The exact number of conjugacy classes of the Sylow $p$-subgroups of $GL(n, q)$ modulo $(q - 1)^{13}$

A. Vera-López a,*, J.M. Arregi a, Leyre Ormaetxea a, F.J. Vera-López b

a Departamento de Matemáticas, Fac. de Ciencia y Tecnología, Universidad del País Vasco, Apartado 644, Bilbao, Spain

b Departamento de Matemática Aplicada, Escuela de Informática, Murcia, Spain

Received 14 June 2007; accepted 17 March 2008

Submitted by J.F. Queiro

Abstract

Let $G$ be a finite $p$-group of order $p^n$. A well known result of $P$. Hall determines the number of conjugacy classes of $G$, $r(G)$, modulo $(p^2 - 1)(p - 1)$. Namely, he proved the existence of a non-negative constant $k$ such that $r(G) = n(p^2 - 1) + p^k + k(p^2 - 1)(p - 1)$.

We denote by $\mathcal{G}_n$ the group of the upper unitriangular matrices over $F_q$, the finite field with $q = p^t$ elements. In [A. Vera-López, J.M. Arregi, F.J. Vera-López, On the number of conjugacy classes of the Sylow $p$-subgroups of $GL(n, q)$. Bull. Austral. Math. Soc. 53 (1996) 431–439] the number $r(\mathcal{G}_n)$ is given modulo $(q - 1)^5$.

In this paper, we introduce the concept of primitive canonical matrix. The knowledge of the number of primitive canonical matrices with connected graph of size less than or equal to $n$ should be sufficient to determine the number of all canonical matrices of size $n$. Moreover, we give explicitly the polynomial formulas $\mu_i = \mu_i(n)$, $i = 0, \ldots, 12$, depending only on $n$, and not on $q$, such that

$$r(\mathcal{G}_n) = \sum_i \mu_i(n)(q - 1)^i + k(n, q)(q - 1)^{13} \quad \forall n \in \mathbb{N}$$

© 2008 Elsevier Inc. All rights reserved.
We use the notion of canonical matrices given in [3,7]. These matrices are, in some sense, the simplest in each conjugacy class.

In the set \( J = \{(i,j)|1 \leq i < j \leq n\} \) of entries over the main diagonal, we define the order \( (i,j) < (k,l) \) if \( i > k \) or \( i = k \) and \( j < l \). Given a matrix \( A = (a_{ij}) \in G_n \), to each entry \( (i,j) \in J \) we associate a linear form \( L_{ij} = \sum_k (a_{ik}x_{kj} - a_{kj}x_{ik}) \). We define that an entry \( (i,j) \) is an inert point (resp. ramification point) if its linear form \( L_{ij} \) is independent (resp. dependent) of the preceding forms \( L_{kl} \), \( (k,l) \prec (i,j) \). We say that a matrix is canonical if it has the zero value at the inert points. It is proved that in each conjugacy class there exists a unique canonical matrix, so counting classes is counting canonical matrices.

**Definition.** Let \( A \in G_n \). We define the associated undirected graph of \( A \), denoted by \( \gamma = \gamma(A) \), the graph \( \gamma = (\nu, \delta) \), where
\[
\nu = \{i \in [1,n]|\exists a_{ij} \neq 0 \text{ or } \exists a_{ji} \neq 0\},
\]
\[
\delta = \{(i,j) \in J|a_{ij} \neq 0\}.
\]

We denote by \( \Gamma_n \) the set of undirected graphs associated to the canonical matrices of \( G_n \).

For each connected component \( \xi \) of \( \gamma \), with vertices \( k_1 < \cdots < k_r \), we define \( A_\xi = A_{k_1 \ldots k_r} \) to be the principal submatrix of \( A \) formed by \( k_1, \ldots, k_r \) rows and columns. For each linear form \( L_{ij} \) of \( A \) we define \( \text{var}(L_{ij}) = \text{var}_A(L_{ij}) \) to be the set of unknowns which appear in the form \( L_{ij} \) with non-zero coefficients. Dually, for each unknown \( x_{uv} \), we define \( \text{form}(x_{uv}) = \text{form}_A(x_{uv}) \) to be the set of linear forms \( L_{ij} \) in which \( x_{uv} \) appears with non-zero coefficient.

