ACTA ARITHMETICA
142.4 (2010)

Roth's theorem on systems of linear forms in function fields

by

YU-RU LIU (Waterloo),
CRAIG V. SPENCER (Princeton, NJ, and Manhattan, KS) and
XIAOMEI ZHAO (Waterloo)

1. Introduction. For $r,s\in\mathbb{N}=\{1,2,\ldots\}$ with $s\ge 2r+1$, let $(b_{i,j})$ be an $r\times s$ matrix whose elements are integers. Suppose that $b_{i,1}+\cdots+b_{i,s}=0$ $(1\le i\le r)$. Suppose further that among the columns of the matrix, there exist r linearly independent columns such that, if any of the r columns are removed, the remaining $n-1$ columns of the matrix can be divided into two sets so that among the columns of each set there are r linearly independent columns. For $k\in\mathbb{N}$, denote by $D([1,k])$ the maximal cardinality of an integer set $A\subseteq[1,k]$ such that the equations $b_{i,1}x_1+\cdots+b_{i,s}x_s=0$ $(1\le i\le r)$ are never satisfied simultaneously by distinct elements $x_1,\ldots,x_s\in A$. Using techniques similar to his work on sets free of three-term arithmetic progressions (see [4]), Roth [5] showed that

$$D([1,k])\ll k/(\log\log k)^{1/r^2}.$$

In this paper, we will build upon the methods in [2] to study an analogous question in function fields.

Let $\mathbb{F}_q[t]$ denote the ring of polynomials over the finite field $\mathbb{F}_q$. For $N\in\mathbb{N}$, let $\mathcal{S}_N$ denote the subset of $\mathbb{F}_q[t]$ containing all polynomials of degree strictly less than N. For $R,S\in\mathbb{N}$ with $S\ge 2R+1$, let $Y=(a_{i,j})$ be an $R\times S$ matrix with elements in $\mathbb{F}_q$. Suppose that Y satisfies the following two conditions.

CONDITION 1. $a_{i,1}+\cdots+a_{i,S}=0$ $(1\le i\le R)$.

CONDITION 2. Y has L columns with $L\ge R$ such that:

- any R of these L columns are linearly independent,

2010 *Mathematics Subject Classification*: Primary 11P55; Secondary 11T55.
Key words and phrases: Roth's theorem, function fields, circle method.

DOI: 10.4064/aa142-4-6

- after removing any $L-R+1$ of these L columns from Y, we can find two disjoint sets of R linearly independent columns among the remaining $S-L+R-1$ columns,
- without loss of generality, we may assume that these L columns are the first L columns of Y.

Consider the system of equations

$$a_{i,1}x_1 + \cdots + a_{i,S}x_S = 0 \quad (1 \le i \le R). \tag{1.1}$$

Let $D_Y(\mathcal{S}_N)$ denote the maximal cardinality of a set $A \subseteq \mathcal{S}_N$ for which the equations in (1.1) are never satisfied simultaneously by distinct elements $x_1, \ldots, x_S \in A$. We write $|V|$ for the cardinality of a set V. In this paper, we employ a variant of the Hardy–Littlewood circle method for $\mathbb{F}_q[t]$ to prove the following result.

THEOREM 1.1. *Assume that Y satisfies Conditions* 1 *and* 2. *There exists an effectively computable constant $C = C(Y) > 0$ such that for $N \in \mathbb{N}$,*

$$D_Y(\mathcal{S}_N) \le q^N \left(\frac{C}{N}\right)^{(L-R+1)/R}.$$

We note that the assumptions in Condition 2 are more general than the corresponding assumptions in [5]. Thus, in the special case when $L = R$, we can derive from Theorem 1.1 a function field analogue of Roth's theorem. In addition, on rewriting the upper bound we obtain in Theorem 1.1 as

$$D_Y(\mathcal{S}_N) \ll \frac{|\mathcal{S}_N|}{(\log_q |\mathcal{S}_N|)^{(L-R+1)/R}},$$

we observe that this result is much sharper than its integer analogue. Our improvement comes from a better estimate of an exponential sum in $\mathbb{F}_q[t]$ than in $\mathbb{Z}$ (see Lemma 2.4).

