$s(N)$-UNIFORM DISTRIBUTION MODULO 1

Michael Drmota and Reinhard Winkler

Technical University of Vienna

Abstract. In this paper we study the notion of $s(N)$-uniform distribution of sequences modulo 1 which sharpens resp. quantifies the notion of complete uniform distribution. A trivial necessary condition for the existence of $s(N)$-u.d. sequences is $s(N) = o(N)$. On the other hand $s(N) = o(\sqrt{N/\log N})$ implies that almost all sequences are $s(N)$-u.d. (It seems to be very difficult to fill the gap between those two results.) By a modification of the metric methods it is shown how to construct a $s(N)$-u.d. sequence if $s(N) = o(\sqrt{N/\log N})$. This construction is rather involved. So a final section discusses some possibilities of simpler approaches for constructing $s(N)$-u.d. sequences.

1. Introduction

In [FKT] and [Gri] distribution properties of random sequences on finite sets are discussed. The purpose of this paper is to consider similar questions for random sequences $\mathbf{x} = (x_1, x_2, \dots)$ in the unit interval $x_n \in [0,1)$. It will turn out that the results to be obtained are completely different from those for sequences on finite sets. One reason for this phenomenon is surely the unavoidable difference between the two definitions of discrepancy.

In order to formulate our results we will introduce the following notations:
For $s \in \mathbf{N}$ let $\mathcal{Q}^{(s)}$ be the set of all generalized rectangles $Q \subseteq [0,1)^s$. This means that

$$Q = \prod_{i=1}^{s} [a_i, b_i), \qquad a_i, b_i \in [0,1],$$

where $[a,b) = [a,1) \cup [0,b)$ for $0 \le b < a \le 1$.
Let $M \in \mathbf{N}$. Then the set of all rectangles $Q = \prod [a_i, b_i) \in \mathcal{Q}^{(s)}$ where all a_i, b_i are rational numbers of the kind k/M ($k \in \mathbf{Z}$) will be denoted as $\mathcal{Q}_M^{(s)}$.
For any sequence $\mathbf{x} = (x_n)_{n \ge 1}$, $x_n \in [0,1)$, set $\mathbf{x}^{(s)} = (x_n^{(s)})_{n\ge 1}$, where $x_n^{(s)} = (x_n, x_{n+1}, \dots, x_{n+s-1}) \in [0,1)^s$, and define

$$\Delta(N, s, \mathbf{x}, Q) = \left| \frac{1}{N} |\{ n \le N \mid x_n^{(s)} \in Q \}| - \lambda^{(s)}(Q) \right|,$$

where $\lambda^{(s)}(Q)$ denotes the s-dimensional Lebesgue measure of the rectangle $Q \in \mathcal{Q}^{(s)}$.
Finally we set

$$\mathbf{x}_{N_1} = (x_{N_1+1}, x_{N_1+2}, \dots)$$

for an infinite subsequence of a given sequence $\mathbf{x} = (x_n)_{n \ge 1}$.

We are mainly interested in the so-called s-discrepancy

$$D(N,s,\mathbf{x}) = \sup_{Q\in\mathcal{Q}^{(s)}} \Delta(N,s,\mathbf{x},Q),$$

where $s = s(N)$ varies in $\mathbf{N}$, especially where $s(N)$ increases monotonically and is unbounded.

A sequence $\mathbf{x}$ is called $s(N)$-uniformly distributed ($s(N)$-u.d.) if

$$\lim_{N\to\infty} D(N,s(N),\mathbf{x}) = 0.$$

This concept was introduced in [Ti] (where the results of [NT] are sharpened.) Since $D(N,s_1,\mathbf{x}) \le D(N,s_2,\mathbf{x})$ for $s_1 \le s_2$ (see Lemma 1) it is a proper question to ask how fast $s(N)$ may increase such that $s(N)$-u.d. sequences still exist. For finite sets this question is completely solved from a measure theoretic point of view in [FKT], [Gri], and [Go2]. Moreover explicit examples are given in [Wi]. However, in finite sets a different notion for the discrepancy is used. For further results see [Gra].

It should be noted that a sequence $\mathbf{x}$ is called completely uniformly distributed if $\lim\limits_{N\to\infty} D(N,s,\mathbf{x}) = 0$ holds for all $s\ge 1$. Hence every $s(N)$-u.d. sequence is completely uniformly distributed if $s(N)\to\infty$.

