VERSITA

Mathematica Slovaca

DOI: 10.2478/s12175-009-0128-1
Math. Slovaca **59** (2009), No. 3, 315–322

PSEUDO-RANDOMNESS OF VAN DER CORPUT'S SEQUENCES

O. Blažeková

(*Communicated by Stanislav Jakubec*)

ABSTRACT. The aim of this paper is to find main terms of the star D_N^* and extremal D_N discrepancies of the two dimensional sequence (x_n, x_{n+1}), $n = 0, 1, 2, \dots, N-1$, where x_n, $n = 0, 1, 2, \dots$, is the van der Corput sequence. This give a quantitative form of a well-known result that van der Corput sequence is not pseudorandom.

1. Introduction

In 1936 van der Corput introduced the following sequence (see [KN, p. 127], [DT, p. 368], and [SP, pp. 2–102]): Let $q \geq 2$ be an integer and

$$n = a_{k(n)}(n)q^{k(n)} + \cdots + a_1(n)q + a_0(n), \qquad a_j(n) \in \{0, 1, \dots, q-1\}, \ a_{k(n)} > 0,$$

be the q-adic digit expansion of integer n in the base q. Then the van der Corput sequence $\gamma_q(n)$, $n = 0, 1, 2, \dots$, in the base q is defined by

$$\gamma_q(n) = \frac{a_0(n)}{q} + \frac{a_1(n)}{q^2} + \cdots + \frac{a_{k(n)}(n)}{q^{k(n)+1}}. \tag{1}$$

In other words, if $n = a_k a_{k-1} \dots a_1 a_0$, then $\gamma_q(n) = 0.a_0 a_1 \dots a_k$. The $\gamma_q(n)$ is called the radical inverse function.

The van der Corput sequence is uniformly distributed (see below) and if $N > q$ then its star discrepancy D_N^* satisfies

$$D_N^*(\gamma_q(n)) < \frac{1}{N}\left(\frac{q\log(qN)}{\log q}\right). \tag{2}$$

2000 Mathematics Subject Classification: Primary 11K31, 11K38, 11K45.
Keywords: van der Corput sequence, discrepancy, pseudo-randomness, q-adic digit expansion of integer, uniformly distributed sequence.
This research was supported by the Slovak Academy of Sciences Vega Grant No. 2/7138/27.

Here, if $x_0, \dots, x_{N-1}$ is a given sequence of real numbers from the unit interval $[0,1)$, then the number

$$D_N^*(x_n) = \sup_{x\in[0,1]} \left| \frac{A([0,x);N;x_n)}{N} - x \right|$$

is called star discrepancy, where the counting function $A(I;N;x_n)$ is defined as the number of terms of x_n with $0 \le n \le N-1$, and with x_n belonging to I, i.e. $A(I;N;x_n) = \#\{0 \le n < N : \{x_n\} \in I\}$. In the following we also use so-called extremal discrepancy D_N defined as

$$D_N(x_n) = \sup_{I\subset[0,1]} \left| \frac{A(I;N;x_n)}{N} - |I| \right|.$$

The infinite sequence x_n, $n = 1,2,\dots$, is uniformly distributed (abbreviating u.d.) if $D_N(x_n) \to 0$ as $N \to \infty$ (or equivalently $D_N^*(x_n) \to 0$). Similar definitions hold for discrepancies $D_N^*(\mathbf{x}_n)$ and $D_N(\mathbf{x}_n)$ of a multidimensional sequence $\mathbf{x}_n$, $n = 1,2,\dots,N$.

The infinite sequence x_n, $n = 1,2,\dots$, is called low discrepancy sequence if $D_N(x_n) = O((\log N)/N)$. By (1) van der Corput sequence $\gamma_q(n)$ is low discrepancy but in Part 2 we prove that it is not pseudorandomness since the two-dimensional sequence

$$(\gamma_q(n), \gamma_q(n+1)), \qquad n = 0,1,2,\dots, \tag{3}$$

is not u.d. (see Theorem 1 or 2). There is no formal fully satisfactory definition of the pseudorandomness of a sequence, we have only a scale of test which such a sequence should satisfy, see [DT, pp. 424–430], [SP, pp. 2–243, 2.25]. By D. E. Knuth's test of pseudorandomness of x_n, $n = 0,1,2,\dots$, all of sequences (x_n, x_{n+1}), $n = 0,1,2,\dots$; (x_n, x_{n+1}, x_{n+2}), $n = 0,1,2,\dots$; etc. must be u.d.

