^{43} + 9q^{42} + 11q^{41} + 15q^{40} + 18q^{39} + 23q^{38} \\ &+27q^{37} + 34q^{36} + 39q^{35} + 47q^{34} + 54q^{33} + 64q^{32} + 72q^{31} + 84q^{30} + 94q^{29} \\ &+108q^{28} + 120q^{27} + 137q^{26} + 152q^{25} + 175q^{24} + 192q^{23} + 217q^{22} + 238q^{21} \\ &+265q^{20} + 287q^{19} + 317q^{18} + 341q^{17} + 372q^{16} + 397q^{15} + 424q^{14} + 446q^{13} \\ &+469q^{12} + 477q^{11} + 485q^{10} + 480q^9 + 474q^8 + 450q^7 + 422q^6 + 362q^5 \\ &+301q^4 + 229q^3 + 168q^2 + 131q + 91 \\ &+\gcd(q-1,3)(2q^{15} + 5q^{14} + 11q^{13} + 19q^{12} + 30q^{11} + 44q^{10} + 60q^9 \\ &+78q^8 + 97q^7 + 109q^6 + 111q^5 + 111q^4 + 114q^3 + 111q^2 + 91q + 43) \\ &+\gcd(q-1,5)(q^9 + 3q^8 + 7q^7 + \frac{21}{2}q^6 + 13q^5 + \frac{35}{2}q^4 + 25q^3 + \frac{63}{2}q^2 + 36q + 18) \\ &+\frac{1}{2}\gcd(q+1,5)(q^6 + q^4 + q^2 - 2) \\ &+\gcd(q-1,7)(q^4 + 4q^3 + 9q^2 + 13q + 8) \\ &+\gcd(q-1,9)(q^2 + 4q + 3) \\ &+\gcd(q-1,11)(q+1) \\ &+\frac{1}{2}\gcd(q-1,3)(\gcd(q-1,5) + \gcd(q+1,5)). \end{aligned}$$

When q is a power of 3 then $g(4,q)$ equals

$$
\begin{aligned}
&q^{48}+q^{47}+2q^{46}+3q^{45}+5q^{44}+6q^{43}+9q^{42}+11q^{41}+15q^{40}+18q^{39}+23q^{38}\\
&+27q^{37}+34q^{36}+39q^{35}+47q^{34}+54q^{33}+64q^{32}+72q^{31}+84q^{30}+94q^{29}\\
&+108q^{28}+120q^{27}+137q^{26}+153q^{25}+177q^{24}+197q^{23}+226q^{22}+253q^{21}\\
&+288q^{20}+320q^{19}+364q^{18}+404q^{17}+456q^{16}+507q^{15}+564q^{14}+621q^{13}\\
&+686q^{12}+736q^{11}+787q^{10}+819q^{9}+846q^{8}+846q^{7}+815q^{6}+731q^{5}\\
&+626q^{4}+506q^{3}+383q^{2}+277q+146\\
&+\gcd(q-1,4)(q^{13}+2q^{12}+5q^{11}+11q^{10}+20q^{9}+32q^{8}+45q^{7}+57q^{6}\\
&+64q^{5}+70q^{4}+82q^{3}+86q^{2}+72q+35)\\
&+\gcd(q-1,8)(q^{6}+q^{5}+q^{4}+2q^{3}+\frac{15}{2}q^{2}+12q+\frac{13}{2})+\frac{1}{2}\gcd(q-3,8)(q^{2}+1)\\
&+\gcd(q-1,5)(q^{9}+3q^{8}+7q^{7}+\frac{21}{2}q^{6}+13q^{5}+\frac{35}{2}q^{4}+25q^{3}+\frac{63}{2}q^{2}+40q+23)\\
&+\frac{1}{2}\gcd(q+1,5)(q^{6}+q^{4}+q^{2})\\
&+\gcd(q-1,7)(q^{4}+4q^{3}+9q^{2}+13q+9)\\
&+\frac{1}{3}(\gcd(q^{3}-1,7)+2)(q^{6}+q^{5}+q^{4}+q^{3}+q^{2}+q+2)\\
&+\gcd(q-1,11)(q+1).
\end{aligned}
$$