In the following we will prove that almost all sequences $\mathbf{x}$ are $s(N)$-u.d. if $s(N) = o(\sqrt{N/\log N})$. Furthermore we will show how it is possible to find explicit examples by in principle finite algorithms using this metric result. The method to be described is rather involved. Therefore we have appended some additional considerations how more efficient algorithms could work.

It seems to be very difficult to decide whether the order of magnitude $s(N) = o(\sqrt{N/\log N})$ is best possible or not. Since

$$D(N,s,\mathbf{x}) \ge 1 - \left(1-\frac{1}{N}\right)^s$$

for any sequence $\mathbf{x}$, the existence of an $s(N)$-u.d. sequence implies $s(N) = o(N)$. But it may be expected that this trivial bound is far away from optimality.

2. Preparatory Lemmata

First we list some properties of the discrepancy:

Lemma 1.

a)

$$D(N,s,\mathbf{x}) \le D(N,s+1,\mathbf{x})$$

b)

$$D(N_2,s,\mathbf{x}) \le \max\left(D(N_1,s,\mathbf{x}), D(N_2-N_1,s,\mathbf{x}_{N_1})\right) \quad (N_1\le N_2)$$

c)

$$D(N_2,s,\mathbf{x}) \le \frac{N_1}{N_2} D(N_1,s,\mathbf{x}) + \frac{|N_2-N_1|}{N_2}$$

d)

$$D(N,s,\mathbf{x}) \le 3\Delta_M + \frac{4s}{M},$$

where

$$\Delta_M = \max_{Q \in \mathcal{Q}_M^{(s)}} \Delta(N, s, \mathbf{x}, Q).$$

Proof. a), b), and c) are trivial.

d) Let $Q \in \mathcal{Q}^{(s)}$. Clearly there exist $Q_1, Q_2 \in \mathcal{Q}_M^{(s)}$ with $Q_1 \subseteq Q \subseteq Q_2$ of the form

$$Q_j = \prod_{i=1}^{s} \left[\frac{k_{j,i}}{M}, \frac{l_{j,i}}{M}\right) \qquad (j = 1, 2)$$

such that $k_{2,i} \equiv k_{1,i} - 1 \mod M$ and $l_{2,i} \equiv l_{1,i} + 1 \mod M$. Now set

$$\begin{aligned} S &= \frac{1}{N} \sum_{n=1}^{N} 1_Q(x_n), & S_j &= \frac{1}{N} \sum_{n=1}^{N} 1_{Q_j}(x_n) \quad (j = 1, 2) \\ \lambda &= \lambda^{(s)}(Q), & \lambda_j &= \lambda^{(s)}(Q_j) \quad (j = 1, 2), \end{aligned}$$

where 1_A denotes the characteristic function of a set A. By definition we have

$$|S - S_1| \le |S_2 - S_1| \le (\lambda_2 - \lambda_1) + 2\Delta_M,$$

and

$$\lambda_2 - \lambda_1 \le 1 - \left(1 - \frac{2}{M}\right)^s \le \frac{2s}{M}.$$

Hence we get

$$\begin{aligned} \Delta(N, s, \mathbf{x}, Q) &= |S - \lambda| \\ &\le |S - S_1| + |S_1 - \lambda_1| + |\lambda_1 - \lambda| \\ &\le 3\Delta_M + \frac{4s}{M}. \end{aligned}$$

□

A quite useful tool is the following proposition which is an easy consequence of [KN, Lemma 4.1, p 208].

Proposition 2. *Let $X_1, \ldots, X_N$ be independent random variables satisfying $|X_i| \le 1$ and $\mathbf{E}X_i = 0$. Then we have*

$$\mathbf{P}\left\{\left|\frac{1}{N}\sum_{i=1}^{N} X_i\right| \ge \eta\right\} \le 2e^{-N\eta^2/2}.$$

Lemma 3. *Assume that $s \ge 1$, $M \ge 1$, and $\eta > 0$ are given. Then we have*

$$\begin{aligned} &\mathbf{P}\left(\left\{\mathbf{x} \,\middle|\, \Delta(N, s, \mathbf{x}, Q) \ge \eta + \frac{s}{N} \text{ for some } Q \in \mathcal{Q}_M^{(s)}\right\}\right) \\ &\quad \le 2M^{2s} s \exp\left(-\frac{N\eta^2}{2s}\right). \end{aligned}$$

