ANNALES ACADEMIAE SCIENTIARUM FENNICAE

Series A

I. MATHEMATICA

386

ON SYSTEMS OF EQUATIONS IN FINITE FIELDS

BY

AIMO TIETÄVÄINEN

HELSINKI 1966
SUOMALAINEN TIEDEAKATEMIA

https://doi.org/10.5186/aasfm.1966.386

Communicated 11 February 1966 by P. J. MYRBERG and K. INKERI

KESKUSKIRJAPAINO
HELSINKI 1966

On systems of equations in finite fields

1. Introduction. Let K be a finite field of q elements where $q = p^n$, p is a prime and n a positive integer. Let $f_{ij}(\xi_j)$ be a polynomial of degree c_{ij} with coefficients in K such that $f_{ij}(0) = 0$ and $f_{ij}(-\alpha) = -f_{ij}(\alpha)$, for every element α of K. Let, furthermore, K_j be a subset of K such that (i) $0 \in K_j$, (ii) $\alpha \in K_j$ implies $-\alpha \in K_j$, and (iii) q_j, the number of elements in K_j, is > 1. We study the non-trivial solvability of the system

$$(1) \qquad \sum_{j=1}^{s} f_{ij}(\xi_j) = 0 \,, \; \xi_j \in K_j \quad (i = 1, \ldots, t) \,,$$

using exponential sums $\sum'_{\xi_j} e(\mathbf{kf}_j(\xi_j))$ where $\mathbf{kf}_j(\xi_j) = \sum_{i=1}^{t} \varkappa_i f_{ij}(\xi_j)$, $e(\alpha) = e^{2\pi i \operatorname{tr}(\alpha)/p}$, $\operatorname{tr}(\alpha)$ is the absolute trace of α, and the summation $\sum'_{\xi_j}$ is over all the elements of K_j. Our main result is

Theorem 1. *Let $r_1, \ldots, r_s$ be real numbers such that*

$$(2) \qquad \sum_{\xi_j}{}' e(\mathbf{kf}_j(\xi_j)) \geqq -r_j \,,$$

for every $\mathbf{k}$. *Then the system* (1) *has a non-trivial solution* $(\xi_1, \ldots, \xi_s)$ *if*

$$(3) \qquad \prod_{j=1}^{s} (q_j + r_j) > q^t \prod_{j=1}^{s} (r_j + 1) \,.$$

As consequences of this theorem we find some results which extend, improve, or sharpen previous results of Chevalley [2], Lewis [10], Gray [9], Chowla ([3]—[8]), Shimura [8], and Tietäväinen ([12], [13]). As a simple example of them we mention here the following corollary of theorem 5.

Let d, the g.c.d. *of c and $q - 1$, be odd. Then the system*

$$\sum_{j=1}^{s} \gamma_{ij} \xi_j^c = 0 \quad (i = 1, \ldots, t)$$

has a non-trivial solution $(\xi_1, \ldots, \xi_s)$ in K if

$$s \geqq 2\,t(1 + \max(\log_2(d-1), 1)) \,.$$

2. Preliminary remarks. Let V be the space of t-tuples over K. Let $\mathbf{a} = (\alpha_1, \ldots, \alpha_t)$ and $\mathbf{b} = (\beta_1, \ldots, \beta_t)$ be elements of V and α an element of K. Define, as usual,

$$\mathbf{a} + \mathbf{b} = (\alpha_1 + \beta_1, \ldots, \alpha_t + \beta_t),$$

$$\alpha\mathbf{a} = (\alpha\alpha_1, \ldots, \alpha\alpha_t),$$

and

$$\mathbf{ab} = \alpha_1\beta_1 + \cdots + \alpha_t\beta_t.$$

The 0-element $(0, \ldots, 0)$ of V will be denoted by $\mathbf{0}$.

