Reinhard Winkler
Some remarks on pseudorandom sequences

Mathematica Slovaca, Vol. 43 (1993), No. 4, 493--512

Persistent URL: http://dml.cz/dmlcz/129391

Math. Slovaca, 43 (1993), No. 4, 493–512

SOME REMARKS ON PSEUDORANDOM SEQUENCES

REINHARD WINKLER

(*Communicated by Robert F. Tichy*)

ABSTRACT. The paper presents some results illustrating the difficulty of giving a general definition of a *pseudorandom sequence*. It seems to indicate that rather tɾchnical distribution properties studied in the theory of uniform distribution may be the best one can expect as criterions. The method is to study different notions of *tests* defining sets of acceptance that are "natural" from a probabilistic point of view. Some results on Baire properties are added.

1 Introduction: The object of this paper

In the past several attempts have been made to give a mathematically satisfactory definition of what we are used to call randomness. Indeed investigations of von Mises, Schnorr, Martin-Löf, Lambalgen, Chaitin and many others give answers to the following problem:

Given a sequence of events (as an example think of quantum physics). Find some laws in it or, if there are none, give the diagnosis of "randomness".

According to this problem definitions are – roughly spoken – of such a kind that "randomness" increases if the shortest formulation of describing laws gets more complicated. In the ideal case a sequence cannot be described in a (much) shorter way than by listing all its members explicitly. The other way round: Every attempt to give a finite definition (a recursive function) of an infinite random sequence must fail. For an extensive treatment cf. for instance [Cha1], [Cha2], [Cha3], [Chu], [D], [Fe2], [Ko1], [Ko2], [Ko3], [vL], [L], [M1], [M2], [vM1], [vM2], [vM3], [P], [S1], [S2], [Sv], [V], [Wa1] and [Wa2] etc., where complexity theory, recursion theory, and even undecidability questions from mathematical logic are playing an important role.

AMS Subject Classification (1991): Primary 65C10. Secondary 60A05.

Key words: Pseudorandom sequence, Invariant, Terminal and Permutable one-sets of sequences, Randomness and Distribution properties.

In this paper we do not ask under which conditions an externally given sequence may be regarded as a random sequence. We are looking for realistic (weaker) conditions for a *pseudorandom sequence* which can be constructed explicitly. The different problem is clear: We want sequences that can be computed in an easy way (that means they are just the opposite of random in the sense described above) but still have some properties we intend to accept as "random-like".

In fact most requested features of pseudorandom sequences are certain distribution properties (cf. the textbook [NS] or [Lev-Sh-So-Tu], [N1], [N2], [N3], [N4], [So1], [So2], [Ta], [Te1], [Te2], [Ti2] etc.). By reasons of better computability up to now random generators are mainly using sequences that have good "average properties" for a long time but then are periodic. For many applications these sequences are absolutely sufficient if the length of the period is long enough with respect to the given problem. But there are fundamental bounds which get obvious as soon as a given sequence is used too long.

According to considerations that can be found for instance in [Kn] or [Ti2] we believe that the notion of completely uniform distribution and its modification $s(N)$-uniform distribution (cf. section 2) is a very useful tool to express some essential aspects. For a general treatment of the theory cf. the textbooks [Ku-N] or [H]. Some results which are of particular interest in this context are contained in [Dr-Ti-Wi], [Fl-Ki-Ti1], [Gra], [Ki-Ti], [Le1], [Le2], [N-Ti], [Ti1] and in some further papers mentioned in more detail in section 2.

2 Motivation: Some recent results on uniform distribution

In this section we present some results as a motivation for everything that follows. It seems useful to distinguish two cases: The case of discrete uniform distribution on a finite set M and continuously uniformly distributed random numbers in the unit interval $[0,1)$. It is well known that all continuous probability distributions on the set $\mathbb{R}$ of the reals which are Borel measures can be transformed to this case.