Proof. Consider the random variables $X_{l,i} = 1_Q(x_{l+is}^{(s)}) - \lambda^{(s)}(Q)$. Notice that for each fixed $l = 1, \dots, s$ the random variables $X_{l,i}$, $i \in \mathbf{N}_0$, are independent and satisfy $\mathbf{E}X_{l,i} = 0$ and $|X_{l,i}| \le 1$. Thus we get by Proposition 2

$$\begin{aligned}
\mathbf{P}&\left\{\left|\frac{1}{N}\sum_{n=1}^{N} 1_Q(x_n^{(s)}) - \lambda^{(s)}(Q)\right| \ge \eta + \frac{s}{N}\right\} \\
&\le \mathbf{P}\left\{\left|\frac{1}{N}\sum_{n=1}^{N'} 1_Q(x_n^{(s)}) - \lambda^{(s)}(Q)\right| \ge \eta\right\} \\
&\le \sum_{l=1}^{s(N)} \mathbf{P}\left\{\left|\frac{1}{N'/s}\sum_{i=0}^{N'/s-1} X_{l,i}\right| \ge \frac{N}{N'}\eta\right\} \\
&\le 2s \exp\left(-\frac{N\eta^2}{2s}\right),
\end{aligned}$$

where $N' = \left[\frac{N}{s}\right] s$. This proves Lemma 3 since $\mathcal{Q}_M^{(s)}$ consists of at most M^{2s} rectangles. $\square$

3. A Metric Theorem

Theorem 4. *Suppose that*

$$s(N) = o\left(\sqrt{\frac{N}{\log N}}\right) \qquad (N \to \infty).$$

Then almost all sequences $\mathbf{x}$ *are* $s(N)$*-u.d.*

Proof. In view of the Borel-Cantelli Lemma, Lemma 3, and Lemma 1.d it suffices to show that there exists a sequence $M(N)$ satisfying $s(N) = o(M(N))$ $(N \to \infty)$ such that the series

$$\sum_{N=1}^{\infty} a_N \tag{3.1}$$

converges for all $\eta > 0$, where

$$a_N = M(N)^{2s(N)} s(N) \exp\left(-\frac{N\eta^2}{2s(N)}\right).$$

Set $M(N) = [\sqrt{N}]$. Then by assumption $s(N) = o(M(N))$ and

$$\begin{aligned}
\log a_N &= 2s(N)\log M(N) + \log s(N) - \frac{N\eta^2}{2s(N)} \\
&\le -\sqrt{N\log N} \le -2\log N
\end{aligned}$$

for sufficiently large N. So we are done. $\square$

4. Construction of an $s(N)$-u.d. sequence

In this section we are going to describe a construction of explicit examples of $s(N)$-u.d. sequences provided that $s(N) = o(\sqrt{N/\log N})$ is explicitly given. This means that, if for every $\varepsilon > 0$ there is an explicitly computable N_ε such that $s(N)\sqrt{\log N/N} < \varepsilon$ for all $N \geq N_\varepsilon$, then there is an $s(N)$-u.d. sequence $\mathbf{x}$ such that for each $N_0 \geq 1$ the elements $x_1, \dots, x_{N_0}$ can be calculated in finitely many steps.

Let us start by listing some useful observations.

Lemma 5. *Let $s(N) = o(\sqrt{N/\log N})$ be explicitly given. Then there exists an explicitly computable function $f(N) = o(1)$, $(N \to \infty)$ such that there exists a sequence $\mathbf{x}$ with*

$$D(N, s(N), \mathbf{x}) \leq f(N) \qquad \textit{for all } N.$$

Proof. Inspecting Lemma 3 ($M(N) = [\sqrt{N}]$) it turns out that it is possible to find some sequence $\eta(N) = o(1)$ $(N \to \infty)$ such that the series (3.1) in the proof of Theorem 4 still converges. Furthermore we can compute some $N_0 \geq 1$ such that

$$\sum_{n=N_0}^{\infty} a_N < \frac{1}{2}.$$

By Lemma 3 this means that there exists a sequence $\mathbf{x}$ satisfying

$$\Delta(N, s(N), \mathbf{x}, Q) < \eta(N) + \frac{s(N)}{N}$$

for all $N \geq N_0$ and all rectangles $Q \in \mathcal{Q}_{M(N)}^{(s(N))}$. Now set $f(N) = 1$ for $N < N_0$ and

$$f(N) = 3\left(\eta(N) + \frac{s(N)}{N}\right) + 4\frac{s(N)}{M(N)}$$

for $N \geq N_0$ which completes the proof by using Lemma 1.d. $\square$

In order to simplify our construction we will use a rounded sequence $\overline{\mathbf{x}} = (\overline{x}_n)_{n\geq 1}$ of a general sequence $\mathbf{x} = (x_n)_{n\geq 1}$, $x_n \in [0,1)$, where $\overline{x}_n = \frac{k}{n}$ $(k = 0, \dots, n-1)$ is the nearest rational number (modulo 1) with denominator n. For such a sequence we have

