ACTA ARITHMETICA
XII (1966)

The number of solutions of a special system of equations in a finite field*

by

L. Carlitz and Charles Wells (Durham, North Carolina)

1. Introduction. Let $GF(q)$, where $q = p^z$, p prime, denote the finite field of order q and let $a_1, \dots, a_n, b_1, \dots, b_n$ be nonzero numbers of $GF(q)$ such that

$$a_i b_j \neq a_j b_i \quad (i, j = 1, 2, \dots, n,\ i \neq j). \tag{1.1}$$

Let $k, k_1, \dots, k_n$ be fixed positive integers and let N denote the number of solutions $x, y_1, \dots, y_n \in GF(q)$ of the system

$$y_i^{k_i} = a_i + b_i x^k \quad (i = 1, 2, \dots, n). \tag{1.2}$$

We shall prove the following

Theorem 1. *The number of solutions of the system* (1.2) *satisfies*

$$N = q + O(q^{1/2}) \quad (q \to \infty). \tag{1.3}$$

The proof of (1.3) makes use of the Riemann hypothesis ([4]) for an algebraic function field over $GF(q)$. If in place of this we make use of the weaker result of Davenport ([2]), we have

$$N = q + O(q^c), \tag{1.4}$$

for some $c < 1$. This result or indeed the still weaker statement implied by (1.4)

$$\lim_{q=\infty} q^{-1} N = 1 \tag{1.5}$$

suffices for the following application.

Let $de = q-1$ and consider the polynomial

$$f(x) = x(x^d + a) \quad (a \in GF(q)). \tag{1.6}$$

The first author ([1]) stated, and for $e = 2, 3$ proved, the following

* Supported in part by NSF grants GP-1593, GP-1881.

THEOREM 2. *Let e be a fixed divisor of $q-1$, $e>1$. Then for sufficiently large q there exist $a\in GF(q)$ for which $f(x)$ is a permutation polynomial.*

We recall that a polynomial $f(x)\in GF[q,x]$ is a permutation polynomial provided the numbers $f(a)$ for $a\in GF(q)$ are distinct.

2. Some lemmas. Let χ, ψ denote characters of the multiplicative group of $GF(q)$ and put

$$e(a)=e^{2\pi i t(a)} \quad (a\in GF(q)), \tag{2.1}$$

where

$$t(a)=a+a^p+\ldots+a^{p^{s-1}} \quad (q=p^s).$$

Also put

$$\tau(\chi)=\sum_a e(a)\chi(a), \tag{2.2}$$

where the summation is over all $a\in GF(q)$.

LEMMA 1. *We have*

$$|\tau(\chi)|=q^{1/2} \quad (\chi\neq\chi_0), \quad \tau(\chi_0)=-1,$$

where χ_0 denotes the principal character.

Put

$$S(a,t)=\sum_b e(ab^t), \tag{2.3}$$

where $t>0$ and a is any number of $GF(q)$.

LEMMA 2. *For $a\neq 0$,*

$$S(a,t)=\sum_\psi \psi(a)\tau(\bar\psi), \tag{2.4}$$

where the summation is over all non principal characters such that $\psi^t=\psi_0$, the principal character.

Lemmas 1 and 2 are well known.

Now let $r>0$ and let $h_1,\ldots,h_r$ be arbitrary positive integers. For $i=1,\ldots,r$ let ψ_i denote a character satisfying

$$\psi_i^{h_i}=\psi_0 \quad (i=1,\ldots,r). \tag{2.5}$$

For $r>1$ put

$$T_r=T_r(c_1,\ldots,c_r)=\sum_{\lambda_1,\ldots,\lambda_r} e(c_1\lambda_1+\ldots+c_r\lambda_r)\psi_1(\lambda_1)\psi_2(\lambda_2)\ldots\psi_r(\lambda_r), \tag{2.6}$$

where the summation is over all $\lambda_1,\ldots,\lambda_r\in GF(q)$ such that $\lambda_1+\ldots+\lambda_r=0$ and the ψ_i are any nonprincipal characters that satisfy (2.5); also $c_1,\ldots,c_r$ are arbitrary numbers in $GF(q)$.

