BULL. AUSTRAL. MATH. SOC. 11L05, 11T71
VOL. 71 (2005) [405–409]

ON THE NONLINEARITY OF THE SEQUENCE OF SIGNS OF KLOOSTERMAN SUMS

IGOR E. SHPARLINSKI

It is known that Kloosterman sums with prime denominator p take real values, so one can define a sequence of signs of such sums. Several pseudorandom properties of this sequence have recently been studied by Fouvry, Michel, Rivat and Sárközy. Here we use one of their results to estimate a certain important characteristic of this sequence which is also of cryptographic interest.

1. INTRODUCTION

Let p be an odd prime. For an integer h we define the Kloosterman sum

$$K(h) = \sum_{u=1}^{p-1} \exp\Big(\frac{2\pi i}{p}(u + h\overline{u})\Big)$$

where $u\overline{u} \equiv 1 \pmod{p}$.

It is easy to show that $K(h)$ takes real values for any integer h, it is also known that $K(h) \neq 0$. Thus, one can define the sequence of signs

$$s_h = \begin{cases} 1, & \text{if } K(h) > 0; \\ -1, & \text{if } K(h) < 0; \end{cases} \qquad h = 0, 1, \ldots . \tag{1}$$

Several results about the distribution of values and autocorrelation and some of measures of pseudorandomess of this sequence have recently been obtained in [8]. Here we show that one of the results of [8] allows us to estimate one more characteristic of this sequence which also has an important cryptographic meaning. To be more precise we define an integer n by the inequalities $2^n \leqslant p < 2^{n+1}$ and denote by $\mathcal{B}_n$ the set of n-bit integers,

$$\mathcal{B}_n = \{h \in \mathbb{Z} : 0 \leqslant h \leqslant 2^n - 1\}.$$

Throughout the paper we do not distinguish between n-bit integers $h \in \mathcal{B}_n$ and their binary expansions. Thus $\mathcal{B}_n$ can be considered as the n-dimensional Boolean cube

Received 29th November, 2004
During the preparation of this paper, the author was supported in part by ARC grant DP0211459.

$\mathcal{B}_n = \{0,1\}^n$. In particular, for $h, r \in \mathcal{B}_n$, $\langle h, r\rangle$ denotes the inner product of h, r considered as the binary vectors.

Given a Boolean function $f : \mathcal{B}_n \to \{0,1\}$ we define its *Fourier coefficients* as

$$\widehat{f}(r) = 2^{-n} \sum_{h \in \mathcal{B}_n} (-1)^{f(h)+\langle h,r\rangle}, \qquad r \in \mathcal{B}_n.$$

Fourier coefficients are closely related to several complexity properties characterising the Boolean function f, such as the circuit complexity, the average sensitivity, the formula size, the average decision tree depth, the degrees of exact and approximate polynomial representations over the reals and several others; see [**1, 2, 9, 10, 11, 12**] and references therein.

Furthermore, we recall that

$$N(f) = 2^{n-1} - 2^{n-1} \max_{r \in \mathcal{B}_n} \left|\widehat{f}(r)\right|$$

is called the *nonlinearity* of f, see [**3, 4, 5, 6, 7, 14, 15**] for the cryptographic significance of this notion. The nonlinearity of f gives the smallest possible Hamming distance between the vector of values of f and the vector of values of a linear function in n variables over the $\mathbb{F}_2$, the field of two elements.

Here we use a certain result of [**8**] to obtain an upper bound on the nonlinearity of a Boolean function which is naturally associated with the sequence (1).