Lemma 6. *For $N \geq 4$ and $s \geq 2$ we have*

$$D(N, s, \overline{\mathbf{x}}) \leq 2\, D(N, s, \mathbf{x}) + \frac{4s}{[\sqrt{N}]}.$$

Proof. As in Lemma 1.c we have

$$D(N, s, \overline{\mathbf{x}}) \leq D(N - [\sqrt{N}], s, \overline{\mathbf{x}}_{[\sqrt{N}]}) + \frac{[\sqrt{N}]}{N}.$$

For any rectangle $Q = \prod_{i=1}^{s}[a_i, b_i) \in \mathcal{Q}^{(s)}$ consider $Q_1 = \prod\limits_{i=1}^{s} I_i$, where

$$I_i = \left[a_i + \frac{1}{[\sqrt{N}]}, b_i - \frac{1}{[\sqrt{N}]}\right) \qquad (\text{or } I_i = \emptyset)$$

and $Q_2 = \prod_{i=1}^{s} J_i$, where

$$J_i = \left[a_i - \frac{1}{[\sqrt{N}]}, b_i + \frac{1}{[\sqrt{N}]}\right) \qquad (\text{or } J_i = [0,1))$$

such that $Q_1 \subseteq Q \subseteq Q_2$ and

$$\lambda^{(s)}(Q_2) - \lambda^{(s)}(Q_1) \le 1 - \left(1 - \frac{2}{[\sqrt{N}]}\right)^s \le \frac{2s}{[\sqrt{N}]}.$$

So we get

$$\begin{aligned}
\Delta(N - [\sqrt{N}], s, \overline{\mathbf{x}}_{[\sqrt{N}]}, Q) &\le \max\Big(\Delta(N - [\sqrt{N}], s, \mathbf{x}_{[\sqrt{N}]}, Q_1) + \frac{2s}{[\sqrt{N}]}, \\
&\qquad \Delta(N - [\sqrt{N}], s, \mathbf{x}_{[\sqrt{N}]}, Q_2) + \frac{2s}{[\sqrt{N}]}\Big) \\
&\le \frac{N}{N - [\sqrt{N}]} D(N, s, \mathbf{x}) + \frac{2s}{[\sqrt{N}]} + \frac{[\sqrt{N}]}{N - [\sqrt{N}]} \\
&\le 2\, D(N, s, \mathbf{x}) + \frac{3s}{[\sqrt{N}]}.
\end{aligned}$$

□

Corollary 7. *For any $N_0 \ge 1$ it is possible to find a finite sequence $\overline{\mathbf{x}} = (\overline{x}_1, \dots, \overline{x}_{N_0 + s(N_0) - 1})$ satisfying*

$$D(N, s(N), \overline{\mathbf{x}}) \le g(N) = 2\, f(N) + \frac{4s}{[\sqrt{N}]} \qquad (N \le N_0) \tag{4.1}$$

in finitely many steps.

Proof. Combining Lemmata 5 and 6 it follows that there exists a rounded sequence $\overline{\mathbf{x}} = (\overline{x}_n)_{n \ge 1}$ with $\overline{x}_n = k_n/n$ ($k_n \in \mathbf{Z}$) satisfying $D(N, s(N), \overline{\mathbf{x}}) \le g(N)$ for all N. Since there are only finitely many rounded sequences of fixed length it suffices to prove that the discrepancy can be computed in finitely many steps. For this purpose observe that $\sup \Delta(N, s, \mathbf{x}, Q)$ is attained formally for one of the finitely many open or closed rectangles

$$Q = \prod_{i=1}^{s} (a_i, b_i) \quad \text{or} \quad Q = \prod_{i=1}^{s} [a_i, b_i],$$

where a_i, b_i are members of the sequence. □

The only disadvantage of Corollary 7 is that it does not provide a method to continue a finite sequence $\overline{x}_1, \dots, \overline{x}_{N_0 + s(N_0) - 1}$ satisfying (4.1) to an $s(N)$-u.d. sequence. In the sequel we will show how to avoid this disadvantage.

Note that it is no loss of generality to assume that the given sequence $s(N)$ is monotonically increasing.