The principal lemma required in the proof of Theorem 1 can now be stated.

LEMMA 3. *If $\psi_1\psi_2\ldots\psi_r\neq\psi_0$ or if $c_1,c_2,\ldots,c_r$ are not all equal, then*

$$T_r(c_1,\ldots,c_r)=O(q^{(r-1)/2}) \quad (r\geqslant 2). \tag{2.7}$$

Proof. It is convenient to put

$$T_r(\lambda)=\sum_{\lambda_1,\ldots,\lambda_r} e(c_1\lambda_1+\ldots+c_r\lambda_r)\psi_1(\lambda_1)\psi_2(\lambda_2)\ldots\psi_r(\lambda_r) \quad (\lambda\in GF(q)),$$

where now the summation is over all $\lambda_1,\ldots,\lambda_r\in GF(q)$ such that $\lambda_1+\ldots+\lambda_r=\lambda$. Then for $t\in GF(q)$

$$\sum_\lambda T_r(\lambda)e(\lambda t)=\sum_{\lambda_1,\ldots,\lambda_r} e\Big\{\sum_{i=1}^r (c_i+t)\lambda_i\Big\}\prod_{i=1}^r\psi_i(\lambda_i), \tag{2.8}$$

where $\lambda_1,\ldots,\lambda_r$ run through all elements of $GF(q)$.

Since, for $\psi\neq\psi_0$,

$$\sum_\lambda e\big((c+t)\lambda\big)\psi(\lambda)=\bar\psi(c+t)\tau(\psi),$$

where $\bar\psi$ denotes the complex conjugate character, it follows from (2.8) that

$$\sum_\lambda T_r(\lambda)e(\lambda t)=\prod_{i=1}^r\{\bar\psi_i(c_i+t)\tau(\psi_i)\}.$$

Summing over t we get

$$qT_r(0)=\prod_{i=1}^r\tau(\psi_i)\sum_t\prod_{j=1}^r\bar\psi_j(c_j+t). \tag{2.9}$$

Now let h denote the least common multiple of $h_1,\ldots,h_r$ and let ψ be a character such that

$$\psi^h=\psi_0, \quad \psi^j\neq\psi_0 \quad (1\leqslant j<h). \tag{2.10}$$

Then there exist positive integers $s_1,\ldots,s_r$ such that

$$\bar\psi_i=\psi^{s_i} \quad (i=1,\ldots,r).$$

Thus (2.9) becomes

$$qT_r(0)=\prod_{i=1}^r\tau(\psi_i)\sum_t\psi\Big\{\prod_{j=1}^r(c_j+t)^{s_j}\Big\}. \tag{2.11}$$

If $c_1 = \dots = c_r = c$ and $\psi_1\psi_2 \dots \psi_r = \psi_0$, then

$$\psi\Big\{\prod_{j=1}^{r}(c_j+t)^{e_j}\Big\} = \prod_{j=1}^{r}\bar{\psi}_j(c+t) = \psi_0(c+t).$$

Since this is ruled out by the hypotheses of the lemma we have by Weil's theorem ([4])

$$\sum_t \psi\Big\{\prod_{j=1}^{r}(c_j+t)^{e_j}\Big\} = O(q^{1/2}). \tag{2.12}$$

Therefore applying Lemma 1, (2.11) implies

$$qT_r(0) = O(q^{(r+1)/2}).$$

Since

$$T_r(0) = T_r(c_1, \dots, c_r),$$

(2.7) follows at once.

