MATHEMATICS OF COMPUTATION
VOLUME 41, NUMBER 163
JULY 1983, PAGES 229-234

An Application of Matrices Over Finite Fields to Algebraic Number Theory

By Frank Gerth III

Abstract. This paper ultilizes properties of matrices over finite fields to obtain information about the rank of the p-class group of certain algebraic number fields.

1. Introduction. Let K be a Galois extension of the field of rational numbers $\mathbf{Q}$ of degree p, where p is a prime number. Let A denote the p-class group of K, i.e., the Sylow p-subgroup of the ideal class group of K. (For $p = 2$, we shall be using the Sylow 2-subgroup of the strict (or narrow) ideal class group of K.) Let v denote the rank of A; i.e., $v = \dim_{\mathbf{F}_p}(A/pA)$, where $\mathbf{F}_p$ is the finite field with p elements. Let t denote the number of primes that ramify in $K/\mathbf{Q}$. It is a classical result that $v = t - 1$ if $p = 2$ (see [1, p. 247]), and in general $t - 1 \leqslant v \leqslant (p - 1)(t - 1)$. (See [5, Satz 30].) When $t = 1$, we get $v = 0$ for all p. For fixed $p \geqslant 3$ and $t \geqslant 2$, we shall show that v is usually equal to $t - 1$ and that in a probabilistic sense the expected value of v, denoted $E(v)$, satisfies $t - 1 < E(v) < t$. The techniques we use are similar to some of the techniques used by Rédei in [6] to specify the 4-rank of A in the quadratic case. In Section 2 we shall develop some results we need about matrices over finite fields, and in Section 3 we shall apply the results in Section 2 to obtain information about v.

2. Ranks of Matrices Over Finite Fields. Let M be an $m \times n$ matrix with entries in the finite field $\mathbf{F}_p$, where $m \leqslant n$ and p is a prime number. Next let N_r be the number of these $m \times n$ matrices M over $\mathbf{F}_p$ with rank $M = r$, where $0 \leqslant r \leqslant m$.

PROPOSITION 2.1.

$$N_r = \left[\prod_{j=1}^{r} (p^n - p^{j-1})\right] \sum_{\substack{i_1 + \cdots + i_r \leqslant m-r \\ \text{each } i_s \geqslant 0}} \left(\prod_{s=1}^{r} p^{s i_s}\right).$$

(For $r = 0$, we interpret this as $N_0 = 1$.)

Proof. Let M be an $m \times n$ matrix over $\mathbf{F}_p$ with rank $M = r$. Let r_i be the rank of the first i rows of M, $1 \leqslant i \leqslant m$. Then $r_1 \leqslant r_2 \leqslant \cdots \leqslant r_m = r$. Thus to each M with rank $M = r$, we have associated an ordered m-tuple $(r_1, r_2, \ldots, r_m)$. To determine N_r, it suffices to determine all possible m-tuples $(r_1, \ldots, r_m)$ and the number of M associated with each m-tuple. Let r_{k_s} be the first term in $(r_1, \ldots, r_m)$ with $r_{k_s} = s$ for

Received October 21, 1981.
1980 *Mathematics Subject Classification.* Primary 12A35, 12A50; Secondary 15A33.

$1 \leqslant s \leqslant r$. Let $i_0 = k_1 - 1$, $i_s = k_{s+1} - k_s - 1$ for $1 \leqslant s \leqslant r-1$, and $i_r = m - k_r$. Then $r_1 = \cdots = r_{i_0} = 0$ (if $k_1 > 1$), and for $s = 1,\ldots,r$, we have $r_{k_s} = r_{k_s+1} = \cdots = r_{k_s+i_s} = s$. We note that each $i_s \geqslant 0$, and