Lemma 8. *Let $s(N) = \varphi(N)\sqrt{N/\log N}$ be explicitely given such that $s(N)$ is monotonically increasing to $+\infty$ and $\varphi(N) = o(1)$ $(N \to \infty)$. Furthermore let N_k be a monotonically increasing sequence such that $\varphi(N) \leq \frac{1}{k}$ for $N \geq N_k$. Set $S(N) = s(\psi(N)N)$, where $\psi(N) = k$ for $N = N_k + 1, \dots, N_{k+1}$ $(N_0 = 0)$. Then we have*

$$S(N) = o\left(\sqrt{\frac{N}{\log N}}\right) \qquad (N \to \infty).$$

Proof. Suppose that $N_k < N \leq N_{k+1}$. Then we get

$$\begin{aligned}\frac{S(N)}{\sqrt{\frac{N}{\log N}}} = \frac{s(kN)}{\sqrt{\frac{N}{\log N}}} &= \varphi(kN)\sqrt{\frac{kN\log N}{N\log kN}} \\ &\leq \varphi(kN)\sqrt{k} \leq \frac{1}{\sqrt{k}}.\end{aligned}$$

□

Lemma 9. *Let $s(N)$ and $S(N)$ be as in Lemma 8. Then one can construct two sequences $a_n < b_n$ of positive integers satisfying*

$$\begin{aligned} b_n &\leq a_{n+1}, \\ s(a_{n+1}) &\leq S(a_n - a_{n-1}), \\ s(b_n + r) &\leq S(b_n - a_n + r) \quad \textit{for } r \in \mathbf{N}_0, \\ \lim_{n\to\infty} b_n - a_n &= \infty, \\ \lim_{n\to\infty} \frac{b_n - a_n}{a_n} &= 0, \quad \textit{and} \\ \lim_{n\to\infty} \frac{a_{n+1}}{a_n} &= \infty. \end{aligned}$$

Proof. We start with $a_1 = 1$. Now choose $a_2 \geq 3$ such that $\psi(a_2 - a_1) \geq 6$. For $n \geq 2$ we recursively set

$$a_{n+1} = (a_n - a_{n-1})\psi(a_n - a_{n-1}).$$

Trivially this sequence satisfies $s(a_{n+1}) \leq S(a_n - a_{n-1})$. Set $\gamma_n = a_n/a_{n-1}$. Then we have

$$\gamma_{n+1} = \left(1 - \frac{1}{\gamma_n}\right)\psi(a_n - a_{n-1})$$

and

$$\delta_n = \frac{a_{n+1} - a_n}{a_n - a_{n-1}} = \frac{\gamma_{n+1} - 1}{1 - \frac{1}{\gamma_n}}.$$

From $\gamma_2 \geq 3$ and $\psi(a_2 - a_1) \geq 6$ we get $\gamma_3 \geq 4$ and $\delta_2 \geq 2$. Hence we get by induction $\delta_n \geq 2$ $(n \geq 2)$ which implies $\lim_{n\to\infty}(a_n - a_{n-1}) = \infty$ and consequently $\lim_{n\to\infty}\gamma_n = \infty$. Now set

$$b_n = \min\{b \in \mathbf{N} \mid b \leq (b - a_n)\psi(b - a_n)\}.$$

Trivially we have $b_n > a_n$ and $b_n \le a_{n+1}$ since $(a_{n+1}-a_n)\psi(a_{n+1}-a_n) > a_{n+1}$. Furthermore by construction $b_n + r \le (b_n - a_n + r)\psi(b_n - a_n + r)$ which gives $s(b_n + r) \le S(b_n - a_n + r)$. Since $s(N) \ge 1$ we always have $\varphi(N) \ge \sqrt{\log N/N}$ which implies $\psi(N) \le \sqrt{N/\log N}$. Hence, from

$$b_n - a_n \ge \frac{b_n}{\psi(b_n - a_n)} \ge \frac{b_n}{\psi(b_n)} \ge \sqrt{b_n \log b_n}$$

we get $\lim\limits_{n\to\infty}(b_n - a_n) = \infty$ and finally from

$$\frac{b_n - a_n - 1}{b_n - 1} < \frac{1}{\psi(b_n - a_n - 1)}$$

it follows $\lim\limits_{n\to\infty} a_n/b_n = 1$ which is equivalent to $\lim\limits_{n\to\infty}(b_n - a_n)/a_n = 0$. $\square$