2. Main Result

We define the Boolean function

$$f(h) = \begin{cases} 0, & \text{if } s_h = 1 \text{ or } h = 0; \\ 1, & \text{if } s_h = -1; \end{cases} \qquad h \in \mathcal{B}_n. \tag{2}$$

In particular,

$$s_h = (-1)^{f(h)}, \qquad h \in \mathcal{B}_n,\ h \neq 0.$$

Theorem: *The nonlinearity $N(f)$ of the Boolean function $f(h)$, given by (2), satisfies the inequality*

$$N(f) = 2^{n-1}\left(1 + O(2^{-n/24} n^{1/12})\right).$$

Proof: We estimate the Fourier coefficients $\widehat{f}(k)$ of f by using the result that for any integers M, d_1, d_2 with $0 \leqslant M \leqslant M + d_1 < M + d_2 < 2^n$ we have

$$\sum_{h=0}^{M} s_{h+d_1} s_{h+d_2} = O\left(p^{5/6} (\log p)^{1/3}\right) = O(2^{5n/6} n^{1/3}), \tag{3}$$

which is a special case $l = 2$ of the result of [8, Theorem 1.1].

https://doi.org/10.1017/S0004972700038405 Published online by Cambridge University Press

We now fix some $m \leqslant n$ and write $h, r \in \mathcal{B}_n$ as

$$h = k + 2^m j \qquad \text{and} \qquad r = s + 2^m t,$$

with $0 \leqslant k, s < 2^m$ and $0 \leqslant j, t < 2^{n-m}$. In particular

$$\langle h, k \rangle = \langle k, s \rangle + \langle j, t \rangle.$$

Therefore,

$$\begin{aligned} |\widehat{f}(r)| &= |\widehat{f}(k + 2^m j)| \\ &= \left| 2^{-n} \sum_{k=0}^{2^m-1} \sum_{j=0}^{2^{n-m}-1} (-1)^{f(k+2^m j)+\langle k,s\rangle+\langle j,t\rangle} \right| \\ &\leqslant 2^{-n} \sum_{k=0}^{2^m-1} \left| \sum_{j=0}^{2^{n-m}-1} (-1)^{f(k+2^m j)+\langle j,t\rangle} \right|. \end{aligned}$$

Furthermore, using the Cauchy inequality we obtain

$$\begin{aligned} |\widehat{f}(k)|^2 &\leqslant 2^{m-2r} \sum_{k=0}^{2^m-1} \left| \sum_{j=0}^{2^{n-m}-1} (-1)^{f(k+2^m j)+\langle j,t\rangle} \right|^2 \\ &= 2^{m-2r} \sum_{k=0}^{2^m-1} \sum_{j_1,j_2=0}^{2^{n-m}-1} (-1)^{f(k+2^m j_1)+f(k+2^m j_2)+\langle j_1,t\rangle+\langle j_2,t\rangle} \\ &\leqslant 2^{m-2r} \sum_{j_1,j_2=0}^{2^{n-m}-1} \left| \sum_{k=0}^{2^m-1} (-1)^{f(k+2^m j_1)+f(k+2^m j_2)} \right| \\ &\leqslant 2^{m-2r} \sum_{j_2,j_2=0}^{2^{n-m}-1} \left| \sum_{k=0}^{2^m-1} s_{k+2^m j_1} s_{k+2^m j_2} + O(1) \right|, \end{aligned}$$

where $O(1)$ takes care of the terms corresponding to $k = j_1 = 0$ or to $k = j_2 = 0$. For 2^{n-m} choices of $j_1 = j_2$ the sums over k is equal to 2^m. For the other choices of j_1 and j_2 we can use the bound (3), getting

$$\begin{aligned} |\widehat{f}(k)|^2 &= O\big(2^{m-2n}(2^{n-m}2^m + 2^{2(n-m)}2^{5n/6}n^{1/3})\big) \\ &= O(2^{m-n} + 2^{5n/6-m}n^{1/3}). \end{aligned}$$

Defining m by the inequalities $2^m \leqslant 2^{11n/12}n^{1/6} < 2^{m+1}$, we conclude the proof. ∎