One can also obtain from Theorem 1.1 some information about irreducible polynomials. Let $\mathcal{P}_N$ denote the set of all monic irreducible polynomials in $\mathbb{F}_q[t]$ of degree strictly less than N, and let A_N denote a subset of $\mathcal{P}_N$. By the prime number theorem for $\mathbb{F}_q[t]$ (see [3, Theorem 2.2]), we have $|\mathcal{P}_N| \asymp q^N/N$. If $L+1 > 2R$, Theorem 1.1 implies that there exists a positive constant $E(Y)$ such that whenever

$$\frac{|A_N|}{|\mathcal{P}_N|} \ge \frac{E(Y)}{N^{(L-2R+1)/R}},$$

then (1.1) has a solution with distinct elements $x_1, \ldots, x_S \in A_N$.

We conclude this section by introducing the Fourier analysis of $\mathbb{F}_q[t]$. Let $\mathbb{K} = \mathbb{F}_q(t)$ be the field of fractions of $\mathbb{F}_q[t]$, and let $\mathbb{K}_\infty = \mathbb{F}_q((1/t))$ be the completion of $\mathbb{K}$ at ∞. We may write each element $\alpha \in \mathbb{K}_\infty$ in the shape $\alpha = \sum_{i \le v} a_i t^i$ for some $v \in \mathbb{Z}$ and $a_i = a_i(\alpha) \in \mathbb{F}_q$ $(i \le v)$. If $a_v \ne 0$, we define $\operatorname{ord} \alpha = v$. We adopt the convention that $\operatorname{ord} 0 = -\infty$. Also, it is often convenient to refer to a_{-1} as being the residue of α, denoted by $\operatorname{res}\alpha$. Consider the compact additive subgroup $\mathbb{T}$ of $\mathbb{K}_\infty$ defined by $\mathbb{T} = \{\alpha \in \mathbb{K}_\infty \mid \operatorname{ord}\alpha < 0\}$. Given any Haar measure $d\alpha$ on $\mathbb{K}_\infty$, we normalize it in such a manner that $\int_{\mathbb{T}} 1\, d\alpha = 1$. We now extend the measure to $\mathbb{K}_\infty^R$ by the standard product measure. Thus, if $\mathfrak{M}$ is the subset of $\mathbb{K}_\infty^R$ defined by

$$\mathfrak{M} = \{\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_R) \in \mathbb{K}_\infty^R \mid \operatorname{ord}\alpha_i < -N \ (1 \le i \le R)\},$$

then the measure of $\mathfrak{M}$, written $\operatorname{mes}(\mathfrak{M})$, is equal to q^{-NR}.

We are now equipped to define the exponential function on $\mathbb{F}_q[t]$. Suppose that the characteristic of $\mathbb{F}_q$ is p. Let $e(z)$ denote $e^{2\pi i z}$, and let $\operatorname{tr} : \mathbb{F}_q \to \mathbb{F}_p$ denote the familiar trace map. There is a non-trivial additive character $e_q : \mathbb{F}_q \to \mathbb{C}^\times$ defined for each $a \in \mathbb{F}_q$ by taking $e_q(a) = e(\operatorname{tr}(a)/p)$. This character induces a map $e : \mathbb{K}_\infty \to \mathbb{C}^\times$ by defining, for each element $\alpha \in \mathbb{K}_\infty$, the value of $e(\alpha)$ to be $e_q(\operatorname{res}\alpha)$. The orthogonality relation underlying the Fourier analysis of $\mathbb{F}_q[t]$, established in [1, Lemma 1], takes the shape

$$\int_{\mathbb{T}} e(h\alpha)\, d\alpha = \begin{cases} 1 & \text{when } h = 0, \\ 0 & \text{when } h \in \mathbb{F}_q[t] \setminus \{0\}. \end{cases}$$

Therefore, for $(h_1, \dots, h_R) \in \mathbb{F}_q[t]^R$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_R) \in \mathbb{K}_\infty^R$, we have

$$\int_{\mathbb{T}^R} e(h_1\alpha_1 + \dots + h_R\alpha_R)\, d\boldsymbol{\alpha} = \prod_{i=1}^{R} \int_{\mathbb{T}} e(h_i\alpha_i)\, d\alpha_i = \begin{cases} 1 & \text{when } h_j = 0 \ (1 \le j \le R), \\ 0 & \text{otherwise}. \end{cases} \tag{1.2}$$