Next set $d_n = a_n - a_{n-1}$ ($a_0 = 0$) and for each $n \ge 1$ construct a finite sequence $\mathbf{z}^{(n)} = (z_1^{(n)}, \dots, z_{d_n+S(d_n)-1}^{(n)})$ of length $d_n + S(d_n) - 1$ such that

$$D(N, S(N), \mathbf{z}^{(n)}) \le g(N) \quad \text{for } N \le d_n$$

by Corollary 7. Now let $\mathbf{y}^{(n)} = (z_1^{(n)}, \dots, z_{d_n}^{(n)})$ consist of the first d_n elements of $\mathbf{z}^{(n)}$. If $\mathbf{y}^{(n)}$ is appended by an arbitrary sequence $\mathbf{w}$ of length $\ge S(d_n) - 1$ then we obviously get

$$D(N, S(N), \mathbf{y}^{(n)}\mathbf{w}) \le g(N) + 2\frac{S(N)}{N} \quad \text{for } N \le d_n.$$

We can further assume that $g(N) + 2s(N)/N \le \overline{g}(N)$, where $\overline{g}(N)$ is monotonically decreasing and satisfies $\lim\limits_{N\to\infty}\overline{g}(N) = 0$.

Theorem 10. *Let $\mathbf{x} = \mathbf{y}^{(1)}\mathbf{y}^{(2)}\mathbf{y}^{(3)}\cdots$, where $\mathbf{y}^{(n)}$ are those finite sequences constructed above. Then $\mathbf{x}$ is $s(N)$-u.d.*

Proof. In order to estimate $D(N, s(N), \mathbf{x})$ we will distinguish two cases. First assume that $a_n \le N < b_n$. Here we have

$$\begin{aligned} D(N, s(N), \mathbf{x}) &\le \frac{a_{n-1}}{N} + D(a_n - a_{n-1}, s(N), \mathbf{x}_{a_{n-1}}) + \frac{N - a_n}{N} \\ &\le \frac{a_{n-1}}{a_n} + D(a_n - a_{n-1}, S(a_n - a_{n-1}), \mathbf{x}_{a_{n-1}}) + \frac{b_n - a_n}{a_n} \\ &\le \frac{1}{\gamma_n} + \overline{g}(d_n) + \frac{b_n - a_n}{a_n}, \end{aligned}$$

where we have used $s(N) \le s(a_{n+1}) \le S(a_n - a_{n-1})$ and $\mathbf{x}_{a_{n-1}} = \mathbf{y}^{(n)}\mathbf{x}_{a_n}$.

Finally suppose that $b_n \le N < a_{n+1}$. Similarly we get

$$\begin{aligned} D(N, s(N), \mathbf{x}) &\le \frac{a_{n-1}}{N} + D(a_n - a_{n-1}, s(N), \mathbf{x}_{a_{n-1}}) \\ &\qquad + D(N - a_n, s(N), \mathbf{x}_{a_n}) \\ &\le \frac{a_{n-1}}{a_n} + D(a_n - a_{n-1}, S(a_n - a_{n-1}), \mathbf{x}_{a_{n-1}}) \\ &\qquad + D(N - a_n, S(N - a_n), \mathbf{x}_{a_n}) \\ &\le \frac{1}{\gamma_n} + \overline{g}(d_n) + \overline{g}(b_n - a_n). \end{aligned}$$

where we have used $s(N) \le S(a_n - a_{n-1})$, $s(N) = s(b_n + (N - b_n)) \le S(b_n - a_n + (N - b_n)) = S(N - a_n)$, and $N - a_n \ge b_n - a_n$.

Hence we get $\lim_{N\to\infty} D(N, s(N), \mathbf{x}) = 0$. □

5. Remarks on Simpler Constructions

The construction in section 4 is very complicated and not useful for practical computation. However, one observes other difficulties by looking for simpler constructions. This concluding section discusses some aspects.

In [Le] it is shown by using Koksma-Erdős-Turan's inequality that sequences of the form $(\{\lambda^n\})_{n\ge1}$ are completely u.d. where λ is constructed appropriately. In fact, it turns out (compare with [DTW]) that this method guarantees $s(N)$-u.d. sequences only for $s(N) = o(\log N)$. Therefore it seems that Koksma-Erdős-Turan's inequality cannot be used to find the correct order of magnitude for $s(N)$ in the sense of Theorem 4.