2. Star and extremal discrepancies

Theorem 1. *The star discrepancy D_N^* of the sequence $(\gamma_q(n), \gamma_q(n+1))$, $n = 0,1,2,\dots,N-1$ satisfies*

$$D_N^*((\gamma_q(n),\gamma_q(n+1))) = \max\left(\frac{1}{q}\left(1-\frac{1}{q}\right), \frac{1}{4}\left(1-\frac{1}{q}\right)^2\right) + O(D_N(\gamma_q(n))).$$

Proof. Let $n = a_k a_{k-1} \dots a_1 a_0$ be the q-adic expression of n and we find $a_j < q-1$ with minimal j. Then $n+1 = a_k a_{k-1} \dots (a_j+1)000\dots0$ and $\gamma_q(n) = 0.a_0 a_1 \dots a_k$ and $\gamma_q(n+1) = 0.00\dots0(a_j+1)a_{j+1}\dots a_k$. Thus

$$\gamma_q(n+1) = \gamma_q(n) + \frac{1}{q^{j+1}} + \frac{1}{q^j} - 1$$

(cf. I. M. Sobol' [So, p. 176]). From it, it follows that every point from the sequence $(\gamma_q(n), \gamma_q(n+1))$, $n = 0, 1, 2, \dots$, lies on the following lines

$$y = x + \frac{1}{q^{j+1}} + \frac{1}{q^j} - 1, \qquad j = 0, 1, 2, \dots,$$

see the Fig. 1. Denote

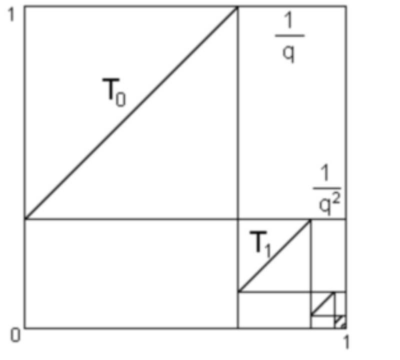

FIGURE 1. Line segments T_j, $j = 1, 2, \dots$

- T_j is the line segment

$$T_j = \left\{ (x,y) \in [0,1]^2 : \ y = x + \frac{1}{q^{j+1}} + \frac{1}{q^j} - 1, \ x \in \left[1 - \frac{1}{q^j}, 1 - \frac{1}{q^{j+1}} \right) \right\};$$

- $T = \bigcup_{j=0}^{\infty} T_j$;
- $T(I) = T \cap I$, where I is a subinterval of the unit square $[0,1]^2$;
- $t(I)$ is the projection of $T(I)$ to the x-axis (or y-axis) i.e. $t(I) = \{x : \ (x,y) \in T(I)\}$;
- $|J|$ is the length of the interval $J \subset [0,1]$ or area of $J \subset [0,1]^2$.

For computing of the star discrepancy $D_N^*((\gamma_q(n), \gamma_q(n+1)))$ we use the extremal discrepancy $D_N(\gamma_q(n))$ and the following relation

$$A\big(I; N; (\gamma_q(n), \gamma_q(n+1))\big) = A\big(t(I); N; \gamma_q(n)\big),$$

where $t(I) \subset [0,1]$ is an interval or a sum of two intervals for every rectangle (i.e. subinterval) $I \subset [0,1]^2$. Furthermore we use u.d. of $\gamma_q(n)$ mod 1 which gives

$$A\big(t(I); N; \gamma_q(n)\big) = N|t(I)| + O\big(N \cdot D_N(\gamma_q(n))\big).$$

Thus for $I = [0,x) \times [0,y)$ the star discrepancy D_N^* of $\big(\gamma_q(n), \gamma_q(n+1)\big)$, $n = 0, 1, \dots$, satisfies

$$D_N^* = \sup_{I \subset [0,1]^2} \big| |t(I)| - |I| \big| + O\big(D_N(\gamma_q(n))\big). \tag{4}$$

By using sets A_i and A'_i, $i = 1, 2, 3$, described in Fig. 2 we differ, for local discrepancy $|t(I)| - |I|$, $I = [0, x) \times [0, y)$, $(x, y) \in \bigcup_{i=1}^{3} A_i \cup \bigcup_{i=1}^{3} A'_i$, the following cases:

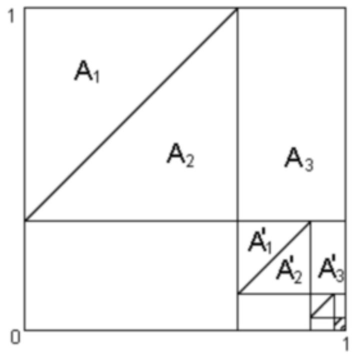

FIGURE 2. Sets A_i and A'_i, $i = 1, 2, 3$

$$|t([0,x) \times [0,y))| - xy = \begin{cases} x - xy & \text{for } (x,y) \in A_1, \\ y - (1/q) - xy & \text{for } (x,y) \in A_2, \\ y + x - 1 - xy & \text{for } (x,y) \in A_3. \end{cases} \tag{5}$$