2. Proof of Theorem 1.1. For $R, S \in \mathbb{N}$ with $S \ge 2R + 1$, let $Y = (a_{i,j}) \in \mathbb{F}_q^{R\times S}$ satisfy Conditions 1 and 2. For $N \in \mathbb{N}$, let $D_Y(\mathcal{S}_N)$ be defined as in Section 1. Write $d_Y(N) = D_Y(\mathcal{S}_N)/q^N$. For convenience, in what follows, we will write $D(\mathcal{S}_N)$ in place of $D_Y(\mathcal{S}_N)$ and $d(N)$ in place of $d_Y(N)$. Hence, to prove Theorem 1.1, it is equivalent to show that $d(N) \le (C/N)^{(L-R+1)/R}$.

For a set $A \subseteq \mathcal{S}_N$, let $T(A) = T_Y(A)$ denote the number of solutions of (1.1) with $x_i \in A$ $(1 \le i \le S)$. Let 1_A be the characteristic function of A, i.e., $1_A(x) = 1$ if $x \in A$ and $1_A(x) = 0$ otherwise. For $1 \le j \le S$ and $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_R) \in \mathbb{K}_\infty^R$, define

$$F_j(\boldsymbol{\alpha}) = \sum_{x\in A} e((a_{1,j}\alpha_1 + \dots + a_{R,j}\alpha_R)x).$$

By (1.2), we see that

$$T(A) = \int_{\mathbb{T}^R} F_1 \cdots F_S(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}.$$

We will estimate $T(A)$ by dividing $\mathbb{T}^R$ into two parts: the *major arc* $\mathfrak{M}$ defined by

$$\mathfrak{M} = \{(\alpha_1, \dots, \alpha_R) \in \mathbb{K}_\infty^R \mid \operatorname{ord} \alpha_i < -N \ (1 \le i \le R)\}$$

and the *minor arc* $\mathfrak{m} = \mathbb{T}^R \setminus \mathfrak{M}$. We have

$$T(A) = \int_{\mathfrak{M}} F_1 \cdots F_S(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} + \int_{\mathfrak{m}} F_1 \cdots F_S(\boldsymbol{\alpha})\, d\boldsymbol{\alpha}. \tag{2.1}$$

Before proving Theorem 1.1, we will need to obtain bounds on $T(A)$ and the contributions of the major and minor arcs.

Lemma 2.1. *Suppose that $Y \in \mathbb{F}_q^{R\times S}$ satisfies Condition 2. Suppose also that $A \subseteq \mathcal{S}_N$ is a set for which the equations in (1.1) are never satisfied simultaneously by distinct elements $x_1, \dots, x_S \in A$. Then*

$$T(A) \le C_1 |A|^{S-R-1},$$

where $C_1 = C_1(Y) = \binom{S}{2}$.

Proof. We have

$$T(A) = |\{\mathbf{x} \in A^S \mid Y\mathbf{x} = \mathbf{0}\}|.$$

Since $A \subseteq \mathcal{S}_N$ is such that the equations in (1.1) are never satisfied simultaneously by distinct elements $x_1, \dots, x_S \in A$, whenever $Y\mathbf{x} = \mathbf{0}$ for some $\mathbf{x} \in A^S$, there exist distinct elements $i, j \in \{1, \dots, S\}$ with $x_i = x_j$. Fix one of the C_1 choices of $\{i, j\}$. Let Y_1 be the matrix obtained from Y by deleting columns i, j. We consider two cases.

Case 1: $\{i, j\} \cap \{1, \dots, L\} = \emptyset$. We denote by $\operatorname{rk} Y_1$ the rank of the matrix Y_1. By Condition 2, we have $\operatorname{rk} Y_1 = R$. It follows that

$$|\{\mathbf{x} \in A^S \mid x_i = x_j \text{ and } Y\mathbf{x} = \mathbf{0}\}| \le |A|^{S-R-1}.$$