Another possibility of construction is the following principle (cf. [Go1]). Find finite sequences $\mathbf{b}_s$ with small s-discrepancy and consider the sequence

$$\mathbf{x} = \mathbf{b}_1^{k_1}\mathbf{b}_2^{k_2}\cdots,$$

where $k_i \in \mathbf{N}$ are chosen properly. Such principles have been used to construct best possible sequences in finite sets (see [Wi]). As stated in the introduction the notion of discrepancy on finite sets is different from that used in the modulo-1-case. There it turns out that the order of magnitude $s(N) = \mathcal{O}(\log N)$ is best possible (see [FKT], [Gri], and [Go2]) and if we try to use a similar construction in the modulo-1-case, namely that $\mathbf{b}_s$ contains all possible subblocks of the form $\left(\frac{l_1}{m}, \dots, \frac{l_s}{m}\right)$ ($l_i = 0, \dots, m-1$, $m = m(s)$), then $\mathbf{b}_s$ must contain at least m^s elements. Since we have to propose $\lim_{s\to\infty} m(s) = \infty$ it is an easy exercise to show that $s(N)$ cannot grow faster than $\log N$. The reason is that $\mathbf{b}_s$ consists of too many elements. Therefore we would like to provide finite sequences $\mathbf{b}_s$ of small s-discrepancy and small length. The simplest case would be a cyclic finite subgroup of $[0,1)^s = (\mathbf{R}/\mathbf{Z})^s$. Hence there is a connection to the theory of good lattice points (see [KN]). As the following theorem shows there are only very special sequences of this kind.

Theorem 11. *Let $\mathbf{x} = (x_1, x_2, \dots, x_L, x_{L+1} = x_1, x_{L+2} = x_2, \dots)$ be a periodic sequence with period L such that all finite subsequences $\mathbf{b}_j = (x_j, x_{j+1}, \dots, x_{j+s-1})$ ($j = 1, \dots, L$) of length s ($2 \le s \le L$) are contained in a cyclic subgroup*

$$G = \langle \mathbf{a} = (a_1, \dots, a_s)\rangle \le (\mathbf{R}/\mathbf{Z})^s$$

of order $|G| = n$, (where n is chosen to be minimal). Then there exist integers m, g with $gcd(m,n) = gcd(g,n) = 1$ such that

$$x_j \equiv \frac{m\, g^{j-1}}{n} \mod 1 \qquad \text{for } j \ge 1. \tag{5.1}$$

Conversely, sequences of the kind (5.1) induce blocks $\mathbf{b}_j$ that are contained in $G = \langle \frac{1}{n}(1, g, \dots, g^{s-1}) \mod 1\rangle$.

Proof. Set $a_i = k_i/n$ ($i = 1, \dots, s$), where $k_i \in \{0, \dots, n-1\}$ and similarly $x_j = r_j/n$ ($j = 1, \dots, L$). Furthermore consider integers m_j satisfying $\mathbf{b}_j = m_j\mathbf{a} \mod 1$. Hence we get

$$r_j \equiv m_j k_1 \equiv m_{j-1}k_2 \equiv \dots \equiv m_{j-s+1}k_s \mod n, \quad (j = 1, \dots, L).$$

Let p be an arbitrary prime divisor of n. If $p|r_j$ for all $j = 1, \ldots, L$ then $n' = n/p$ could be used instead of n in contradiction to the minimality of n. W.l.o.g. assume that p does not divide r_1. Hence p does not divide $k_1, \ldots, k_s$ and $m_1, m_L, m_{L-1}, \ldots, m_{L-s+2}$, either. Thus $s \geq 2$ implies that p does not divide $r_2 \equiv m_1 k_2 \mod n$ and so on. Hence $\gcd(r_j, n) = \gcd(m_j, n) = 1$. Set $g \equiv m_2 m_1^{-1} \equiv k_2 k_1^{-1} \mod n$ and $m = m_1 k_1$. So we get by induction $r_1 \equiv m \equiv mg^0 \mod n$ and

$$r_{j+1} \equiv m_j k_2 \equiv (m_j k_1)(k_2 k_1^{-1}) \equiv r_j g \equiv mg^j \mod n,$$

which directly gives (5.1).

The last assertion is trivial. □

Remark. With respect to uniform distribution it seems to be desirable to cover (almost) the whole group G (of Theorem 11) by finite subsequences $\mathbf{b}_s$ of length s. Hence sequences of the kind $\left(\frac{g^0}{p}, \frac{g^1}{p}, \ldots\right)$, where g is a primitive root modulo p, could play an important role. Hence it might be possible that the construction of simply computable $s(N)$-u.d. sequences (with large $s(N)$) could be tackled by considering primitive roots over finite fields.

Finally we will state a Lemma showing that proper primitive roots $g \mod p$ must be of special size.