Let us start with a finite set M of cardinality $|M| = m$ and the probability measure μ on it, defined by $\mu(N) = \frac{|N|}{m}$ for all subsets $N \subseteq M$ of M. The strong law of large numbers guarantees that for every $a \in M$ and almost all sequences $\mathbf{x} = (x_n)_{n \in \mathrm{N}} \in X$ (w.r.t. the measure P induced by μ on the set X of all sequences on M) the number

$$A(a, N, \mathbf{x}) := \operatorname{card}\{n \leq N \mid x_n = a\}$$

of occurrences of a among the first N members of $\mathbf{x}$ asymptotically equals

$\frac{N}{m}$. This means that the so-called discrepancy

$$D(1,N,\mathbf{x}) := m \cdot \max_{a\in M}\left|\frac{A(a,N,\mathbf{x})}{N} - \frac{1}{m}\right|$$

converges to 0 with probability 1. Sequences $\mathbf{x}$ having this property are called *uniformly distributed* (u.d.) on M. In this notion the measure μ of discrete uniform distribution on M is taken into account but it does not consider the essential property of randomness, i.e. that different events do not have any influence on each other. This can be done by considering for some $s \in \mathbb{N}$ the sequence

$$\mathbf{x}^{(s)} := \left(x_n^{(s)}\right)_{n\in\mathbb{N}}$$

of all s-tuples

$$x_n^{(s)} := (x_n, x_{n+1}, \dots, x_{n+s-1})$$

and requesting its uniform distribution on the set M^s. This means that the s-discrepancy

$$D(s,N,\mathbf{x}) := D\left(1,N,\mathbf{x}^{(s)}\right) = m^s \cdot \max_{a\in M^s}\left|\frac{A\left(a,N,\mathbf{x}^{(s)}\right)}{N} - \frac{1}{m^s}\right|$$

also converges to 0 for $N \to \infty$. Then the original sequence $\mathbf{x}$ is called *s-uniformly distributed.* It is called *completely uniformly distributed* if this is the case for all $s \in \mathbb{N}$. The reason for the factor m^s is that it guarantees the inequality

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

that gives rise to an even more restrictive notion:

Let now $s = s(N)$ not be a fixed number but a sequence of nonnegative integers depending on N. $\mathbf{x}$ is called *$s(N)$-uniformly distributed* if

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

Of course we are interested in $s(N)$-u.d. sequences for an s that is tending to infinity as fast as possible. The following result of Flajolet, Kirschenhofer, Tichy and Grill, cf. [Fl-Ki-Ti2] and [Gri], gives a very precise answer which speed of growth of the sequence $s(N)$ is realistic. (In fact the proof is done for $m = 2$ but can be easily generalized.)

PROPOSITION 1. *Let* $\log_m$ *denote the logarithm to the base* m *and*

$$\varphi(N) := \log_m N - \log_m \log_m N - s(N) .$$

In the case $\varphi(N) \to \infty$ *for* $N \to \infty$ *almost all sequences are* $s(N)$*-uniformly distributed, in the other case almost all are not.*

An explicit construction of a sequence that is $s(N)$-u.d. for a given $s(N)$ with $\varphi(N) \to \infty$ as in the proposition can be found in [Wi].

Looking for corresponding notions for sequences of continuously distributed random numbers in the unit interval $[0,1)$ one gets lead to the following definition of discrepancy:

$$D(s,N,\mathbf{x}) := \sup_{Q \in \mathcal{Q}^s} \left| \frac{A(Q,N,\mathbf{x}^{(s)})}{N} - \lambda(Q) \right| .$$

$\mathcal{Q}^s$ denotes the set of all *generalized rectangles*

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

where $[a_i, b_i) = [0, b_i) \cup [a_i, 1)$ in the case $b_i < a_i$. The abbreviations

$$A(Q,N,\mathbf{x}^{(s)}) = \operatorname{card}\{n \leq N \mid x_n^{(s)} \in Q\}$$

and λ for the Lebesgue measure are clear. In [Dr-Wi] one finds a proof for

PROPOSITION 2. *For*

$$s(N) = o\left(\sqrt{\frac{N}{\log N}}\right)$$

almost all sequences on $[0,1)$ *are* $s(N)$*-u.d. and an explicit construction of such a sequence is in principle possible.*

However the methods up to now are by far too crude for applications. Furthermore it is not known whether the metric result is best possible as it is in the discrete case, cf. Proposition 1. It is only known that $s(N) = o(N)$ is a necessary condition for the existence of $s(N)$-u.d. sequences.

In every case a measure of quality of pseudorandom sequences is given which should be approximated by constructions in the future. But for the moment this seems to be rather difficult. Some remarks on this problem are also given in [Dr-Wi].