3. Proof of Theorem 1. If N denotes the number of solutions of the system (1.2), we have

$$q^n N = \sum_{\substack{\lambda_1,\dots,\lambda_n \\ x,y_1,\dots,y_n}} e\Big\{\sum_{i=1}^{n}\lambda_i(a_i+b_ix^k-y_i^{k_i})\Big\}, \tag{3.1}$$

the summation extending over all $\lambda_1, \dots, \lambda_n, x, y_1, \dots, y_n \in \mathrm{GF}(q)$. It is convenient to rewrite (3.1) in the form

$$q^n N = \sum_{\lambda_1,\dots,\lambda_n} e\Big(\sum_{i=1}^{n}\lambda_i a_i\Big) \sum_{x,y_1,\dots,y_n} e\Big(\sum_{i=1}^{n}\lambda_i b_i x^k\Big)\, e\Big(-\sum_{i=1}^{n}\lambda_i y_i^{k_i}\Big). \tag{3.2}$$

The right member of (3.2) may be broken up according to the number of λ_i that are nonzero. Thus it consists of $\binom{n}{r}$ terms of the type

$$q^{n-r}\sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big) \sum_{x,y_1,\dots,y_r} e\Big(\sum_{i=1}^{r}\lambda_i b_i x^k\Big)\, e\Big(-\sum_{i=1}^{r}\lambda_i y_i^{k_i}\Big). \tag{3.3}$$

We further decompose (3.3) according to whether

$$\lambda_1 b_1 + \dots + \lambda_r b_r = 0$$

or not; we then have terms of the form

$$q^{n-r}\sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big) \sum_{x,y_1,\dots,y_r} e\Big(\sum_{i=1}^{r}\lambda_i b_i x^k\Big)\, e\Big(-\sum_{i=1}^{r}\lambda_i y_i^{k_i}\Big) = q^{n-r}N_r \tag{3.4}$$

$$(0 \leqslant r \leqslant n)$$

and of the form

$$q^{n-r+1}\sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big) \sum_{y_1,\dots,y_r} e\Big(-\sum_{i=1}^{r}\lambda_i y_i^{k_i}\Big) = q^{n-r+1}M_r \quad (0 < r \leqslant n). \tag{3.5}$$

In both (3.4) and (3.5) the λ_i are restricted to nonzero values.

Now, by (2.3), we have

$$N_r = \sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big)\, S\Big(\sum_{i=1}^{r}\lambda_i b_i, k\Big)\prod_{i=1}^{r} S(-\lambda_i, k_i).$$

By Lemma 2 this becomes

$$\begin{aligned} N_r &= \sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big)\sum_{\psi}\psi\Big(\sum_{i=1}^{r}\lambda_i b_i\Big)\tau(\bar{\psi})\prod_{i=1}^{r}\sum_{\psi_i}\psi_i(-\lambda_i)\tau(\bar{\psi}_i) \\ &= \sum_{\psi,\psi_1,\dots,\psi_r}\tau(\bar{\psi})\tau(\bar{\psi}_1)\dots\tau(\bar{\psi}_r)\sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big)\,\psi\Big(\sum_{i=1}^{r}\lambda_i b_i\Big)\times \\ &\qquad \times\psi_1(-\lambda_1)\dots\psi_r(-\lambda_r), \end{aligned}$$

where ψ ranges over the nonprincipal characters that satisfy $\psi^k = \psi_0$ and ψ_i ranges over the nonprincipal characters that satisfy $\psi_i^{k_i} = \psi_0$. The inner sum on the extreme right is equal to

$$\psi_1(b_1^{-1})\dots\psi_r(b_r^{-1})T_{r+1}(-b_1^{-1}a_1, \dots, -b_r^{-1}a_r, 0).$$

Applying Lemmas 2 and 3 we get

$$N_r = O(q^{(r+1)/2+r/2}) = O(q^{r+1/2}) \quad (0 < r \leqslant n). \tag{3.6}$$

On the other hand, where $r = 0$, we have

$$N_0 = q. \tag{3.7}$$

In the second place, we have

$$\begin{aligned} M_r &= \sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big)\, S(-\lambda_1, k_1)\dots S(-\lambda_r, k_r) \\ &= \sum_{\psi_1,\dots,\psi_r}\tau(\bar{\psi}_1)\dots\tau(\bar{\psi}_r)\sum_{\lambda_1,\dots,\lambda_r} e\Big(\sum_{i=1}^{r}\lambda_i a_i\Big)\,\psi_1(-\lambda_1)\dots\psi_r(-\lambda_r). \end{aligned}$$