$$i_0 + i_1 + \cdots + i_r = m - r. \tag{2.1}$$

Also the $(r+1)$-tuple $(i_0, i_1,\ldots,i_r)$ determines the r-tuple $(k_1,\ldots,k_r)$, which determines the m-tuple $(r_1,\ldots,r_m)$. Now how many matrices M are associated with a given $(r_1,\ldots,r_m)$, or equivalently, with a given $(i_0, i_1,\ldots,i_r)$? For rows $1,\ldots,i_0$, there is only one possibility, namely rows with all entries equal to 0. For row k_1 there are $p^n - 1$ possibilities (only the row with all entries equal to 0 is excluded). Each of rows $k_1+1,\ldots,k_1+i_1$ must be contained in the space spanned by row k_1, and hence there are p possibilities for each such row. In general there are $p^n - p^{s-1}$ possibilities for row k_s (i.e., any row vector not in the $(s-1)$-dimensional space spanned by rows $1,\ldots,k_s - 1$) and p^s possibilities for each of rows $k_s+1,\ldots,k_s+i_s$ (i.e., any row vector contained in the space spanned by rows $1,\ldots,k_s$). Thus, for a given $(r+1)$-tuple $(i_0, i_1,\ldots,i_r)$, the number of possible matrices M is

$$1^{i_0}(p^n-1)p^{i_1}\cdots(p^n-p^{s-1})(p^s)^{i_s}\cdots(p^n-p^{r-1})(p^r)^{i_r}.$$

Now allowing for all $(i_0, i_1,\ldots,i_r)$ satisfying Eq. (2.1) with each $i_s \geqslant 0$, we get

$$\begin{aligned} N_r &= \sum_{\substack{i_0+i_1+\cdots+i_r=m-r\\ \text{each } i_s\geqslant 0}} \left[\prod_{s=1}^{r}(p^n-p^{s-1})p^{si_s}\right] \\ &= \left[\prod_{j=1}^{r}(p^n-p^{j-1})\right] \sum_{\substack{i_1+\cdots+i_r\leqslant m-r\\ \text{each } i_s\geqslant 0}} \left(\prod_{s=1}^{r}p^{si_s}\right). \end{aligned}$$

Remark. Proposition 2.1 can be generalized to an arbitrary finite field with p^k elements by replacing p by p^k.

We shall now restrict our attention to $(t-1)\times t$ matrices M over $\mathbf{F}_p$, where $p \geqslant 3$ and $t \geqslant 2$ are fixed. Since there are $p^{t(t-1)}$ such matrices, the probability (which we denote by $R_{t,r}$) that a randomly chosen $(t-1)\times t$ matrix M over $\mathbf{F}_p$ has rank $M = r$ is given by

$$R_{t,r} = \frac{N_r}{p^{t(t-1)}} = \left[\prod_{j=1}^{r}\left(1-\frac{1}{p^{t+1-j}}\right)\right]\cdot\frac{1}{p^{t(t-1-r)}}\cdot\sum_{\substack{i_1+\cdots+i_r\leqslant t-1-r\\ \text{each } i_s\geqslant 0}}\left(\prod_{s=1}^{r}p^{si_s}\right).$$

In subsequent calculations it will be convenient to let $e = t-1-r$ and $B_{t,e} = R_{t,r}$ (thus for example, $e = 0$ when $r = t-1$, and $B_{t,0} = R_{t,t-1}$). Then

$$B_{t,e} = \left[\prod_{j=1}^{t-1-e}\left(1-\frac{1}{p^{t+1-j}}\right)\right]\cdot\frac{1}{p^{te}}\cdot\sum_{\substack{i_1+\cdots+i_{t-1-e}\leqslant e\\ \text{each } i_s\geqslant 0}}\left(\prod_{s=1}^{t-1-e}p^{si_s}\right). \tag{2.2}$$

For $t \geqslant 2$ and $0 \leqslant e \leqslant t-1$, we note that the probability $B_{t,e}$ satisfies $0 < B_{t,e} < 1$.