Define the trace of α as

$$\mathrm{tr}\,(\alpha) = \alpha + \alpha^p + \cdots + \alpha^{p^{n-1}}$$

so that $\mathrm{tr}(\alpha)$ may be considered as an integer (mod p). Define, furthermore,

$$e(\alpha) = e^{2\pi i \mathrm{tr}(\alpha)/p}.$$

Then (see [13], section 3)

$$e(\mathbf{k}(\mathbf{a} + \mathbf{b})) = e(\mathbf{ka})\, e(\mathbf{kb}), \tag{4}$$

for every element $\mathbf{k}$ of V, and, moreover,

$$\sum_{\mathbf{k}} e(\mathbf{ka}) = \begin{cases} q^t \text{ if } \mathbf{a} = \mathbf{0}, \\ 0 \text{ if } \mathbf{a} \neq \mathbf{0}. \end{cases} \tag{5}$$

Here and hereafter, in the sums of type $\sum_{\mathbf{k}}$ and $\sum_{\mathbf{k}\neq\mathbf{0}}$ the summation is over all the elements of V and over all the non-zero elements of V, respectively. Furthermore, in the sums of type $\sum_{\xi_j}$, $\sum'_{\xi_j}$, and $\sum'_{\xi_j\neq 0}$ the variable runs through all the elements of K, through all the elements of K_j, and through all the non-zero elements of K_j, respectively.

Denote

$$\mathbf{f}_j(\xi_j) = (f_{1j}(\xi_j), \ldots, f_{tj}(\xi_j)).$$

Then the system (1) may be written in the form

$$\sum_{j=1}^{s} \mathbf{f}_j(\xi_j) = \mathbf{0}, \ \xi_j \in K_j.$$

It is easy to show that the exponential sum $\sum'_{\xi_j} e(\mathbf{kf}_j(\xi_j))$ is real, for every element $\mathbf{k}$ of V. Indeed, we have, by the definitions of K_j and $\mathbf{f}_j(\xi_j)$,

$$\sum_{\xi_j}{}' e(\mathbf{kf}_j(\xi_j)) = \sum_{\xi_j}{}' e(\mathbf{kf}_j(-\xi_j)) = \sum_{\xi_j}{}' e(-\mathbf{kf}_j(\xi_j)) = \overline{\sum_{\xi_j}{}' e(\mathbf{kf}_j(\xi_j))}$$

where $\bar{z}$ denotes the complex conjugate of z.

3. Proof of theorem 1. Let

$$J = J(\xi_1, \ldots, \xi_s) = \{j \in \{1, \ldots, s\} \mid \xi_j = 0\},$$

(6) $$A(\xi_1, \ldots, \xi_s) = \begin{cases} 1 \text{ if } \xi_j \neq 0, \text{ for every } j, \\ \prod_{j \in J} (r_j + 1) \text{ otherwise,} \end{cases}$$

and

(7) $$B(\xi_1, \ldots, \xi_s) = \begin{cases} A(\xi_1, \ldots, \xi_s) \text{ if } \sum_{j=1}^{s} \mathbf{f}_j(\xi_j) = \mathbf{0}, \\ 0 \text{ otherwise.} \end{cases}$$

Let, furthermore,

(8) $$M = \sum_{\xi_1}{}' \cdots \sum_{\xi_s}{}' B(\xi_1, \ldots, \xi_s).$$

Then (1) has a non-trivial solution if $M > \prod_{j=1}^{s} (r_j + 1)$.