When q is a power of 5 then $g(4,q)$ equals

$$
\begin{aligned}
&q^{48}+q^{47}+2q^{46}+3q^{45}+5q^{44}+6q^{43}+9q^{42}+11q^{41}+15q^{40}+18q^{39}+23q^{38}\\
&+27q^{37}+34q^{36}+39q^{35}+47q^{34}+54q^{33}+64q^{32}+72q^{31}+84q^{30}+94q^{29}\\
&+108q^{28}+120q^{27}+137q^{26}+153q^{25}+177q^{24}+197q^{23}+226q^{22}+253q^{21}\\
&+288q^{20}+320q^{19}+364q^{18}+404q^{17}+456q^{16}+504q^{15}+559q^{14}+610q^{13}\\
&+668q^{12}+707q^{11}+745q^{10}+763q^{9}+773q^{8}+756q^{7}+717q^{6}+628q^{5}\\
&+524q^{4}+394q^{3}+266q^{2}+183q+101\\
&+\gcd(q-1,3)(2q^{15}+5q^{14}+11q^{13}+19q^{12}+30q^{11}+44q^{10}+60q^{9}\\
&+79q^{8}+101q^{7}+117q^{6}+122q^{5}+128q^{4}+143q^{3}+153q^{2}+135q+70)\\
&+\gcd(q-1,9)(q^{2}+4q+3)\\
&+\gcd(q-1,4)(q^{13}+2q^{12}+5q^{11}+11q^{10}+20q^{9}+32q^{8}+45q^{7}+57q^{6}\\
&+64q^{5}+70q^{4}+82q^{3}+86q^{2}+72q+34)\\
&+\gcd(q-1,8)(q^{6}+q^{5}+q^{4}+2q^{3}+\frac{15}{2}q^{2}+12q+\frac{13}{2})+\frac{1}{2}\gcd(q-3,8)(q^{2}+1)\\
&+\gcd(q-1,7)(q^{4}+4q^{3}+9q^{2}+13q+9)\\
&+\frac{1}{3}(\gcd(q^{3}-1,7)+2)(q^{6}+q^{5}+q^{4}+q^{3}+q^{2}+q+2)\\
&+\gcd(q-1,11)(q+1)\\
&+\gcd(q-1,3)\gcd(q-1,4).
\end{aligned}
$$

And finally, when q is a power of p with $p > 5$ then $g(4, q)$ equals

$$
\begin{aligned}
&q^{48} + q^{47} + 2q^{46} + 3q^{45} + 5q^{44} + 6q^{43} + 9q^{42} + 11q^{41} + 15q^{40} + 18q^{39} + 23q^{38} \\
&+27q^{37} + 34q^{36} + 39q^{35} + 47q^{34} + 54q^{33} + 64q^{32} + 72q^{31} + 84q^{30} + 94q^{29} \\
&+108q^{28} + 120q^{27} + 137q^{26} + 153q^{25} + 177q^{24} + 197q^{23} + 226q^{22} + 253q^{21} \\
&+288q^{20} + 320q^{19} + 364q^{18} + 404q^{17} + 456q^{16} + 504q^{15} + 559q^{14} + 610q^{13} \\
&+668q^{12} + 707q^{11} + 745q^{10} + 761q^{9} + 771q^{8} + 749q^{7} + 706q^{6} + 615q^{5} \\
&+507q^{4} + 369q^{3} + 234q^{2} + 143q + 84 \\
&+ \gcd(q-1,3)(2q^{15} + 5q^{14} + 11q^{13} + 19q^{12} + 30q^{11} + 44q^{10} + 60q^{9} \\
&+79q^{8} + 101q^{7} + 117q^{6} + 122q^{5} + 128q^{4} + 143q^{3} + 153q^{2} + 135q + 68) \\
&+ \gcd(q-1,9)(q^2 + 4q + 3) \\
&+ \gcd(q-1,4)(q^{13} + 2q^{12} + 5q^{11} + 11q^{10} + 20q^{9} + 32q^{8} + 45q^{7} + 57q^{6} \\
&+64q^{5} + 70q^{4} + 82q^{3} + 86q^{2} + 72q + 34) \\
&+ \gcd(q-1,8)(q^6 + q^5 + q^4 + 2q^3 + \frac{15}{2}q^2 + 12q + \frac{13}{2}) + \frac{1}{2}\gcd(q-3,8)(q^2+1) \\
&+ \gcd(q-1,5)(q^9 + 3q^8 + 7q^7 + \frac{21}{2}q^6 + 13q^5 + \frac{35}{2}q^4 + 25q^3 + \frac{63}{2}q^2 + 40q + 22) \\
&+\frac{1}{2}\gcd(q+1,5)(q^6 + q^4 + q^2 - 2) \\
&+ \gcd(q-1,7)(q^4 + 4q^3 + 9q^2 + 13q + 9) \\
&+\frac{1}{3}(\gcd(q^3-1,7) + 2)(q^6 + q^5 + q^4 + q^3 + q^2 + q + 2) \\
&+ \gcd(q-1,11)(q+1) \\
&+\frac{1}{2}\gcd(q-1,3)(2\gcd(q-1,4) + \gcd(q-1,5) + \gcd(q+1,5)).
\end{aligned}
$$