LEMMA 2.2.

$$\sum_{e=0}^{t-1} eB_{t,e} \leqslant \sum_{e=0}^{t-1} e(p-2)B_{t,e} < 1.$$

Proof. The first inequality is clear since we are assuming $p \geqslant 3$. We now let

$$W_{t,e} = \sum_{\substack{\iota_1 + \cdots + \iota_{t-1-e} \leqslant e \\ \text{each } \iota_s \geqslant 0}} \left(\prod_{s=1}^{t-1-e} p^{s\iota_s} \right).$$

(For $e = t - 1$, we interpret this as $W_{t,t-1} = 1$.) For $t > e \geqslant 1$, we have

$$\begin{aligned} W_{t,e} &\leqslant \left(1 + p + \cdots + p^{t-1-e}\right) W_{t-1,e-1} \\ &\leqslant \left(1 + p + \cdots + p^{t-1-e}\right) W_{t,e-1} = \frac{p^{t-e} - 1}{p - 1} \cdot W_{t,e-1}. \end{aligned}$$

Using Eq. (2.2), we then get

$$\begin{aligned} B_{t,e} &\leqslant B_{t,e-1}\left(1 - \frac{1}{p^{e+1}}\right)^{-1} \cdot \frac{1}{p^t} \cdot \frac{p^{t-e} - 1}{p - 1} = B_{t,e-1} \cdot \frac{1}{p - 1} \cdot \frac{p^{t-e} - 1}{p^t - p^{t-e-1}} \\ &= B_{t,e-1} \cdot \frac{1}{p - 1} \cdot \frac{1 - \left(p^{t-e}\right)^{-1}}{p^e - p^{-1}} < \left(\frac{1}{p - 1}\right)^2 B_{t,e-1}. \end{aligned}$$

Then by induction we get

$$B_{t,e} < \left(\frac{1}{p - 1}\right)^{2e} B_{t,0} < \left(\frac{1}{p - 1}\right)^{2e} \quad \text{for } t > e \geqslant 1.$$

Finally

$$\begin{aligned} \sum_{e=0}^{t-1} e(p - 2)B_{t,e} &< \sum_{e=0}^{t-1} e(p - 2)\left(\frac{1}{p - 1}\right)^{2e} < \sum_{e=1}^{t-1} \frac{e}{(p - 1)^e} \cdot \frac{1}{(p - 1)^{e-1}} \\ &\leqslant \sum_{e=1}^{t-1} \frac{1}{2} \cdot \frac{1}{2^{e-1}} < \sum_{e=1}^{\infty} \frac{1}{2^e} = 1. \end{aligned}$$

Remark. If X is a random variable which assumes the value e $(0 \leqslant e \leqslant t - 1)$ with $\operatorname{Prob}(X = e) = B_{t,e}$, then the expected value $E(X) = \sum_{e=0}^{t-1} e B_{t,e} < 1$ according to Lemma 2.2. It then follows that for an arbitrarily chosen $(t - 1) \times t$ matrix M over $\mathbf{F}_p$, the expected rank is greater than $t - 2$.

LEMMA 2.3. *Let $t \geqslant 2$ be arbitrary. For $p = 3$, $B_{t,0} > .840$; for $p = 5$, $B_{t,0} > .950$; for $p = 7$, $B_{t,0} > .976$; and for $p \geqslant 11$, $B_{t,0} > .99$.*

Proof. For all $p \geqslant 3$, $B_{t,0} = \prod_{j=1}^{t-1}(1 - 1/p^{t+1-j})$ from Eq. (2.2). By letting $k = t + 1 - j$, we get $B_{t,0} = \prod_{k=2}^{t}(1 - 1/p^k)$. Now for all $t \geqslant 2$,

$$B_{t,0} > \prod_{k=2}^{\infty}\left(1 - \frac{1}{p^k}\right) > 1 - \sum_{k=2}^{\infty} \frac{1}{p^k} = 1 - \left(\frac{1}{p^2}\right)\left(\frac{1}{1 - p^{-1}}\right) = 1 - \frac{1}{p^2 - p}.$$

When $p \geqslant 11$, it is clear that $B_{t,0} > .99$. For the cases $p = 3, 5, 7$, the product $\prod_{k=2}^{\infty}(1 - 1/p^k)$ was evaluated numerically to three decimal places to give the above results.