Case 2: $\{i, j\} \cap \{1, \dots, L\} \ne \emptyset$. Without loss of generality, we may assume that $i \in \{1, \dots, L\}$. By Condition 2, we can find two disjoint subsets I_1 and I_2 of $\{1, \dots, S\} \setminus \{i\}$, each with cardinality R, such that the columns of Y indexed by either set are linearly independent. Since $I_1 \cap I_2 = \emptyset$, we may assume that $j \notin I_1$. Then $\{i, j\} \cap I_1 = \emptyset$. Hence, $\operatorname{rk} Y_1 = R$, which implies that

$$|\{\mathbf{x} \in A^S \mid x_i = x_j \text{ and } Y\mathbf{x} = \mathbf{0}\}| \le |A|^{S-R-1}.$$

On recalling the definition of C_1 and combining Cases 1 and 2, the lemma follows. ∎

Lemma 2.2. *Suppose that $Y \in \mathbb{F}_q^{R\times S}$ and $A \subseteq \mathcal{S}_N$. Then*

$$\int_{\mathfrak{M}} F_1 \cdots F_S(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = q^{-NR} |A|^S.$$

Proof. For $1 \le j \le S$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_R) \in \mathfrak{M}$, and $x \in A \subseteq \mathcal{S}_N$, we have

$$\operatorname{ord}((a_{1,j}\alpha_1 + \cdots + a_{R,j}\alpha_R)x) \le -1 + N + \max_{1\le i\le R} \operatorname{ord}\alpha_i \le -2.$$

Thus,

$$F_j(\boldsymbol{\alpha}) = \sum_{x\in A} e((a_{1,j}\alpha_1 + \cdots + a_{R,j}\alpha_R)x) = \sum_{x\in A} 1 = |A|.$$

Therefore, our major arc contribution is

$$\int_{\mathfrak{M}} F_1 \cdots F_S(\boldsymbol{\alpha})\, d\boldsymbol{\alpha} = \operatorname{mes}(\mathfrak{M})|A|^S = q^{-NR}|A|^S. \qquad \blacksquare$$

LEMMA 2.3. *For $Y \in \mathbb{F}_q^{R\times S}$ and $A \subseteq \mathcal{S}_N$, suppose that the columns of Y indexed by $k_1, \ldots, k_R$ are linearly independent. Then*

$$\int_{\mathbb{T}^R} |F_{k_1} \cdots F_{k_R}(\boldsymbol{\alpha})|^2\, d\boldsymbol{\alpha} = |A|^R.$$

Proof. Let Z denote the matrix $(a_{i,k_j})_{1\le i,j\le R} \in \mathbb{F}_q^{R\times R}$. By (1.2), we have

$$\int_{\mathbb{T}^R} |F_{k_1} \cdots F_{k_R}(\boldsymbol{\alpha})|^2\, d\boldsymbol{\alpha} = |\{(\mathbf{x}, \mathbf{y}) \in A^R \times A^R \mid Z\mathbf{x} = Z\mathbf{y}\}|.$$

Since $\det Z \ne 0$, $Z\mathbf{x} = Z\mathbf{y}$ if and only if $\mathbf{x} = \mathbf{y}$. Thus,

$$\int_{\mathbb{T}^R} |F_{k_1} \cdots F_{k_R}(\boldsymbol{\alpha})|^2\, d\boldsymbol{\alpha} = |\{(\mathbf{x}, \mathbf{y}) \in A^R \times A^R \mid \mathbf{x} = \mathbf{y}\}| = |A|^R. \qquad \blacksquare$$

LEMMA 2.4. *Suppose that $Y \in \mathbb{F}_q^{R\times S}$ satisfies Condition* 1. *Suppose also that $A \subseteq \mathcal{S}_N$ is a set for which the equations in* (1.1) *are never satisfied simultaneously by distinct elements $x_1, \ldots, x_S \in A$. Then*

$$\sup_{-N\le \operatorname{ord}\beta<0} \Big|\sum_{x\in A} e(\beta x)\Big| \le d(N-1)q^N - |A|.$$

Proof. For $-N \le \operatorname{ord}\beta < 0$, let $W = W(\beta) = \{y \in \mathcal{S}_N : \operatorname{res}(\beta y) = 0\}$. Since $-N \le \operatorname{ord}\beta < 0$, we can write $\operatorname{ord}\beta = -l$ and $\beta = \sum_{j\le -l} b_j t^j$ with $-N \le -l \le -1$, $b_j \in \mathbb{F}_q$ $(j \le -l)$, and $b_{-l} \ne 0$. Then the polynomial $y = c_{N-1}t^{N-1} + \cdots + c_0 \in \mathcal{S}_N$ is in W if and only if

$$\operatorname{res}(\beta y) = b_{-l}c_{l-1} + b_{-l-1}c_l + \cdots + b_{-N}c_{N-1} = 0.$$

Hence, $W \simeq \mathbb{F}_q^{N-1}$ as a vector space over $\mathbb{F}_q$.