3 The method

Let A be an n-dimensional algebra over $\mathrm{GF}(q)$, and let $a_1, a_2, \ldots, a_n$ be a basis for A as a vector space over $\mathrm{GF}(q)$. For each pair of basis elements a_i, a_j we can express the product $a_i a_j$ as a linear combination

$$a_i a_j = \sum_{1 \le k \le n} \lambda_{ijk} a_k$$

for some scalars $\lambda_{ijk} \in \mathrm{GF}(q)$. These scalars are *structure constants* for the algebra A, and completely determine the product on A. So there are at most q^{n^3} n-dimensional algebras over $\mathrm{GF}(q)$. However if we pick a different vector space basis for A then we may get a different set of structure constants, so that different sets of structure constants can give the same algebra A.

We investigate how a change of basis affects the structure constants. Let A be an algebra of dimension n over $\mathrm{GF}(q)$, and let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be two bases for A as a vector space over $\mathrm{GF}(q)$. Let the sets of structure constants for these

two bases be $\{\lambda_{ijk} \mid 1 \leq i, j, k \leq n\}$ and $\{\mu_{ijk} \mid 1 \leq i, j, k \leq n\}$. We can express the elements of the second basis as linear combinations of elements of the first basis, and vice versa:

$$\begin{aligned} b_i &= \sum_{j=1}^{n} \alpha_{ji} a_j \ (1 \leq i \leq n), \\ a_j &= \sum_{k=1}^{n} \beta_{kj} b_k \ (1 \leq j \leq n), \end{aligned}$$

where $[\alpha_{ji}]$ and $\left[\beta_{kj}\right]$ are $n \times n$ matrices over $\mathrm{GF}(q)$ which are inverse to each other. So

$$\begin{aligned} b_i b_j &= \sum_{r,s=1}^{n} \alpha_{ri} \alpha_{sj} a_r a_s \\ &= \sum_{r,s,t=1}^{n} \alpha_{ri} \alpha_{sj} \lambda_{rst} a_t \\ &= \sum_{r,s,t,k=1}^{n} \alpha_{ri} \alpha_{sj} \lambda_{rst} \beta_{kt} b_k. \end{aligned}$$

It follows that

$$\mu_{ijk} = \sum_{r,s,t=1}^{n} \alpha_{ri} \alpha_{sj} \lambda_{rst} \beta_{kt}.$$

Each set of structure constants consists of n^3 elements of $\mathrm{GF}(q)$, and we can think of these sets of structure constants as elements in an n^3-dimensional vector space V over $\mathrm{GF}(q)$. The formula above defines an action of $\mathrm{GL}(n, q)$ on V, and if $g \in \mathrm{GL}(n, q)$ then g acts as a linear transformation $T_g : V \to V$. In addition, Higman's methods rely critically on the fact that the entries in the matrix for T_g are rational functions in the entries in the matrix for g.

The number of algebras of dimension n over $\mathrm{GF}(q)$ is the number of orbits in this action of $\mathrm{GL}(n, q)$ on V. This number is given by the Cauchy-Frobenius counting formula

$$\frac{1}{|\operatorname{GL}(n, q)|} \left(\sum_{g \in \operatorname{GL}(n,q)} |\operatorname{fix}(g)| \right),$$

where $\operatorname{fix}(g) = \{v \in V \mid vg = v\}$. Here, $\operatorname{fix}(g)$ is the eigenspace of T_g corresponding to eigenvalue 1.