Directly by computation we find that

$$\max_{(x,y)\in A_i} ||t([0,x) \times [0,y))| - xy| = \begin{cases} \frac{1}{4}\left(1 - \frac{1}{q}\right)^2 & \text{for } i = 1, \\ \max\left(\frac{1}{4}\left(1 - \frac{1}{q}\right)^2, \frac{1}{q}\left(1 - \frac{1}{q}\right)\right) & \text{for } i = 2, \\ \frac{1}{q}\left(1 - \frac{1}{q}\right) & \text{for } i = 3. \end{cases}$$

In the case $(x, y) \in \bigcup_{i=1}^{3} A'_i$ we transform $[1 - (1/q), 1] \times [0, 1/q] \to [0, 1]^2$ such that

$$\begin{aligned} x' &= q(x - (1 - (1/q))), \\ y' &= qy, \\ |t(I')| &= q|t(I)|, \qquad \text{where } I' = [0, x'] \times [0, y'] \text{ and } (x', y') \in \bigcup_{i=1}^{3} A_i, \end{aligned}$$

which gives

$$|t(I)| - xy = \frac{|t(I')|}{q} - \frac{x'y'}{q^2} - y\left(1 - \frac{1}{q}\right).$$

Applying (5) we find again $||t(I)| - xy| \le \frac{1}{q}\left(1 - \frac{1}{q}\right)$.

Finally, for $(x,y) \in [1-(1/q^2),1] \times [0,1/q^2]$ we have either $||t(I)|-xy| \le |t(I)| \le 1/q^2$ or $||t(I)|-xy| \le xy \le 1/q^2$. $\square$

Theorem 2. *The extremal discrepancy D_N of the sequence $\bigl(\gamma_q(n),\gamma_q(n+1)\bigr)$, $n=0,1,2,\dots,N-1$, satisfies*

$$D_N\bigl((\gamma_q(n),\gamma_q(n+1))\bigr) = \frac{1}{4} + O\bigl(D_N(\gamma_q(n))\bigr).$$

Proof. As in (4), for extremal discrepancy D_N of $(\gamma_q(n),\gamma_q(n+1))$ we have

$$D_N = \sup_{I\subset[0,1]^2} ||t(I)|-|I|| + O\bigl(D_N(\gamma_q(n))\bigr).$$

To compute $||t(I)|-|I||$ for interval $I\subset[0,1]^2$ we distinguish two cases: $|t(I)|\le|I|$ or $|t(I)|\ge|I|$. Note that if

(i) I contains a line segment $T\cap I$ having vertexes in opposite sides, then $|t(I)|\ge|I|$.

1. $|t(I)|\le|I|$. In this case, in I we can leave out some subintervals of type (i) (see Fig. 3) such that the resulting interval $I_1\subset I$ satisfies:

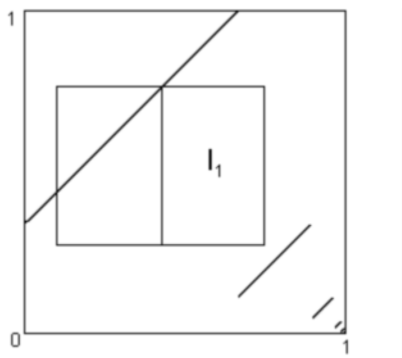

Figure 3. Case 1

- $T\cap I_1=\emptyset$, and
- $|I|-|t(I)|\le|I_1|$.

By Fig. 4, to compute maximum of $|I|$ for $I\cap T=\emptyset$ we differ three cases: $I=I_1$, $I=I_2$ and $I=I_3$ and we have

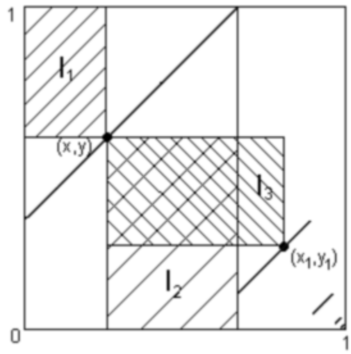

FIGURE 4. Intervals I_i, $i = 1, 2, 3$

$$\max |I_1| = \max_{x \in \left[0, 1-\frac{1}{q}\right]} \left(x \left(1 - \left(x + \frac{1}{q} \right) \right) \right) = \left[\frac{1}{2} \left(1 - \frac{1}{q} \right) \right]^2,$$

$$\max |I_2| = \max_{x \in \left[0, 1-\frac{1}{q}\right]} \left(\left(1 - \frac{1}{q} - x \right) \left(x + \frac{1}{q} \right) \right) = \frac{1}{4},$$

$$\max |I_3| = \max_{x_1 - x \in \left[0, 1-\frac{1}{q^2}\right]} \left[-\left(x_1 - x\right)^2 + (x_1 - x) \left(1 - \frac{1}{q^2} \right) \right] = \left[\frac{1}{2} \left(1 - \frac{1}{q^2} \right) \right]^2.$$