Since $-N \le \operatorname{ord}\beta < 0$, by [1, Lemma 7], we have

$$\sum_{\operatorname{ord} x<N} e(\beta x) = 0.$$

Therefore,

$$|W|\Big|\sum_{x\in A} e(\beta x)\Big| = \Big|\sum_{y\in W}\sum_{\operatorname{ord} x<N} d(N-1)e(\beta x) - \sum_{y\in W}\sum_{\operatorname{ord} x<N} 1_A(x)e(\beta x)\Big|.$$

For $y \in W$, since $e(\beta y) = 1$ and $y \in \mathcal{S}_N$, we deduce by a change of variables that

$$\sum_{\operatorname{ord} x < N} 1_A(x) e(\beta x) = \sum_{\operatorname{ord} x < N} 1_A(x) e(\beta(x+y)) = \sum_{\operatorname{ord} x < N} 1_A(x-y) e(\beta x).$$

It follows that

$$\begin{aligned} |W| \Big| \sum_{x \in A} e(\beta x) \Big| &= \Big| \sum_{\operatorname{ord} x < N} \Big(\sum_{y \in W} d(N-1) - \sum_{y \in W} 1_A(x-y) \Big) e(\beta x) \Big| \\ &\le \sum_{\operatorname{ord} x < N} \Big| \sum_{y \in W} d(N-1) - \sum_{y \in W} 1_A(x-y) \Big| \\ &= \sum_{\operatorname{ord} x < N} \big| d(N-1)|W| - |W \cap (x-A)| \big|. \end{aligned}$$

Since $a_{i,1} + \cdots + a_{i,S} = 0$ $(1 \le i \le R)$ and the equations in (1.1) are never satisfied simultaneously by distinct $x_1, \ldots, x_S \in A$, the equations in (1.1) are never satisfied simultaneously by distinct $x_1, \ldots, x_S \in W \cap (x-A)$. Since $W \simeq \mathcal{S}_{N-1}$ as a vector space over $\mathbb{F}_q$ and $Y \in \mathbb{F}_q^{R\times S}$, any invertible $\mathbb{F}_q$-linear transformation from W to $\mathcal{S}_{N-1}$ maps $W \cap (x-A)$ to a subset of $\mathcal{S}_{N-1}$ for which the equations in (1.1) are never satisfied simultaneously by distinct elements of the subset. This implies that $|W \cap (x-A)| \le d(N-1)|W|$. Therefore

$$\begin{aligned} |W| \Big| \sum_{x \in A} e(\beta x) \Big| &\le \sum_{\operatorname{ord} x < N} (d(N-1)|W| - |W \cap (x-A)|) \\ &= d(N-1)|W| q^N - |W|\,|A|. \end{aligned}$$

Thus, if $-N \le \operatorname{ord} \beta < 0$, we have

$$\Big| \sum_{x \in A} e(\beta x) \Big| \le d(N-1) q^N - |A|. \ \blacksquare$$

Lemma 2.5. *Suppose that* $Y \in \mathbb{F}_q^{R\times S}$ *satisfies Condition* 2. *Let*

$$Q = Q(Y) = \{B \subseteq \{1, \ldots, L\} \mid |B| = L - R + 1\}.$$

For $B \in Q$, *let*

$$\mathfrak{m}_B = \Big\{ \boldsymbol{\alpha} \in \mathbb{T}^R \ \Big| \ \operatorname{ord}\Big(\sum_{i=1}^{R} a_{i,k} \alpha_i \Big) \ge -N \ (k \in B) \Big\}.$$