The dimension of the eigenspace T_g corresponding to eigenvalue 1 depends only on the conjugacy class of g in $\mathrm{GL}(n, q)$, in other words on the rational canonical form of g. However Graham Higman [3] introduces of property of matrices called *type*, which he attributes to Green [1]. The type of a matrix is coarser than the rational canonical form, and records the degree of the irreducible polynomials occurring in the rational

canonical form, together with further information. The rational canonical form of a matrix A is the matrix

$$\begin{bmatrix} C(p_1(x)^{e_1}) & 0 & 0 & 0 \\ 0 & C(p_2(x)^{e_2}) & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & C(p_k(x)^{e_k}) \end{bmatrix},$$

with k blocks down the diagonal denoting the companion matrices of the primary invariant factors of A. These invariant factors are powers $p_i^{e_i}$ of monic irreducible polynomials $p_i(x)$. Let the distinct irreducible polynomials which occur in the rational canonical form of A be $q_1, q_2, \ldots, q_m$ (with $m \leq k$), and for $i = 1, 2, \ldots, m$ let S_i denote the multiset of exponents e such that q_i^e is an invariant factor of A. Then the type of A is the multiset of ordered pairs

$$\{(\deg q_1, S_1), (\deg q_2, S_2), \ldots, (\deg q_m, S_m)\}.$$

For example, if the primary invariant factors of A are $p_1(x)^2$, $p_1(x)^3$, $p_2(x)$, $p_2(x)$, $p_2(x)^4$ where $p_1(x)$ and $p_2(x)$ are distinct monic irreducible polynomials, then the type of A is

$$\{(\deg p_1, \{2, 3\}), (\deg p_2, \{1, 1, 4\})\}.$$

(Note that repeated entries in these multisets are significant.) If $A \in \mathrm{GL}(n, q)$ then the number of conjugacy classes in $\mathrm{GL}(n, q)$ with the same type as A is a polynomial in q. Green [1] proves that the size of the conjugacy class of A is also a polynomial in q, with the polynomial depending only on the type of A. A formula for this polynomial is given on page 181 of [4]. There are only finitely many possible types for $n \times n$ matrices. For example if $n = 3$ then the possible types are

$$\{(3, \{1\})\},\ \{(2, \{1\}), (1, \{1\})\},\ \{(1, \{1\}), (1, \{1\}), (1, \{1\})\},$$
$$\{(1, \{1\}), (1, \{1, 1\})\},\ \{(1, \{1\}), (1, \{2\})\},$$
$$\{(1, \{1, 1, 1\})\},\ \{(1, \{1, 2\})\},\ \{(1, \{3\})\}.$$

The rational canonical forms over a field F for each of these 8 types are

$$[C(a(x))],\ \begin{bmatrix} C(b(x)) & 0 \\ 0 & \lambda \end{bmatrix},\ \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{bmatrix},$$

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix},\ \begin{bmatrix} \lambda & 0 \\ 0 & C((x-\mu)^2) \end{bmatrix},$$

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix},\ \begin{bmatrix} \lambda & 0 \\ 0 & C((x-\lambda)^2) \end{bmatrix},\ \left[C((x-\lambda)^3)\right],$$

where a is an irreducible cubic over F, b is an irreducible quadratic, and where λ, μ, ν are distinct elements of F. As mentioned above, the number of conjugacy classes in $\mathrm{GL}(3,q)$ of each type is given by a polynomial in q. For example the number of conjugacy classes of type $\{(1,\{1\}),(1,\{1\}),(1,\{1\})\}$ is $(q-1)(q-2)(q-3)/6$. Note that this polynomial evaluates to 0 if $q \leq 3$. The dimension of $\mathrm{fix}(g)$ in its action on V is not altered if we extend the ground field to include the eigenvalues of g, and we can write down the Jordan canonical form of each of these 8 types in a suitable extension field. For example if g has type $\{(3,\{1\})\}$ then its Jordan canonical form is

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda^q & 0 \\ 0 & 0 & \lambda^{q^2} \end{bmatrix}$$

for some $\lambda \in \mathrm{GF}(q^3)$ satisfying $\lambda^{q^3-1} = 1$, $\lambda^{q-1} \neq 1$. Note that g is similar to a matrix of this form with $\lambda^{q^3-1} = 1$, $\lambda^{q-1} \neq 1$ if *and only if* g has type $\{(3,\{1\})\}$. Similarly if g has type $\{(2,\{1\}),(1,\{1\})\}$ then its Jordan canonical form is

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda^q & 0 \\ 0 & 0 & \mu \end{bmatrix}$$

with $\lambda^{q^2-1} = 1$, $\lambda^{q-1} \neq 1$, $\mu^{q-1} = 1$. Again, these equations and non-equations characterize elements of type $\{(2,\{1\}),(1,\{1\})\}$. To take one final example, g has type $\{(1,\{1\}),(1,\{1,1\})\}$ if and only if it has Jordan canonical form

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix}$$

with $\lambda^{q-1} = 1$, $\mu^{q-1} = 1$, $\lambda\mu^{-1} \neq 1$.