**Proposition 1.** Let \( A \in G_n \) and \( \xi = (\nu_1, \delta_1) \) a connected component of \( \gamma = \gamma(A) \). The following assertions hold:

1. Let \( i < j, i, j \in \nu_1 \). Then the indices of the unknowns of \( \text{var}(L_{ij}) \) are in \( \nu_1 \).
2. Let \( i < j, i, j \in \nu_1 \). Then the indices of the linear forms of \( \text{form}(x_{ij}) \) are in \( \nu_1 \).

**Proof**

1. We have \( L_{ij} = \sum_{k=i+1}^{j-1} (a_{ik}x_{kj} - a_{kj}x_{ik}) \). Suppose that \( a_{ik} \neq 0 \). Then the indices \( k, j \) of \( x_{kj} \) are in \( \nu_1 \). Similarly, suppose that \( -a_{kj} \neq 0 \). Then \( k \in \nu_1 \) because \( j \in \nu_1 \) and \( a_{kj} \neq 0 \), that is the indices of \( x_{ik} \) are in \( \nu_1 \).
2. The proof is like that of (1). \( \square \)

**Note.** In virtue of the previous proposition, for each \( i, j \in \nu_1 \), we can write
\[
L_{ij} = L_{ij}(A) = \sum_{s \in \nu_1} (a_{is}x_{sj} - a_{sj}x_{is}).
\]

**Theorem 2.** Let \( A \) be a matrix of \( G_n \). If the indices \( i, j, i < j \) belong to the same connected component of \( \gamma = \gamma(A) , \xi = (\nu_1, \delta_1) , i, j \in \nu_1 \), then the entry \( (i, j) \) is an inert (resp. ramification) point of \( A \), if and only if it is an inert (resp. ramification) point as entry of the submatrix \( A_\xi \).
Proof. Suppose that \((i, j)\) is a ramification point of \(A\). Then
\[
L_{ij} = \sum_{(\sigma, \tau) < (i, j)} \lambda_{\sigma \tau} L_{\sigma \tau}.
\] (2)

There is no loss of generality if we suppose that the \(\lambda_{\sigma \tau}\) coefficients in (2) correspond to the inert points \((\sigma, \tau)\) preceding \((i, j)\). If \(\var{L_{ij}} \cap \var{L_{\sigma \tau}} = \emptyset\), then equality (2) implies \(\lambda_{\sigma \tau} = 0\) and \(L_{ij} = 0\). In other case, if \(x_{fg} \in \var{L_{ij}} \cap \var{L_{\sigma \tau}}\), then, using the previous proposition we have \(f, g \in v_1\) and \(\sigma, \tau \in v_1\). Therefore, the linear forms that appear in the second member of (2) with non-zero coefficients have their indices in \(v_1\), so that \((i, j)\) is a ramification point as an entry of the submatrix \(A_{\xi}\).

Conversely, if \((i, j)\) is a ramification point as an entry of the submatrix \(A_{\xi}\), then \(L_{ij}\) depends on the preceding forms of \(A_{\xi}\) which also are preceding forms of \(A\) and consequently, \((i, j)\) is a ramification point as an entry of \(A\). □

We use this theorem to analyze the canonical matrices of \(G_n\), being known the canonical matrices of \(G_m\) for \(m < n\).

From now on, we express the graph \(\gamma = (\nu, \delta)\) as union of its disjoint components:
\[
\gamma = \xi_1 \cup \cdots \cup \xi_r, \quad \xi_k = (\nu_k, \delta_k), \quad v_k = |\nu_k|, \quad v_1 \leq \cdots \leq v_r.
\] (3)

Thus, we have the following result.

**Theorem 3.** Let \(A \in G_n\) and \(\gamma = \gamma(A)\) its graph. Then \(A\) is canonical in \(G_n\) if and only if for each connected component \(\xi_k = (\nu_k, \delta_k)\) of \(\gamma\), the matrix \(A_{\xi_k}\) is canonical in \(G_{v_k}\).