Then

$$\mathfrak{m} \subseteq \bigcup_{B \in Q} \mathfrak{m}_B.$$

Proof. Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_R) \in \mathfrak{m}$. Select any R columns $k_1, \ldots, k_R$ from the first L columns of Y, and denote by $X = (a_{i,k_j})_{1 \le i,j \le R} \in \mathbb{F}_q^{R\times R}$ the matrix formed by these columns. By Condition 2, we have $\det X \neq 0$. Write $\alpha_i = \sum_{m\le -1} b_{i,m}t^m$ $(1 \le i \le R)$ with $b_{i,m} \in \mathbb{F}_q$ $(1 \le i \le R, m \le -1)$. Thus,

$$\sum_{i=1}^{R} a_{i,k_j}\alpha_i = \sum_{m\le -1}\sum_{i=1}^{R} a_{i,k_j} b_{i,m} t^m \qquad (1 \le j \le R).$$

Suppose for the moment that for all $1 \le j \le R$, we have $\operatorname{ord}(\sum_{i=1}^{R} a_{i,k_j}\alpha_i) < -N$. It follows that

$$\sum_{i=1}^{R} a_{i,k_j} b_{i,m} = 0 \qquad (-N \le m \le -1,\ 1 \le j \le R). \tag{2.2}$$

Write $\mathbf{b}_m = (b_{1,m}, \dots, b_{R,m})$. Then (2.2) is equivalent to having $\mathbf{b}_m X = \mathbf{0}$ $(-N \le m \le -1)$. Since $\det X \neq 0$, we have $\mathbf{b}_m = \mathbf{0}$ $(-N \le m \le -1)$. Thus, $\alpha_i = \sum_{m<-N} b_{i,m}t^m$ $(1 \le i \le R)$, contradicting the fact that $\boldsymbol{\alpha} \in \mathfrak{m}$. Thus, $\operatorname{ord}(\sum_{i=1}^{R} a_{i,k_j}\alpha_i) \ge -N$ for at least one $1 \le j \le R$.

Since we can find an element k such that $\operatorname{ord}(\sum_{i=1}^{R} a_{i,k}\alpha_i) \ge -N$ amongst any R-element subset of $\{1, \dots, L\}$, it follows that there are at least $L-R+1$ values $k \in \{1, \dots, L\}$ with $\operatorname{ord}(\sum_{i=1}^{R} a_{i,k}\alpha_i) \ge -N$. That is, there exists $B \subseteq \{1, \dots, L\}$ with $|B| = L-R+1$ such that $\boldsymbol{\alpha} \in \mathfrak{m}_B$. This completes the proof of the lemma. ∎

Lemma 2.6. *Suppose that* $Y \in \mathbb{F}_q^{R\times S}$ *satisfies Conditions* 1 *and* 2. *Suppose also that* $A \subseteq \mathcal{S}_N$ *is such that the equations in* (1.1) *are never satisfied simultaneously by distinct elements* $x_1, \dots, x_S \in A$ *and* $|A| = d(N)q^N$. *Then*

$$\int_{\mathfrak{m}} |F_1 \cdots F_S(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha} \le C_2 (d(N-1) - d(N))^{L-R+1} d(N)^{S-L-1} q^{N(S-R)},$$

where $C_2 = C_2(Y) = \binom{L}{L-R+1}$.

Proof. Let $Q = Q(Y)$ and $\mathfrak{m}_B$ $(B \in Q)$ be as in Lemma 2.5. We have

$$\int_{\mathfrak{m}_B} |F_1 \cdots F_S(\boldsymbol{\alpha})|\, d\boldsymbol{\alpha} \le \Big(\sup_{\boldsymbol{\alpha}\in\mathfrak{m}_B} \prod_{j\in B} |F_j(\boldsymbol{\alpha})|\Big) \int_{\mathbb{T}^R} \Big|\prod_{j\notin B} F_j(\boldsymbol{\alpha})\Big|\, d\boldsymbol{\alpha}.$$