For each of these Jordan canonical forms we can pick a basis for V and compute the matrix for the action T_g of g on V. For example, suppose that g has type $\{(1,\{1\}),(1,\{1,1\})\}$ and Jordan canonical form

$$\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix}.$$

Then the matrix of T_g is a diagonal matrix with eigenvalues λ, $\lambda^2\mu^{-1}$, $\lambda^2\mu^{-1}$, μ, λ, λ, μ, λ, λ, μ, λ, λ, $\lambda^{-1}\mu^2$, μ, μ, $\lambda^{-1}\mu^2$, μ, μ, μ, λ, λ, $\lambda^{-1}\mu^2$, μ, μ, $\lambda^{-1}\mu^2$, μ, μ. So we see that the dimension of $\mathrm{fix}(g)$ is the number of 1's in this sequence of 27 eigenvalues. As another example, suppose that g has type $\{(1,\{2\})\}$ and Jordan canonical form

$$\begin{bmatrix} \lambda & \lambda \\ 0 & \lambda \end{bmatrix}.$$

(Following Higman we take our Jordan blocks to have a rather unusual form, with eigenvalues on the superdiagonal rather than 1's. This is possible since the eigenvalues are non-zero.) If we take a basis for V which expresses $\{\lambda_{ijk}\}$ as an 8-vector $(\lambda_{112}, \lambda_{111}, \lambda_{122}, \lambda_{121}, \lambda_{212}, \lambda_{211}, \lambda_{222}, \lambda_{221})$, then T_g has matrix

$$\lambda \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now we have a slight problem. The Jordan canonical form of this matrix depends on the characteristic of the underlying field and is

$$\lambda \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

in characteristic 2,

$$\lambda \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

in characteristic 3, and

$$\lambda \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

if the characteristic is greater than 3. To see this we proceed as follows. First, we think of the matrix for T_g as a matrix with entries in the rational function field $\mathbb{Q}(\lambda)$, and compute its Jordan canonical form. This is the same as the Jordan canonical form given above for characteristic greater than 3. We keep track of the transforming matrix, which has determinant $-108\lambda^8$. We also need to inspect the denominators of the coefficients in the transforming matrix, but in this case they are all 1. So this computation of the Jordan canonical form of T_g is valid in all characteristics greater than 3. To compute the Jordan canonical form of T_g over fields of characteristic 2 and 3 we think of the matrix for T_g as a matrix with entries in the rational function field $\mathrm{GF}(2)(\lambda)$ and in $\mathrm{GF}(3)(\lambda)$. In this way we see that $\mathrm{fix}(g)$ has dimension 0 if $\lambda \neq 1$. If $\lambda = 1$ then we see that $\mathrm{fix}(g)$ has dimension 4 in characteristic 2 and dimension 3 if the characteristic is greater than 2.