Table 2.1 gives values for $B_{t,e}$ when $t = 2, 3, 4$ and $p = 3, 5, 7, 11$.

TABLE 2.1. *Values of* $B_{t,e}$

	$t \backslash e$	0	1	2	3
$p = 3$	2	.8889	.1111		
	3	.8560	.1427	.0014	
	4	.8454	.1526	.0020	2×10^{-6}
$p = 5$	2	.9600	.0400		
	3	.9523	.0476	.0001	
	4	.9508	.0491	.0001	4×10^{-9}
$p = 7$	2	.9796	.0204		
	3	.9767	.0233	8×10^{-6}	
	4	.9763	.0237	1×10^{-5}	7×10^{-11}
$p = 11$	2	.9917	.0083		
	3	.9910	.0090	6×10^{-7}	
	4	.9909	.0091	6×10^{-7}	3×10^{-13}

LEMMA 2.4. *For all* $t \geqslant 2$ *and* $p \geqslant 3$, $B_{t,0} + B_{t,1} > .99$.

Proof. Since $B_{t,0} > .99$ if $p \geqslant 11$, it suffices to consider $p = 3, 5, 7$. We claim that $B_{t+1,1} > B_{t,1}$ for all $p \geqslant 3$ and $t \geqslant 2$. To show this, we use Eq. (2.2) to get

$$\begin{aligned} B_{t+1,1} &= B_{t,1}\left(1 - \frac{1}{p^{t+1}}\right) \cdot \frac{1}{p} \cdot \frac{p^{t-1} + \cdots + p + 1}{p^{t-2} + \cdots + p + 1} \\ &= B_{t,1} \frac{(p^{t+1} - 1)(p^{t-1} + \cdots + p + 1)}{p^{t+2}(p^{t-2} + \cdots + p + 1)} \\ &= B_{t,1} \frac{p^{2t} + p^{2t-1} + \cdots + p^{t+1} - p^{t-1} - p^{t-2} - \cdots - 1}{p^{2t} + p^{2t-1} + \cdots + p^{t+2}} \\ &> B_{t,1} \end{aligned}$$

since $p^{t+1} - p^{t-1} - p^{t-2} - \cdots - 1 > 0$. We now apply Lemma 2.3 and the results from Table 2.1. If $p = 7$, then for $t \geqslant 2$, $B_{t,0} + B_{t,1} > .976 + B_{2,1} > .99$. If $p = 5$, then for $t \geqslant 2$, $B_{t,0} + B_{t,1} > .950 + B_{2,1} = .99$. If $p = 3$, then for $t \geqslant 4$, $B_{t,0} + B_{t,1} > .840 + B_{4,1} > .99$. Also from Table 2.1 we see that $B_{2,0} + B_{2,1} > .99$ and $B_{3,0} + B_{3,1} > .99$. Hence the proof of Lemma 2.4 is complete.

3. Ranks of p-Class Groups. We first let K be a Galois extension of $\mathbf{Q}$ of degree 3, and we let A be the 3-class group of K. We assume that exactly t primes ramify in $K/\mathbf{Q}$, where $t \geqslant 2$, and we let f_K denote the conductor of K. (*Remark*: The prime divisors of the conductor are the ramified primes.) Employing the techniques described in Chapters IV and VI of [4], we see that $v = \operatorname{rank} A = 2(t - 1) - r$, where r is the rank of a certain $t \times t$ matrix of Hilbert symbols, and we may think of this matrix as a $t \times t$ matrix over $\mathbf{F}_3$. Because of the product formula for Hilbert symbols, the last row of the matrix is completely determined by the preceding $(t - 1)$ rows; hence we are considering a certain $(t - 1) \times t$ matrix M over $\mathbf{F}_3$ associated with K. From [2] and [3], we see that M is equally likely to be any $(t - 1) \times t$ matrix over $\mathbf{F}_3$ in the following sense. Let x be a large positive real number, and let $S_x = \{K \mid \text{exactly } t \text{ primes ramify in } K/\mathbf{Q} \text{ and the conductor } f_K \leqslant x\}$. Assume S_x has the counting measure, and let W_x be the function which assigns to each $K \in S_x$ the associated matrix M. If H is an arbitrary $(t-1) \times t$ matrix over $\mathbf{F}_3$, let $V_x(H)$ be the probability that W_x takes the value H. Then $V_x(H) \to 1/3^{t(t-1)}$ as $x \to \infty$. The fact that this limit probability is the same for all H is the reason we say that each possible choice for M is equally likely.