By Condition 2, there are two disjoint R-element subsets U and V of $\{1, \dots, S\} \setminus B$ such that the columns of Y indexed by either set are linearly independent. By Lemma 2.3 and the Cauchy–Schwarz inequality,

$$\begin{aligned}\int_{\mathbb{T}^R} \Big|\prod_{j\notin B} F_j(\boldsymbol{\alpha})\Big|\, d\boldsymbol{\alpha} &\le |A|^{S-|B|-2R} \int_{\mathbb{T}^R} \Big|\prod_{j\in U} F_j(\boldsymbol{\alpha})\Big| \Big|\prod_{j\in V} F_j(\boldsymbol{\alpha})\Big|\, d\boldsymbol{\alpha} \\ &\le |A|^{S-|B|-2R} \Big(\int_{\mathbb{T}^R} \Big|\prod_{j\in U} F_j(\boldsymbol{\alpha})\Big|^2 d\boldsymbol{\alpha}\Big)^{1/2} \Big(\int_{\mathbb{T}^R} \Big|\prod_{j\in V} F_j(\boldsymbol{\alpha})\Big|^2 d\boldsymbol{\alpha}\Big)^{1/2} \\ &= |A|^{S-|B|-2R} |A|^R = |A|^{S-|B|-R}.\end{aligned}$$

By Lemma 2.4, we see that for $j \in B$,

$$\sup_{\boldsymbol{\alpha} \in \mathfrak{m}_B} |F_j(\boldsymbol{\alpha})| \le (d(N-1) - d(N))q^N.$$

Thus,

$$\int_{\mathfrak{m}_B} |F_1 \cdots F_S(\boldsymbol{\alpha})| \, d\boldsymbol{\alpha} \le (d(N-1) - d(N))^{L-R+1} d(N)^{S-L-1} q^{N(S-R)}.$$

We have seen in Lemma 2.5 that $\mathfrak{m} \subseteq \bigcup_{B \in Q} \mathfrak{m}_B$. Since $|Q| = \binom{L}{L-R+1} = C_2$, we can deduce from the above inequality that

$$\int_{\mathfrak{m}} |F_1 \cdots F_S(\boldsymbol{\alpha})| \, d\boldsymbol{\alpha} \le C_2 (d(N-1) - d(N))^{L-R+1} d(N)^{S-L-1} q^{N(S-R)}.$$

This completes the proof of the lemma. ∎

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Suppose that $A \subseteq \mathcal{S}_N$ is a set for which the equations in (1.1) are never satisfied simultaneously by distinct $x_1, \ldots, x_S \in A$ and $|A| = d(N)q^N$. By (2.1), we have

$$\Big| \int_{\mathfrak{M}} F_1 \cdots F_S(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \Big| - \Big| \int_{\mathfrak{m}} F_1 \cdots F_S(\boldsymbol{\alpha}) \, d\boldsymbol{\alpha} \Big| \le T(A).$$

On applying Lemmas 2.1, 2.2, and 2.6, there exist positive constants C_1 and C_2 such that

$$d(N)^S q^{N(S-R)} - C_2 (d(N-1) - d(N))^{L-R+1} d(N)^{S-L-1} q^{N(S-R)} \le C_1 d(N)^{S-R-1} q^{N(S-R-1)}.$$

Thus,

$$d(N)^S - C_1 d(N)^{S-R-1} q^{-N} - C_2 (d(N-1) - d(N))^{L-R+1} d(N)^{S-L-1} \le 0. \tag{2.3}$$

Let

$$C = \max\{(2C_1)^{R/((R+1)(L-R+1))} \sup_{N \in \mathbb{N}} (N q^{-NR/((R+1)(L-R+1))}), \\ (2C_2)^{1/(L-R+1)} 2^{(L+1)/R} (L-R+1)/R, \ 1\}.$$

We now claim that for all $N \in \mathbb{N}$, one has

$$d(N) \le \left(\frac{C}{N}\right)^{(L-R+1)/R}. \tag{2.4}$$

This statement will follow by induction. Since $d(N) \le 1$, (2.4) holds trivially when $N = 1$. Let $N > 1$, and assume that

$$d(N-1) \le \left(\frac{C}{N-1}\right)^{(L-R+1)/R}.$$

We consider two cases.

