ACTA ARITHMETICA
LXIX.2 (1995)

An improved upper bound for the discrepancy of quadratic congruential pseudorandom numbers

by

Jürgen Eichenauer-Herrmann (Darmstadt) and
Harald Niederreiter (Wien)

1. Introduction and main result. Number theory plays an important role in the theory of various methods for generating uniform pseudorandom numbers in the interval $[0, 1)$. A well-known example is provided by the classical linear congruential method for the generation of uniform pseudorandom numbers, in which number-theoretic techniques are heavily used in the analysis of distribution properties and of the lattice structure (see [6, Chapter 3]). The family of nonlinear congruential methods represents another area of the theory of pseudorandom number generation where significant applications of number theory occur. These nonlinear congruential methods of generating uniform pseudorandom numbers have been studied intensively during the last years. Reviews of the development of this important area can be found in the survey articles [1–3], [7], [8], [10] and in the monograph [9]. The earliest nonlinear congruential approach is the *quadratic congruential method* proposed by Knuth [6, p. 25], which is considered in the present paper in the case of an odd prime power modulus $m = p^\omega$ with some prime $p \geq 3$ and an integer $\omega \geq 2$. Let $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ for integers $n \geq 1$. For parameters $a, b, c, y_0 \in \mathbb{Z}_m$ a *quadratic congruential sequence* $(y_n)_{n\geq 0}$ of elements of $\mathbb{Z}_m$ is defined by

$$y_{n+1} \equiv ay_n^2 + by_n + c \pmod{m}, \qquad n \geq 0.$$

A sequence $(x_n)_{n\geq 0}$ of *quadratic congruential pseudorandom numbers* in the interval $[0, 1)$ is obtained by $x_n = y_n/m$ for $n \geq 0$. The sequences $(x_n)_{n\geq 0}$ and $(y_n)_{n\geq 0}$ are purely periodic with the maximum possible period length m if and only if the conditions $a \equiv 0 \pmod{p}$, $b \equiv 1 \pmod{p}$, $c \not\equiv 0 \pmod{p}$, and $a \not\equiv 3c \pmod{9}$ for $p = 3$ are satisfied [6, p. 34]. We assume from now on that these conditions for the maximum possible period length hold.

Statistical independence properties of the generated sequences, which are very important for their usability in a stochastic simulation, can be analysed based on the discrepancy of s-tuples of successive pseudorandom numbers with $s \ge 2$. For N arbitrary points $\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{N-1} \in [0,1)^s$ the *discrepancy* is defined by

$$D_N(\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{N-1}) = \sup_J |F_N(J) - V(J)|,$$

where the supremum is extended over all subintervals J of $[0,1)^s$, $F_N(J)$ is N^{-1} times the number of points among $\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{N-1}$ falling into J, and $V(J)$ denotes the s-dimensional volume of J. In the present paper the pairs

$$\mathbf{x}_n = (x_n, x_{n+1}) \in [0,1)^2, \qquad n \ge 0,$$

of successive quadratic congruential pseudorandom numbers are considered and the abbreviation

$$D_m^{(2)} = D_m(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{m-1})$$

is used.

In [4] upper and lower bounds for the discrepancy $D_m^{(2)}$ have been established. These results suggest that it is reasonable to choose the parameter a in such a way that $a \not\equiv 0 \pmod{p^2}$. Then the upper bound for $D_m^{(2)}$ has the form

$$D_m^{(2)} < (4 + 5p^{-3/2}) m^{-1/2} p^{1/2} \left(\frac{1}{\pi} \log m + \frac{1}{5} \right)^2 + 2m^{-1},$$

i.e., the upper bound is of an order of magnitude $m^{-1/2} p^{1/2} (\log m)^2$. It should be observed that the discrepancy of m independent and uniformly distributed random points from $[0,1)^2$ is almost always of an order of magnitude between $m^{-1/2}$ and $m^{-1/2}(\log\log m)^{1/2}$ according to the law of the iterated logarithm for discrepancies [5]. The following main result of the present paper provides an improved upper bound for $D_m^{(2)}$ which is of an order of magnitude $m^{-1/2}(p^{1/2} + p^{-1/2}(\log m)^2)$. Hence, for $\omega = \omega(p) \sim p^{1/2}(\log p)^{-1}$, its order of magnitude can be made as small as $m^{-1/2} \log m$.