The inner sum on the extreme right is equal to

$$\psi_1(b_r b_1^{-1})\dots\psi_{r-1}(b_r b_{r-1}^{-1})\psi_r(-1)T_r(-b_r b_1^{-1}a_1, \dots, -b_r b_{r-1}^{-1}a_{r-1}, -a_r).$$

In view of (1.1) we may apply Lemma 3 to T_r. Thus by Lemmas 2 and 3

$$M_r = O(q^{(r-1)/2+r/2}) = O(q^{r-1/2}) \quad (0 < r \leqslant n). \tag{3.8}$$

Combining (3.3), (3.4), (3.5), (3.6), (3.7), (3.8) we get

$$q^n N = q^{n+1} + O(q^{n+1/2}).$$

This evidently completes the proof of the theorem.

4. As remarked in the Introduction, if in place of Weil's theorem we use a weaker result we can prove that

$$N = q + O(q^{\Theta}) \tag{4.1}$$

for some $\Theta < 1$. Indeed the only significant change occurs in (2.12). Thus if we assume only that ([2])

$$\sum_t \psi\Big\{\prod_{j=1}^{r} (c_j + t)^{s_j}\Big\} = O(q^{\Theta}), \tag{4.2}$$

then (2.7) becomes

$$T_r(c_1, \ldots, c_r) = O(q^{(r-2)/2+\Theta}). \tag{4.3}$$

It is now easily verified that (4.2) implies (4.1).

5. Application to permutation polynomials. We shall now apply Theorem 1 to prove the existence of permutation polynomials of a certain type. Let $de = q-1$ and define

$$f(x) = x(x^d + a). \tag{5.1}$$

We shall show that for e fixed, $e > 1$ and q sufficiently large there exist numbers $a \in \mathrm{GF}(q)$ such that $f(x)$ permutes the elements of $\mathrm{GF}(q)$.

To prove this let w be a primitive eth root of unity in $\mathrm{GF}(q)$. Consider the following system of equations:

$$(1-w^s)x^e + (w-1)y_s^e + w^s - w = 0 \quad (s = 2, \ldots, e-1). \tag{5.2}$$

This is a system of the type (1.2). The hypothesis (1.1) now becomes

$$(1-w^s)(w^r-w) - (1-w^r)(w^s-w) \neq 0 \quad (r, s = 2, \ldots, e-1; r \neq s).$$

This reduces to

$$w(w^s - w^r) \neq w^s - w^r \quad (r, s = 2, \ldots, e-1; r \neq s),$$

which is automatically satisfied for $e > 1$.

It therefore follows from Theorem 1 that for q sufficiently large there exists at least one solution of the system (5.2). Put

$$a = \frac{x^e - w}{1 - x^e}. \tag{5.3}$$

Rearranging (5.2) we have

$$w^s(1-x^e) + x^e - w = y_s^e(1 - x^e + x^e - w) \quad (s = 2, \ldots, e-1);$$

in view of (5.3) this becomes

$$w^s + a = y_s^e(1+a) \quad (s = 2, \ldots, e-1).$$

This implies

$$(w^s + a)^d = (w^r + a)^d \quad (r, s = 1, 2, \ldots, e-1). \tag{5.4}$$

Now assume that $f(\xi) = f(\eta)$ for some pair $\xi, \eta \in \mathrm{GF}(q)$. Then by (5.1)

$$\xi^d(\xi^d + a) = \eta^d(\eta^d + a)^d. \tag{5.5}$$

Since w is a primitive eth root of unity in $\mathrm{GF}(q)$, it is evident that $\xi^d = w^r$, $\eta^d = w^s$ for some integers r and s. Thus (5.5) becomes

$$w^r(w^r + a)^d = w^s(w^s + a)^d;$$

because of (5.4) this reduces to $w^r = w^s$ and therefore $\xi^d = \eta^d$. Substitution in $f(\xi) = f(\eta)$ now yields $\xi = \eta$. Hence $f(x)$ as defined by (5.1) and (5.3) is a permutation polynomial. This completes the proof of Theorem 2.