2. $|t(I)| \geq |I|$. We shall investigate two subcases:

2a. Assume that the intersection $T \cap I$ contains only one line segment. Then (see Fig. 5) there exists a subinterval $I_1 \subset I$ such that

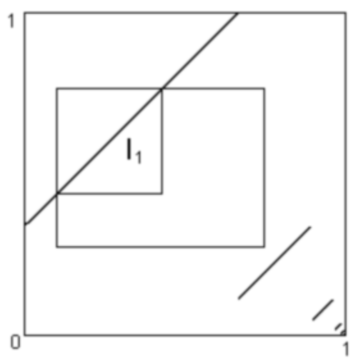

FIGURE 5. Case 2a

- $T \cap I_1 = T \cap I$ and thus
- $|t(I)| - |I| \leq |t(I_1)| - |I_1| = |t(I_1)| - |t(I_1)|^2$.

In this case the maximum is

$$\max\bigl(|t(I_1)| - |t(I_1)|^2\bigr) = \max_{x\in[0,1-(1/q)]}(x - x^2) = \frac{1}{4}.$$

2b. Assume that $T \cap I$ contains two or more line segments, e.g. $T \cap I \subset T_0 \cup T_1 \cup T_2$. By Fig. 6 we have

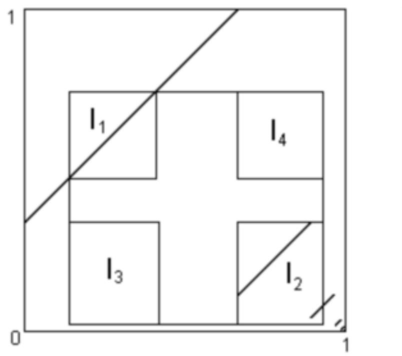

FIGURE 6. Case 2b

$$\begin{aligned}
&|I_1| = |t_1(I_1)|^2,\\
&|I_2| \geq |t(I_2)|^2,\\
&|I_3| \geq |t_1(I_1)||t(I_2)|,\\
&|I_4| \geq |t_1(I_1)||t(I_2)|,
\end{aligned}$$

which gives

$$\begin{aligned}
&\sup_{I\subset[0,1]^2} (|t(I)| - |I|)\\
&\leq \sup\bigl(|t_1(I_1)| + |t(I_2)| - |I_1| - |I_2| - |I_3| - |I_4|\bigr)\\
&\leq \sup\bigl(|t_1(I_1)| - |t_1(I_1)|^2 + |t(I_2)| - |t(I_2)|^2 - 2|t_1(I_1)||t(I_2)|\bigr)\\
&= \max\bigl((|t_1(I_1)| + |t(I_2)|) - (|t_1(I_1)| + |t(I_2)|)^2\bigr)\\
&= \max_{z\in[0,1]}(z - z^2) = \frac{1}{4}.
\end{aligned}$$

□

Note that for every two-dimensional sequence (x_n, y_n), $n = 1, 2, \dots, N$, the extremal D_N and star D_N^* discrepancies are mutually related by ([KN, p. 93])

$$D_N^* \leq D_N \leq 4D_N^*$$

for every $N = 1, 2, \dots$ and the discrepancies from Theorems 1 and 2 satisfy this relation.

REFERENCES

[DT] DRMOTA, M.—TICHY, R. F.: *Sequences, Discrepancies and Applications*. Lecture Notes in Math. 1651, Springer-Verlag, Berlin-Heidelberg, 1997.

[K] KNUTH, D. E.: *Seminumerical Algorithms. The Art of Computer Programming, Vol 2* (2nd ed.), Addison Wesley, Reading, MA, 1981.

[KN] KUIPERS, L.—NIEDERREITER, H.: *Uniform Distribution of Sequences*, John Wiley & Sons, New York, 1974; Reprint edition Dover Publications, Inc. Mineola, New York, 2006.

[So] SOBOL', I. M.: *Multidimensional Quadrature Formulas and Haar Functions*, Izd. Nauka, Moscow, 1969 (Russian).

[SP] STRAUCH, O.—PORUBSKÝ, Š.: *Distribution of Sequences: A Sampler*, Peter Lang, Frankfurt am Main, 2005.

Received 2. 10. 2007

Mathematical Institute
Slovak Academy of Sciences
Štefánikova 49
SK–814 73 Bratislava
SLOVAKIA
E-mail: olgablazekova@gmail.com