THEOREM. *The discrepancy $D_m^{(2)}$ of pairs in the quadratic congruential method with modulus $m = p^\omega$ and $a \not\equiv 0 \pmod{p^2}$ satisfies*

$$D_m^{(2)} < (4 + 5p^{-3/2}) m^{-1/2} \times \left(\frac{1}{9} p^{1/2} + \frac{1}{\pi^2} p^{-1/2} \left(\log m + \frac{2\pi}{3} \log p \right) (\log m + 1.395) \right) + 2m^{-1}.$$

2. Auxiliary results. First, some further notation is necessary. For integers $k \ge 1$ and $q \ge 2$ let $C_k(q)$ be the set of all nonzero lattice points $(h_1, \ldots, h_k) \in \mathbb{Z}^k$ with $-q/2 < h_j \le q/2$ for $1 \le j \le k$. Define

$$r(h,q) = \begin{cases} q \sin(\pi |h|/q) & \text{for } h \in C_1(q), \\ 1 & \text{for } h = 0, \end{cases}$$

and

$$r(\mathbf{h}, q) = \prod_{j=1}^{k} r(h_j, q)$$

for $\mathbf{h} = (h_1, \ldots, h_k) \in C_k(q)$. For real t the abbreviation $e(t) = e^{2\pi i t}$ is used, and $\mathbf{u} \cdot \mathbf{v}$ stands for the standard inner product of $\mathbf{u}, \mathbf{v} \in \mathbb{R}^k$. Subsequently, two known results are stated which follow from [9, Theorem 3.10] and [4, Lemma 7(a)], respectively. The third lemma is crucial for the proof of the main result.

Lemma 1. *Let $N \geq 1$ and $q \geq 2$ be integers. Let $\mathbf{t}_n = \mathbf{y}_n/q \in [0,1)^k$ with $\mathbf{y}_n \in \mathbb{Z}_q^k$ for $0 \leq n < N$. Then the discrepancy of the points $\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{N-1}$ satisfies*

$$D_N(\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{N-1}) \leq \frac{k}{q} + \frac{1}{N} \sum_{\mathbf{h} \in C_k(q)} \frac{1}{r(\mathbf{h}, q)} \Big| \sum_{n=0}^{N-1} e(\mathbf{h} \cdot \mathbf{t}_n) \Big|.$$

Lemma 2. *Let $(\mathbf{x}_n)_{n \geq 0}$ be the sequence of pairs of successive quadratic congruential pseudorandom numbers as defined above. Let $\mathbf{h} = (h_1, h_2) \in C_2(m)$ with $\gcd(h_2, p^{\omega-1}) = p^{\nu}$ and $\nu \in \{0, 1, \ldots, \omega - 1\}$. Then*

$$\Big| \sum_{n=0}^{m-1} e(\mathbf{h} \cdot \mathbf{x}_n) \Big| = \begin{cases} p^{(\omega+\nu+1)/2} & \textit{for } h_1 + h_2 \equiv 0 \ (\operatorname{mod} p^{\nu+1}), \\ 0 & \textit{for } h_1 + h_2 \not\equiv 0 \ (\operatorname{mod} p^{\nu+1}). \end{cases}$$

Lemma 3. *Let $q = p^{\alpha}$ with some prime $p \geq 3$ and an integer $\alpha \geq 1$. Then*

$$\sum_{\substack{\mathbf{h} = (h_1, h_2) \in C_2(q) \\ h_1 h_2 \not\equiv 0 \ (\operatorname{mod} p) \\ h_1 + h_2 \equiv 0 \ (\operatorname{mod} p)}} \frac{1}{r(\mathbf{h}, q)} < \frac{4}{\pi^2 p} \left(\log q + \frac{2\pi}{3} \log p \right) (\log q + 1.395) + \frac{4}{9}.$$

Proof. (i) First, two preliminary estimates are established. Straightforward calculations show that

$$\sum_{d=1}^{(p-1)/2} \frac{p}{d(p-d)} < \frac{p}{p-1} + \int_1^{p/2} \frac{p}{x(p-x)}\, dx$$
$$= \log p + \frac{p}{p-1} - \log\left(\frac{p}{p-1}\right) < \log p + 1.095$$

and

$$\sum_{d=1}^{(p-1)/2} \frac{p^2}{d^2(p-d)^2} < \frac{p^2}{(p-1)^2} + \int_1^{p/2} \frac{p^2}{x^2(p-x)^2}\, dx$$
$$= 2 + \frac{p}{(p-1)^2} + \frac{2}{p} \log(p-1)$$

$$= \frac{2}{p}\log p + 2 + \frac{1}{p}\left(\frac{p^2}{(p-1)^2} - \log\left(\frac{p^2}{(p-1)^2}\right)\right)$$
$$< \frac{2}{p}\log p + 2 + \frac{1.44}{p}.$$