**Proof.** Let \(A\) be a canonical matrix and \(\xi_k\) a connected component of \(\gamma\). Suppose that \(i, j \in v_k\) and \(a_{ij} \neq 0\). Then, since \(A\) is canonical, \((i, j)\) is a ramification point of \(A\) and \(i, j \in v_k\). So, by Theorem 2, \((i, j)\) it is a ramification point of \(A_{\xi_k}\). Therefore, the non-zero entries of \(A_{\xi_k}\) lie at ramification points of \(A_{\xi_k}\) and, consequently, \(A_{\xi_k}\) is canonical.

Conversely, suppose that \(A_{\xi_k}\) are canonical and \(a_{ij} \neq 0\), \((i, j) \in \delta\). Then \((i, j) \in \delta_k\) for some \(k\) and, since \(A_{\xi_k}\) is canonical, the entry \((i, j)\) is a ramification point of \(A_{\xi_k}\) and, by Theorem 2, it a ramification point of \(A\). Therefore, the non-zero entries of \(A\) lie at ramification points of \(A\) and, consequently, \(A\) is canonical. □

**Definition.** A canonical matrix \(A\) (resp. a graph \(\gamma \in \Gamma\)) is said to be **primitive** if for each index \(j = 1, \ldots, n\) there exists some \(a_{ij} \neq 0, i < j\), or \(a_{jk} \neq 0, j < k\) (resp. edge \((i, j)\) or \((j, k)\) \(\in \delta\)).

Given a canonical matrix \(A\) we obtain a primitive canonical matrix by suppressing that indices whose row and column of \(A_0 = A - I\) are zero.

**Definition.** For each graph \(\gamma\) we denote \(\mathcal{A}_\gamma\) the set of primitive canonical matrices of \(G_n\) with graph \(\gamma\) and \(r_\gamma(G_n) = |\mathcal{A}_\gamma|\).

As immediate consequence of the preceding theorem we have the next result.

**Corollary 4.** Let be a graph union of disjoint components: \(\gamma = \xi_1 \cup \cdots \cup \xi_r, \xi_k = (\nu_k, \delta_k)\). Then \(r_\gamma(G_n) = r_{\xi_1}(G_{v_1}) \cdots r_{\xi_r}(G_{v_r})\).
Corollary 5. Let \( A \in \mathcal{G}_n \) be a canonical matrix and \( B \in \mathcal{G}_t \) its primitive matrix. Then \( A \) is canonical in \( \mathcal{G}_n \) if and only if \( B \) is canonical in \( \mathcal{G}_t \).

Proof. Let \( B = A_{s_1, \ldots, s_t} \). Then it is obvious that \( A \) and \( B \) have the same connected components \( \xi \). Moreover, \( A_\xi = B_\xi \) for every connected component \( \xi \). \( \Box \)

Example. Consider the following canonical matrix:

\[
A = \begin{pmatrix}
* & \cdot & \cdot & \cdot & \cdot & \cdot & \\
* & \cdot & \cdot & \cdot & \cdot & \cdot & \\
* & \cdot & \cdot & \cdot & \cdot & \cdot & \\
* & \cdot & \cdot & \cdot & \cdot & \cdot & \\
* & \cdot & \cdot & \cdot & \cdot & \cdot & \\
* & \cdot & \cdot & \cdot & \cdot & \cdot & \\
\end{pmatrix}
\]

The graph of this matrix is:

\begin{align*}
1 & \rightarrow 3 \\
& \downarrow \\
4 & \rightarrow 7 \\
\end{align*}

\begin{align*}
2 & \rightarrow 6 \rightarrow 8 \\
\end{align*}

The index 5 does not appear in these components, so the primitive matrix of \( A \) is of order 7 and its graph has the next two connected components:

\begin{align*}
1 & \rightarrow 3 \\
& \downarrow \\
4 & \rightarrow 6 \\
\end{align*}

\begin{align*}
2 & \rightarrow 5 \rightarrow 7 \\
\end{align*}

The two matrices corresponding to these components are, respectively,

\[
A_1 = \begin{pmatrix}
* & \cdot & \cdot & \\
* & \cdot & \cdot & \\
* & \cdot & \cdot & \\
\end{pmatrix}, \quad A_2 = \begin{pmatrix}
* & \cdot & \cdot \\
* & \cdot & \\
* & \cdot & \\
\end{pmatrix}
\]

Bearing in mind the two last results, from now on, we will study the canonical primitive matrices with connected graphs.