More generally, for any given n there are a finite number of types of $n \times n$ matrices. For each type t we can write down the Jordan canonical form of a matrix of type t, and we can write down a set of *monomial* equations and non-equations which the eigenvalues in the Jordan canonical form must satisfy. We can characterize matrices in $\mathrm{GL}(n,q)$ of type t as matrices with the given Jordan canonical form, where the eigenvalues satisfy the given set of equations and non-equations. As Green [1] proved, there is a polynomial in q, $f_t(q)$, depending on t, such that the conjugacy classes of all elements in $\mathrm{GL}(n,q)$ of type t have $f_t(q)$ elements. Take $g \in \mathrm{GL}(n,q)$ in Jordan canonical form, and pick a basis for V so that it represents a set $\{\lambda_{ijk}\}$ of structure constants as an n^3-vector with the entries indexed lexicographically on the triple $(i,j,-k)$. Then the matrix for T_g will be an upper triangular matrix with diagonal entries (eigenvalues) of the form $\lambda\mu\nu^{-1}$, where λ, μ, ν are eigenvalues of g. The Jordan canonical form of T_g may depend on the characteristic of $\mathrm{GF}(q)$. But there will be some prime p such that the Jordan canonical form is identical for all characteristics greater than p. For each of the eigenvalues $\lambda\mu\nu^{-1}$ of T_g we can compute the dimension of the corresponding eigenspace (though this dimension will depend on the characteristic of $\mathrm{GF}(q)$, as described above). Thus far the calculations only depend on the type of g and on the characteristic, and not on the particular choices of eigenvalues for g. For a given choices of eigenvalues of g, the dimension of $\mathrm{fix}(g)$ is obtained by adding up the dimensions of the eigenspaces of T_g corresponding to eigenvalues $\lambda\mu\nu^{-1}$ such that $\lambda\mu\nu^{-1} = 1$. So if the distinct eigenvalues of g are $\lambda_1, \lambda_2, \ldots, \lambda_m$, then the dimension of $\mathrm{fix}(g)$ is determined by which of the monomial equations $\lambda_i\lambda_j\lambda_k^{-1} = 1$ $(1 \leq i,j,k \leq m)$ are satisfied, and which are not satisfied. Let S be the set of all these equations $\lambda_i\lambda_j\lambda_k^{-1} = 1$. Then for each subset $T \subset S$ we can compute the dimension of $\mathrm{fix}(g)$ for those g of type t which satisfy the equations in T and fail to satisfy the equations in $S \backslash T$. Recall that there is also a set U of monomial equations and non-equations which the eigenvalues of g must satisfy to guarantee that g has type t. The number of choices of eigenvalues of g of type t which satisfy the equations in T and fail to satisfy the equations in $S \backslash T$ is given by the number of choices of $\lambda_1, \lambda_2, \ldots, \lambda_m$ which satisfy the equations and non-equations in $T \cup U$, and fail to satisfy the equations in $S \backslash T$. Higman [3] proves that the number of solutions in a finite field to a finite set of monomial equations and non-equations is PORC. A proof

of this can also be found in [8]. Furthermore the proof in [8] shows that the number of solutions can be expressed as a polynomial in q and in various gcds, like the formulae in Section 2. Each choice of eigenvalues for g of type t determines its conjugacy class, and the size of this conjugacy class is given by the polynomial $f_t(q)$ mentioned above. It follows that for any given dimension d, the number of $g \in \mathrm{GL}(n,q)$ of type t with $\mathrm{fix}(g)$ of dimension d is PORC. It follows that

$$\sum_{g \text{ has type } t} |\operatorname{fix}(g)|$$

is PORC. Since there are only finitely many possible types t for $n \times n$ matrices we see that

$$\sum_{g \in \mathrm{GL}(n,q)} |\operatorname{fix}(g)|$$

is PORC, and hence that $g(n,q)$ is PORC.

As we saw above, some of our calculations depend on the characteristic of $\mathrm{GF}(q)$. Specifically, for any given type t the computation of the Jordan canonical form for T_g when g has type t depends on the characteristic. We showed that there will be a prime p such that the Jordan canonical form is the same for all characteristics greater than p. Let P be the largest prime which arises in this way as we run through all possible types for $n \times n$ matrices. (For $n = 2$ and 3 we have $P = 3$ and for $n = 4$ we have $P = 5$.) Then we get a single PORC formula for $g(n,q)$ valid in all characteristics greater than P, and it can be expressed using gcds in a way similar to the formulae from Section 2. For each separate characteristic $p \leq P$ we similarly obtain a single PORC formula for $g(n,q)$.

4 Conclusions

As we have seen, for any given n there will be a number of different formulae for $g(n,q)$ depending on the characteristic of $\mathrm{GF}(q)$. Specifically, there will be a prime P and one formula for each characteristic $p \leq P$, and one further formula for all characteristics greater than P. The individual formulae will be polynomials in q and in a number of gcds of the kind we have seen in Section 2. There does not seem to be a canonical way of writing formulae involving gcds, since there are relations between them. For example

$$\gcd(q-1,4) + \gcd(q+1,4) + 2 = 4\gcd(q-1,2)$$

and

$$\gcd(q^3-1,7) + 2 = \gcd(q-1,7) + \gcd(q-2,7) + \gcd(q-4,7).$$

It is possible to show that it is sufficient to use gcds of the form $\gcd(q-a,p^k)$ for prime powers p^k and $0 < a < p^k$. Some of the formulae in Section 2 involve products of gcds, such as $\gcd(q-1,3)\gcd(q-1,4)$, which could of course be replaced by $\gcd(q-1,12)$. In fact it is not hard to see that products of gcds of this form can

always be replaced by linear combinations of gcds, though now we need to allow gcds of the form $\gcd(q-a,m)$ for arbitrary integers m, with $0<a<m$.