Now let N_r be the number of $(t-1) \times t$ matrices over $\mathbf{F}_3$ that have rank $= r$, where $0 \leqslant r \leqslant t-1$. Let Y_x be the random variable which assigns to each $K \in S_x$ the rank of the matrix M associated with K. Then $\operatorname{Prob}(Y_x = r) \to N_r/3^{t(t-1)}$ as $x \to \infty$. Now recall that the 3-class group A of K has rank satisfying

$$v = \operatorname{rank} A = 2(t-1) - r = t - 1 + (t - 1 - r) = t - 1 + e,$$

where we have set $e = t - 1 - r$. Then the following proposition is a consequence of our results from Section 2.

PROPOSITION 3.1. *Let an integer $t \geqslant 2$ be fixed, and let x be a positive real number. Let S_x be the set of all cubic Galois extensions K of $\mathbf{Q}$ with exactly t ramified primes over $\mathbf{Q}$ and conductor $f_K \leqslant x$. Assume S_x has counting measure. If Z_x is the random variable which assigns to each $K \in S_x$ the rank of the 3-class group of K, then* $\operatorname{Prob}(Z_x = t - 1 + e) \to B_{t,e}$ *as $x \to \infty$, where $B_{t,e}$ is given by Eq.* (2.2) *with $p = 3$, and $0 \leqslant e \leqslant t-1$. In particular*

$$\operatorname{Prob}(Z_x = t-1) > .840 \quad \textit{and} \quad \operatorname{Prob}(Z_x = t - 1 \textit{ or } t) > .99$$

for all sufficiently large x.

Remark. For $t = 2, 3$, and 4, we can use Table 2.1 to get the limit probabilities for $v = \operatorname{rank} A = t - 1 + e$. For example, when $t = 2$, $\operatorname{Prob}(Z_x = 1)$ is approximately .8889 for large x.

Remark. When $\operatorname{rank} A = t - 1$, it is known that A is an elementary abelian 3-group (cf. [4]). Since $\operatorname{Prob}(Z_x = t-1) > .840$, most cubic Galois extensions of $\mathbf{Q}$ with t ramified primes have elementary abelian 3-class groups with rank $= t - 1$.

From Lemma 2.2, the fact that $v = t - 1 + e$, and the fact that $B_{t,0} < 1$ for $t \geqslant 2$, we get the following result.

PROPOSITION 3.2. *With assumptions as in Proposition* 3.1, $t - 1 < E(Z_x) < t$ *for all sufficiently large x, where $E(Z_x)$ is the expected value of Z_x.*

For these cubic Galois extensions we can also obtain the following result.

PROPOSITION 3.3. *Let assumptions be as in Proposition* 3.1. *Let $L_{t,e,x}$ be the number of elements K in the set S_x whose* 3*-class group has* rank $= t - 1 + e$, *where $0 \leqslant e \leqslant t - 1$. Then*

$$L_{t,e,x} \sim B_{t,e} \cdot \frac{1}{2} \cdot \frac{x(\log\log x)^{t-1}}{(t-1)!\log x}.$$

(Here $F(x) \sim G(x)$ means $F(x)/G(x) \to 1$ as $x \to \infty$.)