Let \( \mathcal{D}_n \) be the set of regular diagonal matrices of size \( n \) over \( \mathbb{F}_q \). For each \( D \in \mathcal{D}_n \), the map \( A \mapsto A^D \) is an automorphism of \( \mathcal{G}_n \) which permutes the conjugacy classes of the same size, moreover, it transforms canonical matrices into canonical matrices having the same non-zero entries, inert and ramification points. We define the \( \mathcal{D}_n \)-class of \( A \) to be the set

\[
A^{\mathcal{D}_n} = \{A^D \mid D \in \mathcal{D}_n\}.
\]

For each canonical matrix \( A \) we have

\[
|A^{\mathcal{D}_n}| = |\mathcal{D}_n; C_{\mathcal{D}_n}(A)|.
\]

We can obtain this cardinality in the following way.

Definition. An edge-subset \( \delta' \subset \delta \) is said to be admissible (for \( A \)) if the corresponding graph is cycle-free. In this case the subgraph \( \xi' = (v, \delta') \) is also said to be admissible.
Theorem 6. Let $A \in \mathcal{G}_n$. Then
\[ |A^{D_n}| = (q - 1)^{|\delta M|}, \]
where $\delta_M$ is a maximal admissible edge-subset of $A$.

Proof. See [8, Theorem 4]. □

Note. In each $D_n$-class of canonical matrices, we choose as representative the one with the value 1 at the entries $(i, j)$ of a maximal admissible edge-subset.

Corollary 7. Let $A \in \mathcal{G}_n$ be a primitive canonical matrix with connected graph. Then
\[ |A^{D_n}| = (q - 1)^{n - 1}. \]

Proof. Let $\xi = (\nu, \delta)$ be the connected graph of $A$. Since $A$ is primitive, $v = n$. By dropping the edges which close cycles, we obtain a maximal admissible subgraph $\xi_M = (\nu, \delta_M)$ which is a tree joining the $n$ indices with $|\delta_M| = n - 1$ edges. □

Remark. In general, let $\gamma = \bigcup_{k=1}^{r} \xi_k$, be the graph of a canonical matrix $A$ expressed as disjoint union of connected components. Then a maximal admissible graph is obtained by taking maximal admissible graphs on these components. These new components are trees with $|\delta_M| = v_k - 1$ edges. So, for any canonical matrix $A$ we have
\[ |A^{D_n}| = (q - 1)^{v_1 + \cdots + v_r - r}. \]

Definition. For each canonical matrix $A \in \mathcal{G}_n$, let $\gamma = \gamma(A)$, as in (3), be its graph and $v_0$ the number of indices not appearing in $\gamma$. Note that $\sum_{k=0}^{r} v_k = n$. We denote by $\mathcal{A}_{v_0; v_1, \ldots, v_r}$ the family of canonical matrices with parameters $v_0; v_1, \ldots, v_r$. To determine totally the graph of $A$ it is necessary and sufficient to fix a partition of the set $\{1, \ldots, n\}$ into $r + 1$ subsets with cardinalities $v_0, v_1, \ldots, v_r$, satisfying $v_1 \leq \cdots \leq v_r$, and then fix the corresponding connected components of the graph, $\xi_1, \ldots, \xi_r$. So, we conclude the proof of the following result.

Proposition 8. The number of canonical matrices with parameters $v_0; v_1, \ldots, v_r, v_0 + v_1 + \cdots + v_r = n$, is given by
\[ |\mathcal{A}_{v_0; v_1, \ldots, v_r}| = \binom{n - v_0}{v_0} P_{v_1, \ldots, v_r}^{n - v_0} |\mathcal{A}_{v_1}| \cdots |\mathcal{A}_{v_r}|, \]
(4)

where $\mathcal{A}_v = \mathcal{A}_{0, v}$ is the set of canonical matrices with connected primitive graph, $\xi_k = (v_k, \delta_k)$, and $P_{v_1, \ldots, v_r}$ is the number of partitions of the set $\{1, \ldots, n - v_0\}$ into $r$ subsets with $v_1, \ldots, v_r$ elements, respectively.