We have, by (5), (7), (8), and (4),

(9) $$\begin{aligned} q^t M &= \sum_{\xi_1}{}' \cdots \sum_{\xi_s}{}' A(\xi_1, \ldots, \xi_s) \sum_{\mathbf{k}} e(\mathbf{k} \sum_{j=1}^{s} \mathbf{f}_j(\xi_j)) \\ &= \sum_{\mathbf{k}} \sum_{\xi_1}{}' \cdots \sum_{\xi_s}{}' A(\xi_1, \ldots, \xi_s) \prod_{j=1}^{s} e(\mathbf{k}\mathbf{f}_j(\xi_j)). \end{aligned}$$

It can be shown, by induction, that

(10) $$\sum_{\xi_1}{}' \cdots \sum_{\xi_s}{}' A(\xi_1, \ldots, \xi_s) \prod_{j=1}^{s} e(\mathbf{k}\mathbf{f}_j(\xi_j)) = \prod_{j=1}^{s} (r_j + \sum_{\xi_j}{}' e(\mathbf{k}\mathbf{f}_j(\xi_j))).$$

Indeed, it is easy to see that the statement (10) is true for $s = 1$, and we assume it to be true for $s - 1$ variables ξ_j. Since, by (6),

$$A(\xi_1, \ldots, \xi_s) = \begin{cases} (r_s + 1) A(\xi_1, \ldots, \xi_{s-1}) \text{ if } \xi_s = 0, \\ A(\xi_1, \ldots, \xi_{s-1}) \text{ if } \xi_s \neq 0, \end{cases}$$

then the left side of (10) equals

$$(r_s + 1) \sum_{\xi_1}{}' \cdots \sum_{\xi_{s-1}}{}' A(\xi_1, \ldots, \xi_{s-1}) \prod_{j=1}^{s-1} e(\mathbf{k}\mathbf{f}_j(\xi_j)) +$$

$$\sum_{\xi_1}{}' \cdots \sum_{\xi_{s-1}}{}' A(\xi_1, \ldots, \xi_{s-1}) \prod_{j=1}^{s-1} e(\mathbf{k}\mathbf{f}_j(\xi_j)) \sum_{\xi_s \neq 0}{}' e(\mathbf{k}\mathbf{f}_s(\xi_s)).$$

Using the equation

$$\sum_{\xi_s \neq 0}{}' e(\mathbf{k}\mathbf{f}_s(\xi_s)) = \sum_{\xi_s}{}' e(\mathbf{k}\mathbf{f}_s(\xi_s)) - 1$$

and the induction hypothesis, we find that this is, moreover, equal to

$$(r_s + \sum_{\xi_s}{}' e(\mathbf{k}\mathbf{f}_s(\xi_s)) \sum_{\xi_1}{}' \cdots \sum_{\xi_{s-1}}{}' A(\xi_1, \ldots, \xi_{s-1}) \prod_{j=1}^{s-1} e(\mathbf{k}\mathbf{f}_j(\xi_j))$$

$$= \prod_{j=1}^{s} (r_j + \sum_{\xi_j}{}' e(\mathbf{k}\mathbf{f}_j(\xi_j))) .$$

Thus we have proved the equation (10).

Using (9) and (10), we get

$$q^t M = \sum_{\mathbf{k}} \prod_{j=1}^{s} (r_j + \sum_{\xi_j}{}' e(\mathbf{k}\mathbf{f}_j(\xi_j)))$$

$$= \prod_{j=1}^{s} (q_j + r_j) + \sum_{\mathbf{k} \neq 0} \prod_{j=1}^{s} (r_j + \sum_{\xi_j}{}' e(\mathbf{k}\mathbf{f}_j(\xi_j))) .$$

We have hence, by (2) and (3).

$$M \geqq q^{-t} \prod_{j=1}^{s} (q_j + r_j) > \prod_{j=1}^{s} (r_j + 1)$$

which is the required inequality.

4. Consequences of theorem 1. Since $e(\mathbf{k}\mathbf{f}_j(0)) = e(0) = 1$ then $\sum_{\xi_j}{}' e(\mathbf{k}\mathbf{f}_j(\xi_j)) \geqq 2 - q_j$. Therefore we may take $r_j = q_j - 2$ in theorem 1. Then

$$\prod_{j=1}^{s} (q_j + r_j) = 2^s \prod_{j=1}^{s} (q_j - 1) = 2^s \prod_{j=1}^{s} (r_j + 1) .$$

Consequently we have the following corollary of theorem 1.