3. Problems: A strange situation

On one hand the results mentioned in section 2 seem to give guidelines what one may and should require of a good pseudorandom sequence. On the other hand one realizes a certain kind of difficulty which can be explained in a simple form by the following problem left open by Proposition 2 (notations as there):

Let Y_s denote the set of all $s = s(N)$-u.d. sequences. Then, by Proposition 2, we have $P(Y_s) = 1$ if

$$\varphi(N) := s(N)^{-1}\sqrt{\frac{N}{\log N}} \to \infty .$$

If $s_1(N) \leq s_2(N)$ for all $N \in \mathbb{N}$ we have $Y_{s_2} \subseteq Y_{s_1}$. This is an immediate consequence of $D(s, N, \mathrm{x}) \leq D(s+1, N, \mathrm{x})$. Nevertheless a priori it is not trivial to conclude that the set

$$Y := \bigcap \{Y_s \mid \varphi(N) \to \infty\}$$

is a one-set too because the intersection must be taken over an uncountable set. A recent result of M. Goldstern, cf. [Gl], shows by using some set theory that indeed $P(Y) = 1$ (the analogue holds for Proposition 1 instead of Proposition 2). But some results of this paper will show that in several general situations the phenomenon does occur that the intersection of a “natural” family of one-sets is empty. They may be regarded as generalizations of the fact that there is no sequence satisfying every condition that is fulfilled by almost all sequences.

Considering the argument that $s(N)$-u.d. is a rather technical condition one may feel the wish to find a “natural” set of properties that can be regarded as a definition of a pseudorandom sequence. If one keeps in mind the preceding it is very natural to look at those properties that can be “tested” in a certain way like the convergence of the discrepancy of a sequence may be regarded as a test for its uniform distribution. The following sections are containing investigations whether more general approaches than that of uniform distribution (cf. section 2) can give rise to notions that are even more satisfactory. The results seem to indicate that this is not the case and that one cannot get something that is essentially better than the several notions of uniform distribution. Nevertheless the subsequent considerations are far away from being complete or being the only possible way to go. Therefore this paper may also be understood as an idea that could (maybe after some modifications that are only vaguely perceivable at the moment) still lead to surprising and interesting views.

4 Notions: A hierarchy of tests

The notion of discrepancy introduced in section 2 can be considered as a test selecting certain features of "randomness". We are using this idea to give a more general definition of a test. The notations will be used throughout the paper.

Definition 1. *Let $(M, \mathcal{A}, \mu)$ be a probability space, $\left(X, \mathcal{A}^{(\omega)}, P\right)$, $X := M^\omega$, the induced probability space of all infinite sequences $\mathbf{x} = (x_n)_{n \in \mathrm{N}}$, $\mathbf{y} = (y_n)_{n \in \mathrm{N}}$, etc., on M and*

$$M' := \bigcup_{n \in \mathrm{N}} M^n$$

the set of all finite sequences. A test is a function $t\colon M' \to [0,1]$ whose restriction to each M^n is measurable with respect to the product σ-algebra $\mathcal{A}^{(n)}$ on M^n. If t takes only the values 0, 1 we call it a discrete test. Mostly we will write

$$t_n(\mathbf{x}) = t_n(x_1, \ldots, x_n, \ldots) := t(x_1, \ldots, x_n)$$

and

$$t = (t_n)_{n \in \mathrm{N}} .$$

Every test defines the sets

$$X_{t,a} := \left\{\mathbf{x} \mid \lim_{n \to \infty} t_n(\mathbf{x}) = a\right\}$$

of convergence to $a \in [0,1]$. t is called test for $X_{t,1}$ and refutation for $X \backslash X_{t,1}$. It is called strict if

$$X_{t,0} \cup X_{t,1} = X .$$

A set $A \subseteq X$ is called (*strictly, discretely*) *testable* resp. *refutable* (*by* t) if there is a (strict, discrete) test resp. refutation t for A.

We are going to study sets $A \subseteq X$ which are characterized by the property to be accepted (or refuted) by some natural classes of tests (or refutations).

To get a better view over all these notions we state some important relations between them. Some of them are not proved immediately but are corollaries of results of later sections. Note that there are indeed no circular arguments.

Theorem 1.

1. *A is strictly testable $\Longleftrightarrow$ A is discretely strictly testable.*
2. *A is strictly testable $\Longrightarrow$ A is discretely testable $\Longrightarrow$ A is testable.*
3. *A is strictly testable $\Longrightarrow$ A is discretely refutable $\Longrightarrow$ A is refutable.*
4. *A is strictly testable $\Longleftrightarrow$ A is discretely testable and discretely refutable.*

5. *A is (discretely) testable (refutable) $\iff$ $X \setminus A$ is (discretely) refutable (testable).*
6. *A is discretely testable (refutable) $\Longrightarrow$ A is refutable (testable).*
7. *There are 1-sets that are discretely testable but not discretely refutable.*
8. *There are 1-sets that are discretely refutable but not discretely testable.*

Proof.