Remark. The combinatorial expression for cited number of partitions is
\[ P_{v_1, \ldots, v_r}^{n - v_0} = \binom{n - v_0}{v_1, \ldots, v_r} \frac{1}{\rho(1)! \cdots \rho(n - v_0)!}, \quad \rho(i) = |\{k|v_k = i\}|. \]

Definition. We define $r_{pc}(\mathcal{G}_n)$ as the number of conjugacy classes of $\mathcal{G}_n$ whose canonical matrices are primitive and have connected graph.
To compute \( r(\mathcal{G}_n) \), we add the cardinalities (4) corresponding to the different sets of parameters and we obtain the following proposition.

**Theorem 9.** Suppose that, for each \( t \leq n \), there exists a polynomial \( r_{pc,t}(x) \) whose coefficients are independent of \( q \), such that \( r_{pc}(\mathcal{G}_t) = r_{pc,t}(q - 1) \). Then the number of conjugacy classes of \( \mathcal{G}_n \) may be computed by the following equation:

\[
 r(\mathcal{G}_n) = \sum_{v_0, v_1, \ldots, v_r \geq 0} \binom{n}{v_0} p_{v_1, \ldots, v_r} \prod_{i=0}^{r} r_{pc,i}(q - 1). 
\]

Comparing coefficients in this equality, we have the explicit expressions for the coefficients \( \mu_i(n) \):

**Corollary 10.** Suppose that, for \( t \leq n \), \( r_{pc}(\mathcal{G}_t) \) is a polynomial in \((q - 1)\) whose coefficients depend on \( n \) but not on \( q \). Then the number of conjugacy classes of \( \mathcal{G}_n \) may be computed by the following equation:

\[
 r(\mathcal{G}_n) = \sum_{i=0}^{\infty} \mu_i(n)(q - 1)^i, \quad (5)
\]

where

\[
 \mu_i(n) = \sum_{v_0, v_1, \ldots, v_r \geq 0} \binom{n}{v_0} p_{v_1, \ldots, v_r} \text{coeff}(r_{pc,t_1}(x) \cdots r_{pc,t_r}(x), x^i). \quad (6)
\]

**Remark.** Higman has conjectured that, for each \( n \), the number of conjugacy classes of elements of \( \mathcal{G}_n \) is a polynomial expression in \( q \) with coefficients independent of \( q \). In the hypothesis of Higman’s conjecture, the polynomial \( r(\mathcal{G}_n) \) is given by means of expressions (5) and (6). These expressions show that the computation of \( r(\mathcal{G}_n) \) depends only on the concepts of canonical primitive matrices with connected graph.

We list the first thirteen \( r_{pc,k}(x) \) polynomials:

\[
\begin{align*}
 r_{pc,1}(x) &= 1, \\
r_{pc,2}(x) &= x, \\
r_{pc,3}(x) &= x^2, \\
r_{pc,4}(x) &= 2x^3, \\
r_{pc,5}(x) &= 5x^4, \\
r_{pc,6}(x) &= 18x^5 + x^6, \\
r_{pc,7}(x) &= 77x^6 + 8x^7, \\
r_{pc,8}(x) &= 404x^7 + 74x^8 + 4x^9, \\
r_{pc,9}(x) &= 2451x^8 + 665x^9 + 72x^{10} + 3x^{11}, \\
r_{pc,10}(x) &= 17100x^9 + 6462x^{10} + 1140x^{11} + 110x^{12} + 5x^{13}, \\
r_{pc,11}(x) &= 134145x^{10} + 66584x^{11} + 16632x^{12} + 2563x^{13} + 242x^{14} + 11x^{15},
\end{align*}
\]
Corollary 11. The first 12 coefficients of the development of \( r(\mathcal{D}_n) \) into powers of \( q - 1 \) are polynomial functions only depending on \( n \) and not on \( q \), because of the following equalities:

\[
\begin{align*}
\mu_0(n) &= 1, \\
\mu_1(n) &= n(n - 1)/2, \\
\mu_2(n) &= n(n - 1)(n - 2)(3n - 5)/24, \\
\mu_3(n) &= n(n - 1)(n - 2)(n - 3)(n^2 - 5n + 8)/48, \\
\mu_4(n) &= n(n - 1)(n - 2)(n - 3)(n - 4)(3n^3 - 30n^2 + 121n - 182)/1152, \\
\mu_5(n) &= n(n - 1)(n - 2)(n - 3)(n - 4)(n - 5) \\
&\quad \times (3n^4 - 50n^3 + 365n^2 - 1310n + 1920)/11520, \\
\mu_6(n) &= n(n - 1)(n - 2)(n - 3)(n - 4)(n - 5)(9n^6 - 279n^5 \\
&\quad + 3915n^4 - 31405n^3 + 150060n^2 - 401372n + 465888)/414720, \\
\mu_7(n) &= n(n - 1)(n - 2)(n - 3)(n - 4)(n - 5)(n - 6) \\
&\quad \times (9n^7 - 378n^6 + 7350n^5 - 84700n^4 + 618625n^3 \\
&\quad - 2842154n^2 + 7556672n - 8917632)/580680, \\
\mu_8(n) &= \frac{1}{278691840} n(n - 1)(n - 2)(n - 3)(n - 4)(n - 5)(n - 6)(n - 7) \\
&\quad \times (27n^8 - 1476n^7 + 37926n^6 - 591528n^5 + 6074075n^4 \\
&\quad - 41775748n^3 + 186904996n^2 - 494895824n + 591057792), \\
\mu_9(n) &= \frac{1}{557383680} n(n - 1)(n - 2)(n - 3)(n - 4)(n - 5)(n - 6)(n - 7) \\
&\quad \times (3n^{10} - 231n^9 + 8442n^8 - 191614n^7 + 2975371n^6 \\
&\quad - 32874695n^5 + 260680104n^4 - 1459329876n^3 \\
&\quad + 5500310480n^2 - 12561067392n + 13155760128), \\
\mu_{10}(n) &= \frac{1}{33443020800} n(n - 1)(n - 2)(n - 3)(n - 4)(n - 5)(n - 6)(n - 7)(n - 8) \\
&\quad \times (9n^{11} - 846n^{10} + 38025n^9 - 1072470n^8 + 20988735n^7
\end{align*}
\]
\[ \mu_{11}(n) = \frac{1}{735746457600} n(n-1)(n-2)(n-3)(n-4)(n-5)(n-6)(n-7)(n-8) \times (9n^{13} - 1095n^{12} + 64071n^{11} - 2377815n^{10} + 62230685n^9 - 1209104369n^8 + 17892479077n^7 - 20365182242n^6 + 1778477794390n^5 - 11741788530164n^4 + 56836649446232n^3 - 190597857038880n^2 + 396040317099264n - 384327254476800), \]

\[ \mu_{12}(n) = \frac{1}{52973744947200} n(n-1)(n-2)(n-3)(n-4)(n-5)(n-6)(n-7) \times (n-8)(n-9)(27n^{14} - 3861n^{13} + 266679n^{12} - 11748825n^{11} + 367673625n^{10} - 8621844759n^9 + 155837722381n^8 - 2200418288739n^7 + 24336685840632n^6 - 209372512972392n^5 + 1376482974616848n^4 - 6692871949029936n^3 + 22709660250220544n^2 - 48046159820733696n + 47735892549734400). \]

**Remark.** For \( n \leq 13 \) and each \( q \), \( r(\mathcal{G}_n) \) is totally determined in [8]. In particular, the coefficients \( \mu_i(n) \) of \((q-1)^i\) for \( i \leq 12 \) agree with those obtained by Corollary 11.

Consequently, the number \( r(\mathcal{G}_n) \) is determined modulo \((q-1)^{13}\) for every \( n \in \mathbb{N} \).

**References**