Bull. Aust. Math. Soc. **79** (2009), 259–271
doi:10.1017/S0004972708001184

ON THE GOWERS NORM OF PSEUDORANDOM BINARY SEQUENCES

HARALD NIEDERREITER and JOËL RIVAT✉

(Received 16 June 2008)

Abstract

We study the Gowers norm for periodic binary sequences and relate it to correlation measures for such sequences. The case of periodic binary sequences derived from inversive pseudorandom numbers is considered in detail.

2000 *Mathematics subject classification*: primary 11K45, 11L40; secondary 65C10.
Keywords and phrases: Gowers norm, pseudorandom, binary sequence, additive character, correlation.

1. Introduction

Considerable interest in pseudorandom binary sequences was stimulated by the paper of Mauduit and Sárközy [4] which introduced several new measures of pseudorandomness for binary sequences. The present paper is inspired by [4] as well as by the work of Gowers [2] on combinatorial and additive number theory. In fact, as can be seen from [2, Section 2], Gowers already had in mind some notion of pseudorandomness for subsets of $\mathbb{Z}/N\mathbb{Z}$. We consider the essentially equivalent case of periodic binary sequences with period N and we use what is now called the Gowers norm (see Section 2 below for its definition) as a measure of pseudorandomness for such sequences.

The link between the Gowers norm and the work of Mauduit and Sárközy [4] is established via Theorem 6 which bounds the Gowers norm in terms of a suitable correlation measure in the spirit of [4]. The proof of Theorem 6 leads to interesting combinatorial problems involving polynomials over the binary field.

To provide an example of how the Gowers norm can be treated for specific periodic binary sequences, we consider the case of inversive sequences, that is, of binary sequences derived from inversive pseudorandom numbers. Concretely, we analyse the inversive generator that was recently introduced by the authors in [7] since it has advantages over the classical inversive generator.

2. Gowers norm and correlation

Let $d \geq 1$ and $N \geq 2$ be integers and identify $\mathbb{Z}_N = \mathbb{Z}/N\mathbb{Z}$ with $\{0, 1, \ldots, N-1\}$. For $f : \mathbb{Z}_N \to \{-1, +1\}$ we define

$$G_d(f) := \sum_{x_1,\ldots,x_d \in \mathbb{Z}_N} \sum_{y \in \mathbb{Z}_N} \prod_{S \subseteq \{1,\ldots,d\}} f\Bigg(y + \sum_{i \in S} x_i\Bigg).$$

Note that $|G_d(f)|^{1/2^d}$ is the norm $\|f\|_d$ introduced by Gowers in [2, Lemma 3.9]. We observe that we can equivalently view f as a periodic binary sequence with period N.

We can write $G_d(f)$ as

$$G_d(f) = \sum_{x_1,\ldots,x_d \in \mathbb{Z}_N} \sum_{y \in \mathbb{Z}_N} f(y) \Bigg(\prod_{1 \leq i_1 \leq d} f(y + x_{i_1})\Bigg)\Bigg(\prod_{1 \leq i_1 < i_2 \leq d} f(y + x_{i_1} + x_{i_2})\Bigg) \cdots \Bigg(\prod_{1 \leq i_1 < \cdots < i_d \leq d} f(y + x_{i_1} + \cdots + x_{i_d})\Bigg). \tag{1}$$

For fixed $x_1, \ldots, x_d \in \mathbb{Z}_N$, we define

$$t_1 = 0, \quad t_2 = x_1, \ldots, t_{d+1} = x_d, \quad t_{d+2} = x_1 + x_2, \ldots, \quad t_{2^d} = x_1 + \cdots + x_d,$$

where all these identities are viewed in $\mathbb{Z}_N$, that is, they are in fact congruences mod N. Then

$$\begin{aligned} &\sum_{y \in \mathbb{Z}_N} f(y) \Bigg(\prod_{1 \leq i_1 \leq d} f(y + x_{i_1})\Bigg)\Bigg(\prod_{1 \leq i_1 < i_2 \leq d} f(y + x_{i_1} + x_{i_2})\Bigg) \\ &\qquad \cdots \Bigg(\prod_{1 \leq i_1 < \cdots < i_d \leq d} f(y + x_{i_1} + \cdots + x_{i_d})\Bigg) \\ &\quad = \sum_{y \in \mathbb{Z}_N} f(y + t_1) \cdots f(y + t_{2^d}). \end{aligned} \tag{2}$$

With $x_1, \ldots, x_d$ still fixed, we partition the t_j's according to their values modulo N and we put for $0 \leq i \leq N-1$,

$$m_i := \#\{j \in \mathbb{Z} \mid 1 \leq j \leq 2^d, t_j \equiv i \pmod N\},$$

so that $\sum_{i=0}^{N-1} m_i = 2^d$. Then

$$\forall y \in \mathbb{Z}_N, \quad f(y + t_1) \cdots f(y + t_{2^d}) = \prod_{i=0}^{N-1} f(y + i)^{m_i}.$$