Theorem 2. *The system* (1) *has a non-trivial solution if*

$$2^s > q^t .$$

This theorem is an extension of a result of CHOWLA's (see, for example, [5]) .For some related theorems, see [11], theorem 1, and [13], lemma 3. Theorem 2 can be proved also by using CHOWLA's method but it is interesting to see that all the theorems 1—5 can be proved by using exponential sum methods only.

If we put $K_1 = \cdots = K_s = K$, we get immediately, by theorem 1, the following result.

Theorem 3. *Let* $r_1, \ldots, r_s$ *be real numbers such that* $\sum_{\xi_j} e(\mathbf{k}\mathbf{f}_j(\xi_j)) \geqq - r_j$, *for every* $\mathbf{k}$. *Then the system*

$$\sum_{j=1}^{s} f_{ij}(\xi_j) = 0 \quad (i = 1, \ldots, t) \tag{11}$$

has a non-trivial solution in K *if*

$$\prod_{j=1}^{s} (q + r_j) > q^t \prod_{j=1}^{s} (r_j + 1) .$$

CARLITZ and UCHIYAMA [1] have proved

Lemma 1. *The inequality*

$$|\sum_{\xi} e(f(\xi))| \leqq (c - 1)q^{\frac{1}{2}}$$

holds on the assumption that f *is a polynomial of degree* c *over* K *such that*

$$f \neq g^p - g + \beta ,$$

for every polynomial g *over* K *and for every element* β *of* K.

In the following theorem we must suppose, because of the assumption of lemma 1, that the system (11) satisfies the subsequent condition (cf. [13]).

Condition B. *For any value of* j *no non-zero linear combination of the polynomials* $f_{1j}, \ldots, f_{tj}$ *over* K *can be written in the form* $g^p - g + \beta$ *where* g *is a polynomial over* K *and* β *is an element of* K.

It should be noted that condition B is satisfied at least in the case where $c_{ij} \leqq p - 1$, for every i and j. Therefore (see [13]) condition B is no restriction in prime fields.

Define the degree of the 0-polynomial as $-\infty$ and suppose that there exists at least one non-zero polynomial $f_{ij}(\xi_j)$. Combining theorem 2 with theorem 3 and lemma 1, we then find

Theorem 4. *Assume that the system* (11) *satisfies condition* B. *Then it has a non-trivial solution in* K *if*

$$s \geqq 2\, t(1 + \max(\log_2(c - 1), 1)) \tag{12}$$

where $c = \max c_{ij}$.

This theorem sharpens theorem 1 of [13]. For some related results, see corollary 1 of theorem 2 of [13] and theorems I and II of [12]. For small values of c our method gives better results than that mentioned in theorem 4. For example, we may replace the inequality (12) by $s \geqq 1 + 2t$ in case $c = 2$ and by $s \geqq 3t$ in case $c = 3$.

Proof of theorem 4. If $c \leqq 2$, our assertion is a consequence of a well-known result of CHEVALLEY's [2] (and it is easy to prove also by a slight modification of the following proof). Therefore we may assume that $c \geqq 3$.

Suppose that, contrary to our assertion, the system (11) has only the trivial solution in K. Then we have, by theorem 2,

$$2^s \leqq q^t .$$

Combining this with (12), we find

$$q^{s-2t} \geqq (c-1)^{2s} . \tag{13}$$

We may take, by lemma 1, $r_j = (c-1)q^{\frac{1}{2}}$, for every j. Then

$$\begin{aligned}\prod_{j=1}^{s}(q + r_j) &= q^{\frac{1}{2}s}(q^{\frac{1}{2}} + (c-1))^s \\ &= q^{\frac{1}{2}s}(c-1)^{-s}((c-1)q^{\frac{1}{2}} + (c-1)^2)^s \\ &> q^t \cdot q^{\frac{1}{2}(s-2t)}(c-1)^{-s}((c-1)q^{\frac{1}{2}} + 1)^s\end{aligned}$$

from which we get, by (13),

$$\prod_{j=1}^{s}(q + r_j) > q^t((c-1)q^{\frac{1}{2}} + 1)^s = q^t \prod_{j=1}^{s}(r_j + 1) .$$

This is, by theorem 3, an impossible inequality. Hence theorem 4 is true.