(ii) Now, for $\alpha \ge 2$ and any integer $d \in \{1, \ldots, p-1\}$ one obtains

$$\sum_{\substack{h=1\\ h\equiv d \,(\mathrm{mod}\, p)}}^{(q-1)/2} \frac{1}{r(h,q)}$$
$$< \frac{1}{r(d,q)} + \int\limits_0^{(q-2d)/(2p)} \frac{1}{q\sin(\pi(px+d)/q)}\,dx$$
$$= \frac{1}{r(d,q)} - \frac{1}{\pi p}\log(\tan(\pi d/(2q))) < \frac{1}{r(d,q)} - \frac{1}{\pi p}\log(\pi d/(2q))$$
$$< \frac{1}{q\sin(\pi d/q)} + \frac{1}{\pi p}\log q - \frac{0.143}{p} < \frac{1}{3d} + \frac{1}{\pi p}\log q - \frac{0.143}{p},$$

where in the last step $6d \le q$ has to be assumed. It follows by inspection that the resulting estimate remains valid for $d = 2$ and $q = 9$. Therefore

$$\sum_{\substack{h\in C_1(q)\\ h\equiv d\,(\mathrm{mod}\, p)}} \frac{1}{r(h,q)} = \sum_{\substack{h=1\\ h\equiv d\,(\mathrm{mod}\, p)}}^{(q-1)/2} \frac{1}{r(h,q)} + \sum_{\substack{h=1\\ h\equiv p-d\,(\mathrm{mod}\, p)}}^{(q-1)/2} \frac{1}{r(h,q)}$$
$$< \frac{p}{3d(p-d)} + \frac{2}{\pi p}\log q - \frac{0.286}{p}$$

for $\alpha \ge 2$ and any $d \in \{1, \ldots, p-1\}$.

(iii) Finally, it follows from the estimates in (ii) and (i) that for $\alpha \ge 2$,

$$\sum_{\substack{\mathbf{h}=(h_1,h_2)\in C_2(q)\\ h_1h_2\not\equiv 0\,(\mathrm{mod}\, p)\\ h_1+h_2\equiv 0\,(\mathrm{mod}\, p)}} \frac{1}{r(\mathbf{h},q)}$$
$$= \sum_{d=1}^{p-1} \sum_{\substack{h_1\in C_1(q)\\ h_1\equiv d\,(\mathrm{mod}\, p)}} \sum_{\substack{h_2\in C_1(q)\\ h_2\equiv p-d\,(\mathrm{mod}\, p)}} \frac{1}{r(h_1,q)r(h_2,q)}$$
$$< \sum_{d=1}^{p-1} \left(\frac{p}{3d(p-d)} + \frac{2}{\pi p}\log q - \frac{0.286}{p}\right)^2$$
$$< \frac{2}{9}\left(\frac{2}{p}\log p + 2 + \frac{1.44}{p}\right) + \frac{4}{3}\left(\frac{2}{\pi p}\log q - \frac{0.286}{p}\right)(\log p + 1.095)$$

$$+ p\left(\frac{2}{\pi p}\log q - \frac{0.286}{p}\right)^2$$
$$< \frac{4}{\pi^2 p}\left(\log q + \frac{2\pi}{3}\log p\right)(\log q + 1.395) + \frac{4}{9},$$

which is the desired result.

(iv) For $\alpha = 1$, it follows from $\sin x > x(\pi - x)/\pi$ for $x \in (0,\pi)$ and the second part of (i) that

$$\sum_{\substack{\mathbf{h}=(h_1,h_2)\in C_2(p)\\ h_1h_2\not\equiv 0\ (\mathrm{mod}\, p)\\ h_1+h_2\equiv 0\ (\mathrm{mod}\, p)}} \frac{1}{r(\mathbf{h},p)} = 2\sum_{d=1}^{(p-1)/2} \frac{1}{(p\sin(\pi d/p))^2}$$
$$< \frac{2}{\pi^2}\sum_{d=1}^{(p-1)/2} \frac{p^2}{d^2(p-d)^2} < \frac{4}{\pi^2 p}(\log p + 0.72) + \frac{4}{\pi^2},$$

which completes the proof. ■

3. Proof of the Theorem.

First, Lemma 1 is applied with $k = 2$, $q = N = m$, and $\mathbf{t}_n = \mathbf{x}_n$ for $0 \le n < m$. This yields