Let $k = k(x_1, \ldots, x_d)$ be the number of i, $0 \leq i \leq N-1$, such that m_i is odd, and let $0 \leq d_1 < d_2 < \cdots < d_k \leq N-1$ be the values of i such that m_i is odd. Note that $0 \leq k \leq \min(N, 2^d)$. Then

$$\forall y \in \mathbb{Z}_N, \quad f(y + t_1) \cdots f(y + t_{2^d}) = f(y + d_1) \cdots f(y + d_k),$$

with an empty product on the right-hand side being interpreted as 1.

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

Now we introduce a correlation measure for the periodic binary sequence $e_n = f(n)$, $n \in \mathbb{Z}_N$, with period N. For $k = 0$ we define $P_0(f) := N$ and for $1 \le k \le N$ we define

$$P_k(f) := \max_D \left| \sum_{y \in \mathbb{Z}_N} f(y + d_1) \cdots f(y + d_k) \right|,$$

where the maximum is over all $D = (d_1, \ldots, d_k) \in \mathbb{Z}^k$ with $0 \le d_1 < d_2 < \cdots < d_k \le N - 1$.

With the notation above we have then

$$\left| \sum_{y \in \mathbb{Z}_N} f(y + t_1) \cdots f(y + t_{2^d}) \right| \le P_k(f)$$

whenever $k(x_1, \ldots, x_d) = k$. In view of (1) and (2), this yields

$$|G_d(f)| \le \sum_{k=0}^{\min(N, 2^d)} P_k(f) \cdot \#\{(x_1, \ldots, x_d) \in \mathbb{Z}_N^d \mid k(x_1, \ldots, x_d) = k\}.$$

For $0 \le k \le \min(N, 2^d)$, we define

$$B_d(k, N) := \#\{(x_1, \ldots, x_d) \in \mathbb{Z}_N^d \mid k(x_1, \ldots, x_d) = k\}.$$

Then

$$|G_d(f)| \le B_d(0, N)N + \sum_{k=1}^{\min(N, 2^d)} B_d(k, N) P_k(f). \tag{3}$$

Now we define the correlation measure $M_d(f)$ by

$$M_d(f) := \max_{1 \le k \le \min(N, 2^d)} P_k(f). \tag{4}$$

Then

$$\begin{aligned} |G_d(f)| &\le B_d(0, N)N + M_d(f) \sum_{k=1}^{\min(N, 2^d)} B_d(k, N) \\ &= B_d(0, N)N + (N^d - B_d(0, N))M_d(f). \end{aligned} \tag{5}$$

The numbers $B_d(k, N)$ can be described in the following equivalent manner. We write $\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$ for the finite field of order 2.

LEMMA 1. *For $0 \le k \le \min(N, 2^d)$, $B_d(k, N)$ is equal to the number of $(x_1, \ldots, x_d) \in \{0, 1, \ldots, N - 1\}^d$ such that in the polynomial ring $\mathbb{F}_2[z]$*

$$(z^{x_1} - 1) \cdots (z^{x_d} - 1) \equiv k\text{-nomial (of degree } < N) \pmod{(z^N - 1)}.$$

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

PROOF. For given $(x_1, \ldots, x_d) \in \{0, 1, \ldots, N-1\}^d$,

$$(z^{x_1}+1)\cdots(z^{x_d}+1) = 1 + z^{x_1} + \cdots + z^{x_d} + z^{x_1+x_2} + \cdots + z^{x_1+\cdots+x_d} = \sum_{j=1}^{2^d} z^{t_j}.$$

Consider this polynomial identity modulo $z^N - 1$. Then by the definition of the m_i,

$$(z^{x_1}+1)\cdots(z^{x_d}+1) \equiv \sum_{i=0}^{N-1} m_i z^i \pmod{(z^N-1)}.$$

Now we consider this congruence also modulo 2, that is, we work in the residue class ring $\mathbb{F}_2[z]/(z^N-1)$. Then

$$(z^{x_1}-1)\cdots(z^{x_d}-1) \equiv k\text{-nomial (of degree } < N) \pmod{(z^N-1)}$$

by the definition of k. This yields the desired result. □

Now we study $B_d(0, N)$ and write $B_d(N) := B_d(0, N)$. The determination of the counting function $B_d(N)$ seems to be a nontrivial combinatorial problem. The following lemma is an easy consequence of Lemma 1.