Proof. The factor

$$\frac{1}{2}\,\frac{x(\log\log x)^{t-1}}{(t-1)!\log x}$$

is an asymptotic estimate for the number of elements in S_x (see [3] for details), and the factor $B_{t,e}$ is introduced because we are counting only the elements K of S_x that have 3-class group with rank $= t - 1 + e$.

We are now ready to consider primes $p \geqslant 5$. We suppose that K is a Galois extension of $\mathbf{Q}$ of degree p; A is the p-class group of K; t is the number of primes that ramify in $K/\mathbf{Q}$ (and we are assuming $t \geqslant 2$); f_K is the conductor of K. Then employing the techniques from [4], we see that $v = \operatorname{rank} A$ satisfies $t - 1 + e \leqslant v \leqslant t - 1 + e(p-2)$, where $e = t - 1 - r$ and r is the rank of a certain $(t-1) \times t$ matrix over $\mathbf{F}_p$. Thus for $p \geqslant 5$ we have the inequalities $t - 1 + e \leqslant v \leqslant t - 1 + e(p-2)$ instead of the equality $v = t - 1 + e$. However when $e = 0$, we do have the equality $v = t - 1$, and from our calculations in Section 2, we know that the cases $e = 0$ has the highest probability. Using our results from Section 2, we can obtain the following result.

PROPOSITION 3.4. *Let $p \geqslant 5$ be a prime number. Let an integer $t \geqslant 2$ be fixed, and let x be a positive real number. Let S_x be the set of all Galois extensions K of $\mathbf{Q}$ of degree p with exactly t ramified primes over $\mathbf{Q}$ and conductor $f_K \leqslant x$. Assume S_x has counting measure. If Z_x is the random variable which assigns to each $K \in S_x$ the rank of the p-class group A of K, then* $\operatorname{Prob}(Z_x = t - 1) \to B_{t,0}$ *as $x \to \infty$, where $B_{t,0}$ is given by Eq.* (2.2). *In particular, for all sufficiently large x,* $\operatorname{Prob}(Z_x = t - 1) > .950$ (*resp*., .976; *resp*., .99) *when $p = 5$* (*resp*., $p = 7$; *resp*., $p \geqslant 11$). *Furthermore $t - 1 < E(Z_x) < t$ for all sufficiently large x, where $E(Z_x)$ is the expected value of Z_x. Finally if $L_{t,x}$ is the number of elements K in S_x whose p-class group has rank $= t - 1$, then*

$$L_{t,x} \sim B_{t,0} \cdot \frac{1}{p-1} \cdot \frac{x(\log\log x)^{t-1}}{(t-1)!\log x}.$$

Remark. When $\operatorname{rank} A = t - 1$, it is known that A is an elementary abelian p-group (cf. [4]). Thus most Galois extensions of $\mathbf{Q}$ of degree p with t ramified primes have elementary abelian p-class groups with rank $= t - 1$.

Department of Mathematics
The University of Texas
Austin, Texas 78712

1. Z. BOREVICH & I. SHAFAREVICH, *Number Theory*, Academic Press, New York, 1966.
2. F. GERTH, "Consequences of l-th power reciprocity." (To appear.)
3. F. GERTH, "Counting certain number fields with prescribed l-class numbers," *J. Reine Angew. Math.*, v. 337, 1982, pp. 195–207.
4. G. GRAS, "Sur les l-classes d'idéaux dans les extensions cycliques relatives de degré premier l," *Ann. Inst. Fourier* (*Grenoble*), v. 23, no. 3, 1973, pp. 1–48, and v. 23, no. 4, 1973,pp. 1–44.
5. E. INABA, "Über die Struktur der l-Klassengruppe zyklischer Zahlkörper von Primzahlgrad l", *J. Fac. Sci. Imp. Univ. Tokyo*, *Sect.* I, v. 4, 1940, pp. 61–115.
6. L. RÉDEI, "Über einige Mittelwertfragen im quadratischen Zahlkörper," *J. Reine Angew. Math.*, v. 174, 1936, pp. 15–55.