$$D_m^{(2)} \le \frac{2}{m} + \frac{1}{m}\sum_{\mathbf{h}\in C_2(m)} \frac{1}{r(\mathbf{h},m)}\Big|\sum_{n=0}^{m-1} e(\mathbf{h}\cdot\mathbf{x}_n)\Big|$$
$$= \frac{2}{m} + \frac{1}{m}\sum_{\nu=0}^{\omega-1}\sum_{\substack{\mathbf{h}=(h_1,h_2)\in C_2(m)\\ \gcd(h_2,p^{\omega-1})=p^\nu}} \frac{1}{r(\mathbf{h},m)}\Big|\sum_{n=0}^{m-1} e(\mathbf{h}\cdot\mathbf{x}_n)\Big|$$
$$= \frac{2}{m} + \frac{p^{1/2}}{m^{1/2}}\sum_{\nu=0}^{\omega-1} p^{\nu/2}\sum_{\substack{\mathbf{h}=(h_1,h_2)\in C_2(m)\\ \gcd(h_2,p^{\omega-1})=p^\nu\\ h_1+h_2\equiv 0\ (\mathrm{mod}\, p^{\nu+1})}} \frac{1}{r(\mathbf{h},m)}$$
$$= \frac{2}{m} + \frac{p^{1/2}}{m^{1/2}}\sum_{\nu=0}^{\omega-1} p^{-3\nu/2}\sum_{\substack{\mathbf{g}=(g_1,g_2)\in C_2(p^{\omega-\nu})\\ g_1g_2\not\equiv 0\ (\mathrm{mod}\, p)\\ g_1+g_2\equiv 0\ (\mathrm{mod}\, p)}} \frac{1}{r(\mathbf{g},p^{\omega-\nu})},$$

where in the penultimate step Lemma 2 has been used. Now, it follows from Lemma 3 that

$$D_m^{(2)} < \frac{2}{m} + \frac{p^{1/2}}{m^{1/2}}\sum_{\nu=0}^{\omega-1} p^{-3\nu/2}$$
$$\times\left(\frac{4}{\pi^2 p}\left(\log p^{\omega-\nu} + \frac{2\pi}{3}\log p\right)(\log p^{\omega-\nu} + 1.395) + \frac{4}{9}\right)$$

$$
\begin{aligned}
&< \frac{2}{m} + \frac{p^{1/2}}{m^{1/2}}\Big(\sum_{\nu=0}^{\infty}(p^{-3/2})^{\nu}\Big)\\
&\quad\times\left(\frac{4}{\pi^2 p}\left(\log m + \frac{2\pi}{3}\log p\right)(\log m + 1.395) + \frac{4}{9}\right)\\
&< \frac{2}{m} + \frac{p^{1/2}}{m^{1/2}}\left(1 + \frac{5}{4p^{3/2}}\right)\\
&\quad\times\left(\frac{4}{\pi^2 p}\left(\log m + \frac{2\pi}{3}\log p\right)(\log m + 1.395) + \frac{4}{9}\right),
\end{aligned}
$$

which yields the desired result. ■

References

[1] J. Eichenauer-Herrmann, *Inversive congruential pseudorandom numbers: a tutorial*, Internat. Statist. Rev. 60 (1992), 167–176.

[2] —, *Inversive congruential pseudorandom numbers*, Z. Angew. Math. Mech. 73 (1993), T644–T647.

[3] —, *Pseudorandom number generation by nonlinear methods*, Internat. Statist. Rev., to appear.

[4] J. Eichenauer-Herrmann and H. Niederreiter, *On the discrepancy of quadratic congruential pseudorandom numbers*, J. Comput. Appl. Math. 34 (1991), 243–249.

[5] J. Kiefer, *On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm*, Pacific J. Math. 11 (1961), 649–660.

[6] D. E. Knuth, *The Art of Computer Programming*, Vol. 2, *Seminumerical Algorithms*, 2nd ed., Addison-Wesley, Reading, Mass., 1981.

[7] H. Niederreiter, *Recent trends in random number and random vector generation*, Ann. Oper. Res. 31 (1991), 323–345.

[8] —, *Nonlinear methods for pseudorandom number and vector generation*, in: Simulation and Optimization, G. Pflug and U. Dieter (eds.), Lecture Notes in Econom. and Math. Systems 374, Springer, Berlin, 1992, 145–153.

[9] —, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, Penn., 1992.

[10] —, *Pseudorandom numbers and quasirandom points*, Z. Angew. Math. Mech. 73 (1993), T648–T652.

FACHBEREICH MATHEMATIK
TECHNISCHE HOCHSCHULE DARMSTADT
SCHLOSSGARTENSTRASSE 7
D-64289 DARMSTADT, F.R.G.

INSTITUT FÜR INFORMATIONSVERARBEITUNG
ÖSTERREICHISCHE AKADEMIE
DER WISSENSCHAFTEN
SONNENFELSGASSE 19
A-1010 WIEN, AUSTRIA
E-mail: NIED@QIINFO.OEAW.AC.AT

Received on 18.7.1994 (2640)