LEMMA 2. *For any $d \geq 1$ and $N \geq 2$,*

$$B_d(N) = N^d - (N-1)^d + \#\{(x_1, \ldots, x_d) \in \{1, \ldots, N-1\}^d : z^N - 1 \text{ divides } (z^{x_1}-1)\cdots(z^{x_d}-1) \text{ in } \mathbb{F}_2[z]\}.$$

PROOF. There are $N^d - (N-1)^d$ choices of $(x_1, \ldots, x_d) \in \{0, 1, \ldots, N-1\}^d$ with at least one $x_j = 0$, and for each of these it is trivial that $z^N - 1$ divides $(z^{x_1}-1)\cdots(z^{x_d}-1)$ in $\mathbb{F}_2[z]$. The rest follows from Lemma 1. □

Any positive integer x can be written in the form $x = 2^e y$ with an integer $e \geq 0$ and an odd integer y. We put

$$\rho(x) := 2^e, \quad \lambda(x) := y.$$

LEMMA 3. *Let $(x_1, \ldots, x_d) \in \{1, \ldots, N-1\}^d$. Then $z^N - 1$ divides $(z^{x_1}-1)\cdots(z^{x_d}-1)$ in $\mathbb{F}_2[z]$ if and only if*

$$\sum_{\substack{j=1 \\ \lambda(N)\mid\lambda(x_j)}}^{d} \rho(x_j) \geq \rho(N).$$

PROOF. Write $N = 2^e n$ with integers $e \geq 0$ and n odd. Note that $z^N - 1 = (z^n - 1)^{2^e}$ in $\mathbb{F}_2[z]$. For given $(x_1, \ldots, x_d) \in \{1, \ldots, N-1\}^d$, we can write $x_j = \rho(x_j)\lambda(x_j)$ for $1 \leq j \leq d$. Then $z^N - 1$ divides $(z^{x_1}-1)\cdots(z^{x_d}-1)$ in $\mathbb{F}_2[z]$ if and only if

$$(z^n-1)^{2^e} \text{ divides } (z^{\lambda(x_1)}-1)^{\rho(x_1)}\cdots(z^{\lambda(x_d)}-1)^{\rho(x_d)} \quad \text{in } \mathbb{F}_2[z]. \tag{6}$$

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

Note that each of the binomials $z^n - 1, z^{\lambda(x_1)} - 1, \ldots, z^{\lambda(x_d)} - 1$ has only simple roots in the algebraic closure $\overline{\mathbb{F}_2}$. Choose a primitive nth root of unity β in $\overline{\mathbb{F}_2}$. Then (6) holds if and only if each β^r, $r = 1, \ldots, n$, is a root of multiplicity at least 2^e of $(z^{\lambda(x_1)} - 1)^{\rho(x_1)} \cdots (z^{\lambda(x_d)} - 1)^{\rho(x_d)}$. Now β^r is a root of $z^{\lambda(x_j)} - 1$ if and only if $\beta^{r\lambda(x_j)} = 1$, that is, if and only if $n \mid r\lambda(x_j)$. Thus, it suffices to take $r = 1$. This leads to the condition in the lemma. □

From Lemmas 2 and 3 it follows that

$$B_d(N) = N^d - (N-1)^d + E_d(N), \tag{7}$$

where

$$E_d(N) := \#\left\{(x_1, \ldots, x_d) \in \{1, \ldots, N-1\}^d : \sum_{\substack{j=1 \\ \lambda(N)\mid\lambda(x_j)}}^{d} \rho(x_j) \geq \rho(N)\right\}. \tag{8}$$

Now we prove an upper bound on $E_d(N)$.

LEMMA 4. *For any $d \geq 1$ and $N \geq 2$,*

$$E_d(N) \leq c_d N^{d-2}$$

with a constant $c_d > 0$ depending only on d.

PROOF. We proceed by induction on d. For $d = 1$ we note that if $x_1 \in \{1, \ldots, N-1\}$ is counted by $E_1(N)$, then $\lambda(N) \mid \lambda(x_1)$ and $\rho(x_1) \geq \rho(N)$, hence $x_1 = \rho(x_1)\lambda(x_1) \geq \rho(N)\lambda(N) = N$. Therefore $E_1(N) = 0$.