Lemma 12. *Let p be a prime, g a primitive root $\mod p$, and m an arbitrary non-zero integer. Suppose that $g^k \equiv \pm r \mod p$ $(1 < r < p)$ for some integer k with $0 < |k| < s$. Then the s-discrepancy $D(N, s, \mathbf{x})$ $(N \geq 1)$ of the sequence $\mathbf{x} = (x_1, x_2, \ldots)$ with $x_j = mg^{j-1}/p \mod 1$ $(j \geq 1)$ is bounded below by*

$$D(N, s, \mathbf{x}) \geq \frac{1}{4r}.$$

Proof. First let $g^k \equiv +r \mod p$ and $k > 0$. Consider the rectangle $Q = \prod_{i=1}^{s} Q_i$, where $Q_1 = \left[\frac{1}{2r}, \frac{1}{r}\right)$, $Q_{k+1} = \left[0, \frac{1}{2}\right)$, and $Q_i = [0, 1)$ otherwise. Then $\lambda^{(s)}(Q) = \frac{1}{4r}$. Since $x_j \in Q_1$ if and only if $x_{j+k} \in [0,1) \setminus Q_{k+1}$, it follows that $(x_j, \ldots, x_{j+s-1}) \notin Q$ for all $j \geq 1$. Hence $D(N, s, \mathbf{x}) \geq \frac{1}{4r}$ in this case. Obviously the case $g^k \equiv -r \mod p$ can be treated similarly by considering $Q_{k+1} = \left(\frac{1}{2}, 1\right] \mod 1$. (Note that the use of intervals of the kind $(a, b]$ instead of $[a, b)$ does not effect the value of the discrepancy.)

If $k < 0$ then set $Q_s = \left[\frac{1}{2r}, \frac{1}{r}\right)$, $Q_{s-k} = \left[0, \frac{1}{2}\right)$ for the case $g^k \equiv +r \mod p$ resp. $Q_{s-k} = \left(\frac{1}{2}, 1\right] \mod 1$ for the case $g^k \equiv -r \mod p$, and $Q_i = [0, 1)$ otherwise. As above, $x_{j+s-1} \in Q_s$ if and only if $x_{j+s-k-1} \in [0,1) \setminus Q_{s-k}$ and hence $(x_j, \ldots, x_{j+s-1}) \notin Q$ for all $j \geq 1$. This completes the proof of Lemma 13. □

References

[DTW] M. Drmota, R. F. Tichy, and R. Winkler, *Completely Uniformly Distributed Sequences of Matrices*, in: Number-Theoretic Analysis (E. Hlawka and R. F. Tichy eds.), Lecture Notes in Math. (Springer, Berlin) **1452** (1990), 43–57.

[FKT] Ph. Flajolet, P. Kirschenhofer, and R. F. Tichy, *Deviations from Uniformity in Random Strings*, Probab. Th. Rel. Fields **80** (1988), 139–150.

[Go1] M. Goldstern, *Vollständige Gleichverteilung in diskreten Räumen*, in: Zahlentheoretische Analysis II (E. Hlawka ed.), Lecture Notes in Math. (Springer, Berlin) **1262** (1987), 46–49.

[Go2] M. Goldstern, *An Application of Shoenfield's Absoluteness Theorem to the Theory of Uniform Distribution*, Monatsh. Math., to appear (1993).

[Gra] P. Grabner, *Block Distribution in Random Strings*, Ann. Inst. Fourier, to appear (1993).

[Gri] K. Grill, *A Note on Randomness*, Stat. and Probab. Letters **14** (1992), 229–233.

[KN] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.

[Le] M. B. Levin, *Completely Uniform Distribution of Fractional Parts of the Exponential Function*, Trudy Sem. Petrovsk. **7** (1981), 245–256 (Russian, English translation: J. Sov. Math. **31** (1985), 3247–3256).

[NT] H. Niederreiter and R. F. Tichy, *Solution of a Problem of Knuth on Complete Uniform Distribution of Sequences*, Mathematika **32** (1985), 26–32.

[Ti] R. F. Tichy, *Ein metrischer Satz über vollständig gleichverteilte Folgen*, Acta Arithmetica **48** (1987), 197–207.

[Wi] R. Winkler, *Some Constructive Examples in Uniform Distribution of Finite Sets and Normal Numbers*, Anz. d. Österr. Akad. d. Wiss. **126** (1989), 1–8.

Department for Algebra and Discrete Mathematics, Technical University of Vienna, Wiedner Hauptstrasse 8–10/118, A-1040 Vienna, AUSTRIA