6. Some additional remarks. If $(c, q-1) = 1$ it is well known that x^c is a permutation polynomial. Using this fact one may prove using an argument like the above that for a defined by (5.3), the polynomial

$$f(x) = x^c(x^d + a)^k \tag{6.1}$$

is a permutation polynomial, where k is an arbitrary integer. We may therefore state the following result.

THEOREM 3. *Let $de = q-1$, $(c, q-1) = 1$ and k arbitrary. Then for sufficiently large q there exist permutation polynomials of the form*

$$f(x) = x^c g(x^d),$$

where $g(x)$ is a polynomial of degree k.

Let c, r, s be positive integers. We note that if $(r, s) \neq 1$ and $(r, s) | q-1$ then

$$g(x) = x^r + ax^s \tag{6.2}$$

is *not* a permutation polynomial for any a in $\mathrm{GF}(q)$. Indeed if $d = (r, s)$, $d | q-1$, $d > 1$, then there are distinct numbers ξ, η in $\mathrm{GF}(q)$ such that $\xi^d = \eta^d$. Then clearly (6.2) implies $g(\xi) = g(\eta)$.

We note also that

$$h(x) = x^{c+s} - ax^s \quad (s > 0, a \neq 0) \tag{6.3}$$

is not a permutation polynomial if either $(c, q-1) = 1$ or a is a cth power in $\mathrm{GF}(q)$. For if $(c, q-1) = 1$ there is a unique ξ in $\mathrm{GF}(q)$ such that $\xi^c = a$. This gives $h(0) = h(\xi) = 0$. If $(c, q-1) > 1$ and $a = \xi^c$ we again have $h(0) = h(\xi) = 0$.

References

[1] L. Carlitz, *Some theorems on permutation polynomials*, Bull. Amer. Math. Soc. 68 (1962), pp. 120-122.

[2] H. Davenport, *On character sums in finite fields*, Acta Math. 71 (1939), pp. 99-121.

[3] L. E. Dickson, *Linear Groups with an Exposition of Galois Field Theory*, New York, Dover, 1958.

[4] A. Weil, *On the Riemann hypothesis in function fields*, Proc. Nat. Acad. Sci. 27 (1941), pp. 97-98.

Reçu par la Rédaction le 19. 1. 1966

ACTA ARITHMETICA
XII (1966)

Further developments in the comparative prime-number theory VI

Accumulation theorems for residue-classes representing quadratic residues mod k

by

S. Knapowski (Poznań) and P. Turán (Budapest)

1. In this paper we return to the "modified Abelian means", introduced in paper [2] and further studied throughout [3] and [4]. Our present aim is to compare, in the sense of this means, the number of primes belonging to progressions $\equiv l_1 \pmod k$ resp. $\equiv l_2 \pmod k$, where *both* l_1 and l_2 are quadratic residues mod k. As before, we have to assume the Haselgrove-condition: there is an $E = E(k) > 0$ such that none of the $L(s, \chi)$-functions mod k vanishes in

$$\sigma \geqslant \tfrac{1}{2}, \quad |t| \leqslant E(k), \quad s = \sigma + it. \tag{1.1}$$

In addition to (1.1), we have to assume what we call "a finite Riemann-Piltz hypothesis": with a suitable η satisfying [1]

$$0 < \eta < \min\left(c_1, \left(\frac{E(k)}{8\pi}\right)^2\right) \tag{1.2}$$

none of the $L(s, \chi)$-functions mod k vanishes in

$$\sigma > \frac{1}{2}, \quad |t| \leqslant \frac{3}{\sqrt{\eta}}. \tag{1.3}$$

There is no loss of generality in supposing

$$E(k) \leqslant k^{-15}; \tag{1.4}$$

this and (1.2) give automatically

$$\eta < k^{-30}. \tag{1.5}$$

With these provisions we can state the following:

[1] c_1 and later $c_2, \ldots$ stand for positive numerical constants.