Assume that the lemma is shown for all dimensions $\leq d$, for some $d \geq 1$. Now consider the dimension $d+1$. As before, write $N = 2^e n$ with integers $e \geq 0$ and n odd. For $e = 0$ we have $\lambda(N) = N$, and so it is trivial by (8) that $E_d(N) = 0$. So we can assume that $e \geq 1$. Note that $(x_1, \ldots, x_{d+1}) \in \{1, \ldots, N-1\}^{d+1}$ is counted by $E_{d+1}(N)$ if and only if

$$\sum_{\substack{j=1 \\ n\mid\lambda(x_j)}}^{d+1} \rho(x_j) \geq 2^e. \tag{9}$$

For any j with $n \mid \lambda(x_j)$ we must have $\rho(x_j) \leq 2^{e-1}$, since otherwise $x_j = \rho(x_j)\lambda(x_j) \geq 2^e n = N$. In particular, if m is the number of j, $1 \leq j \leq d+1$, with $n \mid \lambda(x_j)$, then $2 \leq m \leq d+1$. Therefore

$$E_{d+1}(N) = \sum_{m=2}^{d+1} \sum_{\substack{J \subseteq \{1,\ldots,d+1\} \\ |J| = m}} E_{d+1}(N; J),$$

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

where

$$E_{d+1}(N;\, J) := \#\{(x_1, \ldots, x_{d+1}) \in \{1, \ldots, N-1\}^{d+1} : (x_1, \ldots, x_{d+1}) \text{ satisfies (9), } n \mid \lambda(x_j) \text{ for } j \in J, \text{ and } n \nmid \lambda(x_j) \text{ for } j \in \{1, \ldots, d+1\} \setminus J\}.$$

By symmetry, it suffices to consider $J = \{1, \ldots, m\}$. If we write

$$E_{d+1}(N; m) := E_{d+1}(N; \{1, \ldots, m\}),$$

then

$$E_{d+1}(N) = \sum_{m=2}^{d+1} \binom{d+1}{m} E_{d+1}(N; m). \tag{10}$$

Note that if $(x_1, \ldots, x_{d+1}) \in \{1, \ldots, N-1\}^{d+1}$ is counted by $E_{d+1}(N; m)$ for a fixed m, then $n \mid \lambda(x_j)$ for $1 \le j \le m$ and $n \nmid \lambda(x_j)$ for $m+1 \le j \le d+1$. Thus, the condition (9) becomes

$$\sum_{j=1}^{m} \rho(x_j) \ge 2^e. \tag{11}$$

Let j_0 with $1 \le j_0 \le m$ be such that

$$\rho(j_0) = \max(\rho(x_1), \ldots, \rho(x_m)).$$

Without loss of generality we assume that $j_0 = m$. Then the condition (11) is equivalent to $\sum_{j=1}^{m-1} \rho(x_j) \ge 2^e - \rho(x_m)$, and since $\rho(x_m) \le 2^{e-1}$, a weaker condition is $\sum_{j=1}^{m-1} \rho(x_j) \ge 2^{e-1}$. It follows therefore that

$$E_{d+1}(N; m) \le m F_m(N) N^{d+1-m}, \tag{12}$$

where

$$F_m(N) := \#\left\{(x_1, \ldots, x_m) \in \{1, \ldots, N-1\}^m : \sum_{j=1}^{m} \rho(x_j) \ge 2^e, \sum_{j=1}^{m-1} \rho(x_j) \ge 2^{e-1}, \right.$$
$$\left. n \mid \lambda(x_j) \text{ for } 1 \le j \le m, \text{ and } \rho(x_m) \ge \rho(x_j) \text{ for } 1 \le j \le m-1 \right\}$$

and the factor m on the right-hand side of (12) takes care of the fact that there are m choices for j_0.

We put

$$H_{m-1}(N) := \#\left\{(x_1, \ldots, x_{m-1}) \in \{1, \ldots, N-1\}^{m-1} : \sum_{j=1}^{m-1} \rho(x_j) \ge 2^{e-1}, n \mid \lambda(x_j) \text{ for } 1 \le j \le m-1 \right\}.$$

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

Note that in view of Lemmas 2 and 3, $B_{m-1}(N/2) = B_{m-1}(2^{e-1}n)$ can be written in the form

$$B_{m-1}\left(\frac{N}{2}\right) = \#\left\{(x_1, \ldots, x_{m-1}) \in \left\{1, \ldots, \frac{N}{2}\right\}^{m-1} : \sum_{\substack{j=1 \\ n|\lambda(x_j)}}^{m-1} \rho(x_j) \geq 2^{e-1}\right\}.$$

Furthermore, any $(x_1, \ldots, x_{m-1}) \in \{1, \ldots, N-1\}^{m-1}$ can be written in the form

$$\left(w_1 + \delta_1 \frac{N}{2}, \ldots, w_{m-1} + \delta_{m-1}\frac{N}{2}\right)$$