We say (cf. [13]) that the system

$$\sum_{j=1}^{s} \gamma_{ij}\, \xi_j^c = 0 \quad (i = 1, \ldots, t) , \tag{14}$$

where c is a positive integer, is an A-system if -1 is a cth power in K (for $t = 1$, cf. paper [5] by CHOWLA). Using the same method as in the proof of theorem 4, we can prove

Theorem 5. *The* A-*system* (14) *has a non-trivial solution in* K *if*

$$s \geqq 2t(1 + \max(\log_2(d-1), 1))$$

where d *is the* g.c.d. *of* c *and* $q - 1$.

Theorem 5 is an extension of some results by CHOWLA ([3]—[8]) and SHIMURA [8] and an improvement for theorem 4 of [13] (see also theorem III of [12]). It is, practically, a corollary of our theorem 4. It should be noted, however, that in the proof of theorem 5 we may use, in place of the deep lemma 1, the following well-known lemma 2 which can be proved elementarily.

Lemma 2. *If* γ *is a non-zero element of* K *then*

$$\left|\sum_{\xi} e(\gamma \xi^c)\right| \leqq (d - 1)q^{\frac{1}{2}}$$

where d *is the* g.c.d. *of* c *and* $q - 1$.

Theorem 5 implies immediately

Corollary. *Let* d , *the* g.c.d. *of* c *and* $q - 1$, *be odd. Then the system* (14) *has a non-trivial solution in* K *if*

$$s \geqq 2\,t(1 + \max(\log_2(d - 1)\,,\,1))\,.$$

University of Turku
Turku, Finland

References

[1] CARLITZ, L. and UCHIYAMA, S.: Bounds for exponential sums. - Duke Math. J. 24 (1957), 37–41.

[2] CHEVALLEY, C.: Démonstration d'une hypothèse de M. Artin. - Abh. Math. Sem. Univ. Hamburg 11 (1936), 73–75.

[3] CHOWLA, S.: Some results in number theory. - Norske Vid. Selsk. Forh. (Trondheim) 33 (1960), 43–44.

[4] —»— A generalization of Meyer's theorem on indefinite quadratic forms in five or more variables. - J. Ind. Math. Soc. 25 (1961), 41.

[5] —»— On the congruence $\sum_{i=1}^{s} a_i x_i^k \equiv 0 \pmod{p}$. - J. Ind. Math. Soc. 25 (1961), 47–48.

[6] —»— On a conjecture of Artin (I). - Norske Vid. Selsk. Forh. (Trondheim) 36 (1963), 135–138.

[7] —»— On a conjecture of Artin (II). - Norske Vid. Selsk. Forh. (Tronheim) 36 (1963), 139–141.

[8] —»— and SHIMURA, G.: On the representation of zero by a linear combination of kth powers. - Norske Vid. Selsk. Forh. (Trondheim) 36 (1963), 169–176.

[9] GRAY, J. F.: Diagonal forms of odd degree over a finite field. - Michigan Math. J. 7 (1960), 297–302.

[10] LEWIS, D. J.: Cubic congruences. - Michigan Math. J. 4 (1957), 85–95.

[11] STEVENS, H.: Linear homogeneous equations over finite rings. - Canad. J. Math. 16 (1964), 532–538.

[12] TIETÄVÄINEN, A.: On the non-trivial solvability of some systems of equations in finite fields. - Ann. Univ. Turku., Ser. A I 71 (1964).

[13] —»— On the non-trivial solvability of some equations and systems of equations in finite fields. - Ann. Acad. Sci. Fenn., Ser. A I 360 (1965).

Printed March 1966.