CASE 1: $d(N)^S - C_1 d(N)^{S-R-1} q^{-N} \le \frac{1}{2} d(N)^S$. Then

$$d(N) \le (2C_1)^{1/(R+1)} q^{-N/(R+1)}.$$

Since

$$C \ge (2C_1)^{R/((R+1)(L-R+1))} (N q^{-NR/((R+1)(L-R+1))}),$$

we obtain

$$d(N) \le (C/N)^{(L-R+1)/R}.$$

CASE 2: $d(N)^S - C_1 d(N)^{S-R-1} q^{-N} > \frac{1}{2} d(N)^S$. We may deduce from (2.3) that

$$d(N)^{L+1} < 2C_2 (d(N-1) - d(N))^{L-R+1}.$$

By setting $C_3 = (2C_2)^{-1/(L-R+1)}$, we have

$$C_3 d(N)^{(L+1)/(L-R+1)} + d(N) < d(N-1). \tag{2.5}$$

Let $f(x) = (C/x)^{(L-R+1)/R}$. By the mean value theorem, there exists $\theta_N \in [0,1]$ such that

$$\begin{aligned} f(N-1) - f(N) &= f'(N-\theta_N)(-1) \\ &= C^{(L-R+1)/R} (L-R+1) R^{-1} (N-\theta_N)^{-(L+1)/R}. \end{aligned}$$

Since $C \ge C_3^{-1} 2^{(L+1)/R} (L-R+1)/R$, it follows that

$$\begin{aligned} f(N-1) - f(N) &\le C^{(L-R+1)/R} (L-R+1) R^{-1} (N-1)^{-(L+1)/R} \\ &= C^{(L+1)/R} C^{-1} (L-R+1) R^{-1} (N-1)^{-(L+1)/R} \\ &\le C^{(L+1)/R} C_3 2^{-(L+1)/R} (N-1)^{-(L+1)/R} \\ &\le C_3 C^{(L+1)/R} N^{-(L+1)/R}. \end{aligned} \tag{2.6}$$

From the induction hypothesis and (2.6), we obtain

$$\begin{aligned} d(N-1) \le f(N-1) &\le C_3 (C/N)^{\frac{L+1}{R}} + f(N) \\ &= C_3 (C/N)^{\frac{L-R+1}{R} \cdot \frac{L+1}{L-R+1}} + (C/N)^{\frac{L-R+1}{R}}. \end{aligned}$$

On recalling (2.5), we have

$$\begin{aligned} C_3 d(N)^{\frac{L+1}{L-R+1}} + d(N) &< d(N-1) \\ &\le C_3 (C/N)^{\frac{L-R+1}{R} \cdot \frac{L+1}{L-R+1}} + (C/N)^{\frac{L-R+1}{R}}. \end{aligned}$$

Since $C_3 x^{\frac{L+1}{L-R+1}} + x$ is an increasing function in x, we have

$$d(N) \le (C/N)^{(L-R+1)/R}.$$

On combining Cases 1 and 2, the inequality (2.4) follows. This completes the proof of Theorem 1.1. ■

Acknowledgements. The research of the first author is supported in part by an NSERC discovery grant. The research of the second author is supported in part by NSF grant DMS-0635607 and NSA Young Investigators Grant H98230-10-1-0155. The authors would like to thank the referee for carefully reading the paper and making numerous valuable suggestions.

References

[1] R. M. Kubota, *Waring's problem for* $\mathbb{F}_q[x]$, Dissertationes Math. (Rozprawy Mat.) 117 (1974).
[2] Y.-R. Liu and C. V. Spencer, *A generalization of Roth's theorem in function fields*, Int. J. Number Theory 5 (2009), 1149–1154.
[3] M. Rosen, *Number Theory in Function Fields*, Springer, New York, 2002.
[4] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. 28 (1953), 104–109.
[5] —, *On certain sets of integers (II)*, ibid. 29 (1954), 20–26.

Yu-Ru Liu, Xiaomei Zhao
Department of Pure Mathematics
Faculty of Mathematics
University of Waterloo
Waterloo, Ontario, Canada N2L 3G1
E-mail: yrliu@math.uwaterloo.ca
x8zhao@math.uwaterloo.ca

Craig V. Spencer
School of Mathematics
Institute for Advanced Study
1 Einstein Drive
Princeton, NJ 08540, U.S.A.
and
Department of Mathematics
Kansas State University
138 Cardwell Hall
Manhattan, KS 66506, U.S.A.
E-mail: craigvspencer@gmail.com

Received on 16.7.2009
and in revised form on 1.11.2009 (6088)