with $(w_1, \ldots, w_{m-1}) \in \{1, \ldots, N/2\}^{m-1}$ and $\delta_1, \ldots, \delta_{m-1} \in \{0, 1\}^{m-1}$. If $(x_1, \ldots, x_{m-1})$ is counted by $H_{m-1}(N)$, then $n \mid x_j$ for $1 \leq j \leq m-1$, hence $n \mid (x_j - \delta_j 2^{e-1}n) = w_j$ for $1 \leq j \leq m-1$, and so $n \mid \lambda(w_j)$ since n is odd. Moreover, $\rho(w_j) = \rho(x_j - \delta_j 2^{e-1}n) \geq \rho(x_j)$ since $\rho(x_j) \leq 2^{e-1}$ for $1 \leq j \leq m-1$. Hence $(w_1, \ldots, w_{m-1})$ is counted by $B_{m-1}(N/2)$, and so

$$H_{m-1}(N) \leq 2^{m-1} B_{m-1}\left(\frac{N}{2}\right).$$

By the induction hypothesis we have $E_{m-1}(N/2) \leq c_{m-1}(N/2)^{m-3}$, hence (7) implies that $B_{m-1}(N/2) \leq c_{m-1}^{(1)} N^{m-2}$, and so

$$H_{m-1}(N) \leq c_{m-1}^{(2)} N^{m-2} \tag{13}$$

with constants $c_{m-1}^{(1)} > 0$ and $c_{m-1}^{(2)} > 0$ depending only on $m-1$.

Now we consider $F_m(N)$. If $(x_1, \ldots, x_m)$ is counted by $F_m(N)$, then $(x_1, \ldots, x_{m-1})$ is counted by $H_{m-1}(N)$. We fix $(x_1, \ldots, x_{m-1})$ and consider the number of choices for x_m. We have $1 \leq x_m < N = 2^e n$, $n \mid \lambda(x_m)$, and $\rho(x_m) \geq (1/m)2^e$. The latter inequality implies that $\rho(x_m) \geq 2^r$ with r being the least integer such that $2^r \geq (1/m)2^e$. Then both 2^r and n divide x_m, and so $x_m = 2^r nk$ with an integer k satisfying $1 \leq k < 2^{e-r} \leq m$. This yields at most $m-1$ choices for x_m. Therefore $F_m(N) \leq (m-1)H_{m-1}(N)$, and so (10), (12), and (13) yield the desired bound on $E_{d+1}(N)$. □

The following result is a consequence of (7) and Lemma 4.

PROPOSITION 5. *For any $d \geq 1$ and $N \geq 2$,*

$$N^d - (N-1)^d \leq B_d(N) \leq N^d - (N-1)^d + c_d N^{d-2}$$

with a constant $c_d > 0$ depending only on d.

By combining (5) and Proposition 5, we obtain the following theorem which provides a bound on $G_d(f)$ in terms of the correlation measure $M_d(f)$ defined in (4).

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

THEOREM 6. *For any* $f : \mathbb{Z}_N \to \{-1, +1\}$ *and any* $d \geq 1$,

$$|G_d(f)| \leq (N-1)^d M_d(f) + dN^d + c_d N^{d-1}$$

with a constant $c_d > 0$ *depending only on* d.

3. Pseudorandom numbers defined by inversive methods

Let q be a prime power and let $\mathbb{F}_q$ be the finite field of order q. Let $\alpha, \beta \in \mathbb{F}_q$ be such that the quadratic polynomial $X^2 - \beta X - \alpha$ is primitive over $\mathbb{F}_q$. Then by [7, Lemma 1 and Theorem 1], the sequence $R_0, R_1, \ldots$ of rational functions over $\mathbb{F}_q$ defined by

$$R_0(X) = X, \quad R_i(X) = R_{i-1}(\alpha X^{-1} + \beta) \quad \text{for } i = 1, 2, \ldots$$

is purely periodic of least period $q+1$. Furthermore, by [7, (8)], for $1 \leq i \leq q$,

$$R_i(X) = \frac{(\beta - \varepsilon_i)X + \alpha}{X - \varepsilon_i},$$

where $\varepsilon_1, \ldots, \varepsilon_q$ are distinct elements of $\mathbb{F}_q$, with $\varepsilon_1 = 0$ and $\varepsilon_q = \beta$, such that for $i = 2, \ldots, q$,

$$\varepsilon_i = \frac{\alpha}{\varepsilon_{i-1} - \beta}.$$

We let ε_0 be an arbitrary element *not* in $\mathbb{F}_q$ and extend $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_q$ to a sequence with period $q+1$. As in [7, (4)], we consider the permutations of $\mathbb{F}_q$ defined by $\psi_0(\gamma) = \gamma$, and for $1 \leq i \leq q$,

$$\psi_i(\gamma) = \begin{cases} R_i(\gamma) & \text{if } \gamma \neq \varepsilon_i, \\ \beta - \varepsilon_i & \text{if } \gamma = \varepsilon_i. \end{cases} \tag{14}$$

We extend the definition of ψ_i to all $i \geq 0$ by periodicity: for all $i \geq 0$, $\psi_{i+q+1} = \psi_i$. Then we build from a seed $\gamma_0 \in \mathbb{F}_q$ a sequence $\gamma_0, \gamma_1, \ldots$ of elements of $\mathbb{F}_q$, purely periodic of least period $q+1$, by putting for $i \geq 0$,

$$\gamma_i = \psi_i(\gamma_0). \tag{15}$$

By [7, Lemma 2] the sequence $\gamma_0, \gamma_1, \ldots$ contains all elements of $\mathbb{F}_q$.