3. Remarks

We recall, that the Legendre symbol satisfies

$$G(h) = \left(\frac{h}{p}\right) G(1)$$

https://doi.org/10.1017/S0004972700038405 Published online by Cambridge University Press

and thus can be viewed as the "sign" of the Gauss sums

$$G(h) = \sum_{u=0}^{p-1} \exp\Big(\frac{2\pi i}{p} h u^2\Big).$$

The Fourier coefficients of the Boolean function

$$g(h) = \begin{cases} 0, & \text{if } \left(\frac{h}{p}\right) = 1 \text{ or } h = 0; \\ 1, & \text{if } \left(\frac{h}{p}\right) = -1; \end{cases} \qquad h \in \mathcal{B}_n,$$

has been estimated in [**13**, Theorem 10.1] which immediately yields the bound

$$N(g) = 2^{n-1}\big(1 + O(2^{-n/8} n^{1/4})\big).$$

Several other properties of g have been studied in [**13**] as well, for example its linear complexity (that is, the length of the shortest linear recurrence over $\mathbb{F}_2$ satisfied by this sequence). Obtaining analogues of these results for f, given by (2) would be of interest too.

Certainly generating the sequence s_h is too complicated to lead to a practical pseudorandom number generator. Nevertheless, a more detailed study of various properties of this sequence is of ultimate interest.

References

[1] A. Bernasconi, 'On the complexity of balanced Boolean functions', *Inform. Process Lett.* **70** (1999), 157–163.

[2] R.B. Boppana, 'The average sensitivity of bounded-depth circuits', *Inform. Process Lett.* **63** (1997), 257–261.

[3] C. Carlet, 'On cryptographic complexity of Boolean functions', in *Finite Fields with Applications to Coding Theory, Cryptography and Related Areas* (Springer-Verlag, Berlin, 2002), **pp.** 53–69.

[4] C. Carlet, 'On the degree, nonlinearity, algebraic thickness, and nonnormality of Boolean functions, with developments on symmetric functions', *IEEE Trans. Inform. Theory* **50** (2004), 2178–2185.

[5] C. Carlet and C. Ding, 'Highly nonlinear mappings', *J. Complexity* **20** (2004), 205–244.

[6] C. Carlet and P. Sarkar, 'Spectral domain analysis of correlation immune and resilient Boolean functions', *Finite Fields Appl.* **8** (2002), 120–130.

[7] C. Carlet and Y. Tarannikov, 'Covering sequences of Boolean functions and their cryptographic significance', *Des. Codes Cryptogr.* **25** (2002), 263–279.

[8] É. Fouvry, P. Michel, J. Rivat and A. Sárközy, 'On the pseudorandomness of the signs of Kloosterman sums', *J. Austral. Math. Soc.* **77** (2004), 425–436.

[9] M. Goldmann, 'Communication complexity and lower bounds for simulating threshold circuits', in *Theoretical Advances in Neural Computing and Learning* (Kluwer Acad. Publ., Dordrecht, 1994), **pp.** 85–125.

[10] N. Linial, Y. Mansour and N. Nisan, 'Constant depth circuits, Fourier transform, and learnability', *J. Assoc. Comput. Mach.* **40** (1993), 607–620.

[11] Y. Mansour, 'Learning Boolean functions via the Fourier transform', in *Theoretical Advances in Neural Computing and Learning* (Kluwer Academic Publ., Dordrecht, 1994), pp. 391–424.

[12] V. Roychowdhry, K.-Y. Siu and A. Orlitsky, 'Neural models and spectral methods', in *Theoretical Advances in Neural Computing and Learning* (Kluwer Academic Publ., Dordrecht, 1994), pp. 3–36.

[13] I.E. Shparlinski, *Cryptographic applications of analytic number theory* (Birkhäuser, Basel, 2003).

[14] P. Štanikă, 'Nonlinearity, local and global avalanche characteristics of balanced Boolean functions', *Discrete Math.* **248** (2002), 181–193.

[15] Y. Zheng, and X.M. Zhang, 'Connections among nonlinearity, avalanche and correlation immunity', *Theoret. Comput. Sci.* **292** (2003), 697–710.

Department of Computing
Macquarie University
Sydney, NSW 2109
Australia
e-mail: igor@ics.mq.edu.au