Perhaps more interesting is the asymptotic form of $g(n,q)$, and in fact we can say something about this. Clearly $\frac{q^{n^3}}{|\mathrm{GL}(n,q)|}$ is a lower bound for $g(n,q)$, but it is relatively easy to provide quite a reasonable upper bound. To see this consider the formula

$$\frac{1}{|\,\mathrm{GL}(n,q)|}\left(\sum_{g\in\mathrm{GL}(n,q)}|\,\mathrm{fix}(g)|\right)$$

for $g(n,q)$. If g is the identity element then $|\mathrm{fix}(g)|=q^{n^3}$. If $g\neq 1$ then the maximum dimension of $\mathrm{fix}(g)$ is n^3-3n^2+6n-4, which is attained when g is a diagonal matrix with all eigenvalues equal to 1, except for one eigenvalue equal to -1. It follows that

$$\frac{q^{n^3}}{|\,\mathrm{GL}(n,q)|}<g(n,q)<\frac{q^{n^3}}{|\,\mathrm{GL}(n,q)|}+q^{n^3-3n^2+6n-4}.$$

Perhaps a better way of expressing this, using the fact that $|\,\mathrm{GL}(n,q)|<q^{n^2}$, is

$$\frac{q^{n^3}}{|\,\mathrm{GL}(n,q)|}<g(n,q)<\frac{q^{n^3}+q^{n^3-2n^2+6n-4}}{|\,\mathrm{GL}(n,q)|}.$$

For $n=4$, for example this gives

$$\frac{q^{64}}{|\,\mathrm{GL}(4,q)|}<g(4,q)<\frac{q^{64}+q^{52}}{|\,\mathrm{GL}(4,q)|}.$$

We can push this argument further if we note that the conjugacy class of the diagonal matrix with $n-1$ eigenvalues 1 and one eigenvalue -1 has size $\frac{|\,\mathrm{GL}(n,q)|}{|\,\mathrm{GL}(n-1,q)|(q-1)}$, so that this conjugacy class only contributes $\frac{q^{n^3-3n^2+6n-4}}{|\,\mathrm{GL}(n-1,q)|(q-1)}$ to $g(n,q)$. The next largest possibility for the dimension of $\mathrm{fix}(g)$ is n^3-3n^2+5n-3 which arises when g is conjugate to a diagonal matrix with $n-1$ eigenvalues 1 and one eigenvalue which is different from 1 and -1. There are $q-3$ of these conjugacy classes, and they all have size $\frac{|\,\mathrm{GL}(n,q)|}{|\,\mathrm{GL}(n-1,q)|(q-1)}$. So these conjugacy classes contribute $\frac{q^{n^3-3n^2+5n-3}(q-3)}{|\,\mathrm{GL}(n-1,q)|(q-1)}$ to $g(n,q)$. We could push this argument even further, but the law of diminishing returns takes over pretty quickly.

References

[1] J.A. Green, *The characters of the finite general linear groups*, Trans. Amer. Math. Soc. **80** (1955), 402–447.

[2] G. Higman, *Enumerating p-groups. I: Inequalities*, Proc. London Math. Soc. **(3) 10** (1960), 24–30.

[3] G. Higman, *Enumerating p-groups. II: Problems whose solution is PORC*, Proc. London Math. Soc. **(3) 10** (1960), 566–582.

[4] I.G. Macdonald, *Symmetric functions and Hall polynomials*, The Clarendon Press, Oxford University Press, New York, 1979.

[5] M.F. Newman, E.A. O'Brien, and M.R. Vaughan-Lee, *Groups and nilpotent Lie rings whose order is the sixth power of a prime*, J. Algebra **278** (2004), 383–401.

[6] Holger P Petersson and Matthias Scherer, *The number of nonisomorphic two-dimensional algebras over a finite field*, Results Math. **45** (2004), 137–152.

[7] Michael Vaughan-Lee, *Graham Higman's PORC conjecture*, Jahres. Dtsch. Math.-Ver. **114** (2012), 89–106.

[8] Michael Vaughan-Lee, *On Graham Higman's famous PORC paper*, Internat. J. Group Theory **1** (2012), 65–79.