The advantage of this construction over the classical inversive generator (see [6, Section 8.2] and [9] for this generator) is that for $i, j \geq 0$,

$$\psi_i(\psi_j(\gamma)) = R_{i+j}(\gamma) \quad \text{for } \gamma \neq \varepsilon_j \text{ and } \psi_j(\gamma) \neq \varepsilon_i. \tag{16}$$

The price is a slightly more complicated algorithm to compute γ_i. This construction allowed us to prove strong distribution and correlation properties of the generated sequences [7]. Further properties of these sequences were derived in [8] and [11].

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

In this section our aim is to study the character sum

$$S := \sum_{n=0}^{q} \chi\left(\sum_{j=1}^{k} \mu_j \gamma_{n+d_j}\right), \tag{17}$$

where χ is a nontrivial additive character of $\mathbb{F}_q$, $\mu_1, \ldots, \mu_k \in \mathbb{F}_q$, and $0 \le d_1 < \cdots < d_k \le q$.

We need the following result which is obtained by combining special cases of [5, Theorem 2] and [10, Lemma 2] (see also [7, Lemma 3]).

LEMMA 7. *Let $G(X) = h(X)/g(X)$ be a nonzero rational function over $\mathbb{F}_q$ such that g is a product of distinct monic linear polynomials over $\mathbb{F}_q$ and $\deg(h) < \deg(g)$ or $\deg(h) = \deg(g) + 1$. Then for any nontrivial additive character χ of $\mathbb{F}_q$,*

$$\left| \sum_{\substack{\phi \in \mathbb{F}_q \\ g(\phi) \neq 0}} \chi(G(\phi)) \right| \le 2 \deg(g) q^{1/2}.$$

THEOREM 8. *If $\mu_1, \ldots, \mu_k \in \mathbb{F}_q$ are not all 0, then for the character sum in (17) we have*

$$|S| \le 2kq^{1/2} + 5k + 5.$$

PROOF. First we consider

$$S_1 := \sum_{n=0}^{q} \chi\left(\sum_{j=1}^{k} \mu_j \psi_{d_j}(\psi_n(\gamma_0))\right).$$

We have

$$|S - S_1| \le 2\#\{n \in \{0, \ldots, q\} : \exists j \in \{1, \ldots, k\},\ \psi_{n+d_j}(\gamma_0) \neq \psi_{d_j}(\psi_n(\gamma_0))\}.$$

Since $\varepsilon_1, \ldots, \varepsilon_q$ are distinct elements of $\mathbb{F}_q$, there is exactly one $n_0 \in \{1, \ldots, q\}$ such that $\gamma_0 = \varepsilon_{n_0}$. By (16) it might happen that $\psi_{d_j}(\psi_{n_0}(\gamma_0)) \neq R_{n_0+d_j}(\gamma_0)$. For $n \in \{0, \ldots, q\} \setminus \{n_0\}$ we have $\psi_{d_j}(\psi_n(\gamma_0)) = R_{n+d_j}(\gamma_0)$ except possibly if $\psi_n(\gamma_0) = \gamma_n = \varepsilon_{d_j}$. Since the sequence $\gamma_0, \ldots, \gamma_q$ contains all elements of $\mathbb{F}_q$,

$$\#\{n \in \{0, \ldots, q\} \setminus \{n_0\} : \gamma_n \in \{\varepsilon_{d_1}, \ldots, \varepsilon_{d_k}\}\} \le k + 1.$$

Thus, we have $\psi_{d_j}(\psi_n(\gamma_0)) = R_{n+d_j}(\gamma_0)$ for $1 \le j \le k$, except for at most $k + 2$ values of n.

For $n \in \{0, \ldots, q\}$ we have $\psi_{n+d_j}(\gamma_0) = R_{n+d_j}(\gamma_0)$ except if $\gamma_0 = \varepsilon_{n+d_j}$. Since $\varepsilon_1, \ldots, \varepsilon_q$ are distinct elements of $\mathbb{F}_q$, for each j there is at most one $n \in \{0, \ldots, q\}$ such that $\gamma_0 = \varepsilon_{n+d_j}$. Thus, we have $\psi_{n+d_j}(\gamma_0) = R_{n+d_j}(\gamma_0)$ for $1 \le j \le k$, except for at most k values of n. Hence we obtain

$$|S - S_1| \le 2(2k + 2). \tag{18}$$

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

Since the sequence $\gamma_0, \gamma_1, \ldots$ contains all elements of $\mathbb{F}_q$, if we write

$$S_2 := \sum_{\phi \in \mathbb{F}_q} \chi\left(\sum_{j=1}^{k} \mu_j \psi_{d_j}(\phi)\right),$$

then

$$|S_1 - S_2| \le 1.$$

Let $\mathcal{G} = \{\varepsilon_{d_1}, \ldots, \varepsilon_{d_k}\}$ and

$$S_3 := \sum_{\phi \in \mathbb{F}_q \setminus \mathcal{G}} \chi(G(\phi)),$$

where for $\phi \in \mathbb{F}_q \setminus \mathcal{G}$,

$$G(\phi) := \sum_{j=1}^{k} \mu_j \, R_{d_j}(\phi).$$

Then

$$|S_2 - S_3| \le k.$$

By Lemma 7

$$|S_3| \le 2kq^{1/2},$$

which implies $|S_2| \le 2kq^{1/2} + k$, then $|S_1| \le 2kq^{1/2} + k + 1$, and finally $|S| \le 2kq^{1/2} + 5k + 5$. □

4. The Gowers norm of inversive sequences

Let p be an odd prime and let $\gamma_0, \gamma_1, \ldots$ be the sequence of elements of $\mathbb{F}_p$ of least period $p+1$ constructed in (15). We derive a binary sequence $e_0, e_1, \ldots$ of period $p+1$ by setting

$$e_n = \begin{cases} +1 & \text{if } 0 \le \gamma_n \le (p-1)/2, \\ -1 & \text{if } (p+1)/2 \le \gamma_n \le p-1. \end{cases}$$

Consider $f : \mathbb{Z}_{p+1} \to \{-1, +1\}$ defined by $f(n) = e_n$ for all $n \in \mathbb{Z}_{p+1}$. We want to bound the quantity $G_d(f)$ in (1) for this function f, or equivalently for the sequence $e_0, e_1, \ldots$.

The main tools in estimating $G_d(f)$ are Theorem 6 and a connection between $M_d(f)$ and discrepancy. For integers $0 \le d_1 < \cdots < d_k \le p$, consider the $p+1$ points

$$\mathbf{x}_n = \frac{1}{p}(\gamma_{n+d_1}, \ldots, \gamma_{n+d_k}) \in [0, 1)^k, \quad n = 0, 1, \ldots, p,$$

and let $D_{p+1}(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_p)$ be the discrepancy of $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_p$. Then [3, Theorem 1] implies the following result.

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

LEMMA 9. *For any* $0 \le d_1 < \cdots < d_k \le p$,

$$\left|\sum_{n=0}^{p} e_{n+d_1} \cdots e_{n+d_k}\right| \le 2^k (p+1) D_{p+1}(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_p).$$

THEOREM 10. *Let* $f : \mathbb{Z}_{p+1} \to \{-1, +1\}$ *be as given above. Then for any* $d \ge 1$ *we have*

$$|G_d(f)| = O(p^{d+1/2} (\log p)^{2^d})$$

with an implied constant depending only on d.

PROOF. Since it is trivial that $|G_d(f)| \le (p+1)^{d+1}$, we can assume without loss of generality that $2^d \le p+1$. For integers $1 \le k \le 2^d$ and $0 \le d_1 < \cdots < d_k \le p$, we first consider the discrepancy $D_{p+1}(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_p)$. We will bound this discrepancy by means of [6, Corollary 3.11]. To this end, let $\mathbf{h} = (\mu_1, \ldots, \mu_k) \in \mathbb{F}_p^k$ with $\mu_1, \ldots, \mu_k$ not all 0. Put $\mathrm{e}(t) = \mathrm{e}^{2\pi i t}$ for $t \in \mathbb{R}$ and let $\langle \cdot, \cdot \rangle$ denote the standard inner product in $\mathbb{R}^k$. Then

$$\sum_{n=0}^{p} \mathrm{e}(\langle \mathbf{h}, \mathbf{x}_n \rangle) = \sum_{n=0}^{p} \mathrm{e}\left(\frac{1}{p} \sum_{j=1}^{k} \mu_j \gamma_{n+d_j}\right) = \sum_{n=0}^{p} \chi\left(\sum_{j=1}^{k} \mu_j \gamma_{n+d_j}\right),$$

where χ is the canonical additive character of $\mathbb{F}_p$. Hence Theorem 8 yields

$$\left|\sum_{n=0}^{p} \mathrm{e}(\langle \mathbf{h}, \mathbf{x}_n \rangle)\right| \le 2kp^{1/2} + 5k + 5.$$

Thus, we can apply [6, Corollary 3.11] with $B = 2kp^{1/2} + 5k + 5$, and we obtain

$$D_{p+1}(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_p) \le \frac{2kp^{1/2} + 5k + 5}{p+1} \left(\frac{4}{\pi^2} \log p + 1.72\right)^k + \frac{k}{p}.$$

Lemma 9 implies that

$$\left|\sum_{n=0}^{p} e_{n+d_1} \cdots e_{n+d_k}\right| \le 2^k (2kp^{1/2} + 5k + 5) \left(\frac{4}{\pi^2} \log p + 1.72\right)^k + \frac{2^k k(p+1)}{p}.$$

Since this bound does not depend on $d_1, \ldots, d_k$ and since $k \le 2^d \le p+1$, we get

$$M_d(f) \le 2^{2^d} (2^{d+1} p^{1/2} + 5 \cdot 2^d + 5) \left(\frac{4}{\pi^2} \log p + 1.72\right)^{2^d} + \frac{2^{2^d + d}(p+1)}{p}.$$

Now an application of Theorem 6 completes the proof. □

As we have noted in the proof of Theorem 10, the trivial bound on $G_d(f)$ is $|G_d(f)| \le (p+1)^{d+1}$. Thus, for any given $d \ge 1$, Theorem 10 yields a nontrivial bound on $G_d(f)$ for all sufficiently large primes p.

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

5. Concluding remarks

We have chosen to present the estimation of the Gowers norm of inversive sequences because the estimation of the complete character sums is much sharper than the estimation of the incomplete character sums we can obtain by present techniques for this construction. Of course, we can estimate the Gowers norm of many other sequences, including the Legendre symbol construction $e_n = (n/p)$ of [4], its generalisation $e_n = (f(n)/p)$ where f is a suitable polynomial [1], and so on. In all these constructions, the saving in the bound on the Gowers norm in comparison with that on the correlation measures of Mauduit and Sárközy [4] would be at most a log factor.

Acknowledgements

The research of the first author was carried out while he was hosted by CNRS-FR2291 (FRUMAM) at the Université de la Méditerranée in Marseille-Luminy. We thank Huaning Liu for discussions on the topic of this paper.

References

[1] L. Goubin, Ch. Mauduit and A. Sárközy, 'Construction of large families of pseudorandom binary sequences', *J. Number Theory* **106** (2004), 56–69.
[2] W. T. Gowers, 'A new proof of Szemerédi's theorem', *Geom. Funct. Anal.* **11** (2001), 465–588.
[3] Ch. Mauduit, H. Niederreiter and A. Sárközy, 'On pseudorandom [0, 1) and binary sequences', *Publ. Math. Debrecen* **71** (2007), 305–324.
[4] Ch. Mauduit and A. Sárközy, 'On finite pseudorandom binary sequences I: Measure of pseudorandomness, the Legendre symbol', *Acta Arith.* **82** (1997), 365–377.
[5] C. J. Moreno and O. Moreno, 'Exponential sums and Goppa codes: I', *Proc. Amer. Math. Soc.* **111** (1991), 523–531.
[6] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods* (SIAM, Philadelphia, PA, 1992).
[7] H. Niederreiter and J. Rivat, 'On the correlation of pseudorandom numbers generated by inversive methods', *Monatsh. Math.* **153** (2008), 251–264.
[8] H. Niederreiter, J. Rivat and A. Sárközy, 'Pseudorandom sequences of binary vectors', *Acta Arith.* **133** (2008), 109–125.
[9] H. Niederreiter and I. E. Shparlinski, 'Recent advances in the theory of nonlinear pseudorandom number generators', in: *Monte Carlo and Quasi-Monte Carlo Methods 2000* (Springer, Berlin, 2000), pp. 86–102.
[10] H. Niederreiter and A. Winterhof, 'Incomplete exponential sums over finite fields and their applications to new inversive pseudorandom number generators', *Acta Arith.* **93** (2000), 387–399.
[11] ———, 'On the structure of inversive pseudorandom number generators', in: *AAECC 2007*, Lecture Notes in Computer Science, 4851 (Springer, Berlin, 2007), pp. 208–216.

HARALD NIEDERREITER, Department of Mathematics,
National University of Singapore, 2 Science Drive 2, Singapore 117543,
Republic of Singapore
e-mail: nied@math.nus.edu.sg

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press

JOËL RIVAT, Institut de Mathématiques de Luminy, CNRS-UMR 6206,
Université de la Méditerranée, 163 avenue de Luminy, Case 907, 13288 Marseille Cedex 9, France
e-mail: rivat@iml.univ-mrs.fr

https://doi.org/10.1017/S0004972708001184 Published online by Cambridge University Press