1. $\Longrightarrow$: Let t be a strict test for A. We define $t'(x) = 0$ for $t(x) < \frac{1}{2}$ and $t'(x) = 1$ else. It is obvious that t' is a discrete strict test for A.

$\Longleftarrow$: Trivial.

2. Trivial by 1.
3. Trivial by 1.
4. $\Longrightarrow$: Trivial.

$\Longleftarrow$: Let t be a discrete test, s a discrete refutation for A. We define $t' = (t'_n)_{n \in \mathbb{N}}$ by considering two cases. The first one: $t_k(\mathbf{x}) = s_k(\mathbf{x})$ for all $k \leq n$. Note that this can happen only for finitely many n. Hence $t'_n(\mathbf{x})$ may be arbitrary, say $t'_n(\mathbf{x}) = 0$. Otherwise consider the maximal $k \leq n$ such that $t_k(\mathbf{x}) \neq s_k(\mathbf{x})$ and put $t'_n(\mathbf{x}) := t_k(\mathbf{x})$. Indeed t' is a strict test for A. To see this we study both possible cases:

$\mathbf{x} \in A$: Since t is a discrete test for A there is an $N \in \mathbb{N}$ such that $t_n(\mathbf{x}) = 1$ for all $n \geq N$. s is a discrete refutation for A, hence $s_n(\mathbf{x}) = 0$ for infinitely many n, i.e. also for some $K \geq N$. By construction this implies $t'_n(\mathbf{x}) = 1$ for all $n \geq K$, i.e. $t'_n(\mathbf{x}) \to 1$.

$\mathbf{x} \notin A$: An argument totally symmetric to that for the case above shows $t'_n(\mathbf{x}) \to 0$.

5. Trivial by the definition.

6. If $t = (t_n)_{n \in \mathbb{N}}$ is a discrete test (refutation) for A, then

$$t'_n(\mathbf{x}) = 1 - \left(\max\left(\operatorname{card}\{k \leq n \mid t_k(\mathbf{x}) = 0\}, 1\right)\right)^{-1}$$

is a refutation (test) for A.

7. The 1-set A constructed in Theorem 5 is discretely testable and terminal. If A were discretely refutable, then by claim 4 of this theorem strictly testable, contradicting Theorem 4.

8. Like 7, only using the set A from Theorem 6. □

We are interested in tests defining sets that are "typical" in the sense of probability theory, i.e. tests such that at least one of the numbers $P(X_{t,0})$ and

$P(X_{t,1})$ equals 0 (or 1). The forthcoming results show that we have to make some further restrictions. It seems natural to be interested in those sets which, in some sense, do not depend on finitely many members of the sequence. For this reason we use the following notions.

Definition 2. *A set $A \subseteq X$ is called permutable if for every finite permutation $\pi \in S_n$ (symmetric group acting on the set $\{1, \ldots, n\}$) and for every*

$$\mathbf{x} = (x_1, x_2, \ldots) \in A$$

we also have

$$\pi(\mathbf{x}) := (x_{\pi(1)}, \ldots, x_{\pi(n)}, x_{n+1}, \ldots) \in A.$$

A is called terminal if

$$(x_1, \ldots, x_n, x_{n+1}, \ldots) \in A$$

implies

$$(y_1, \ldots, y_n, x_{n+1}, x_{n+2}, \ldots) \in A$$

for arbitrary $n \in \mathbb{N}$ and $y_i \in M$, $i = 1, \ldots, n$. A is called invariant if

$$(x_1, x_2, \ldots, x_n, x_{n+1}, \ldots) \in A$$

implies

$$(y_1, y_2, \ldots, y_k, x_{n+1}, x_{n+2}, \ldots) \in A$$

for arbitrary $n, k \in \mathbb{N}$ and $y_i \in M$, $i = 1, \ldots, k$.

A test (refutation) t for the set A is called permutable (terminal, invariant) if A has this property.

Proposition 3. *The following chain of implications holds:*
A invariant $\Longrightarrow$ A terminal $\Longrightarrow$ A permutable $\Longrightarrow$ $P(A) \in \{0, 1\}$.
In general none of the implications can be replaced by equivalence.

Proof. The first two implications are trivial. The third one follows from the zero-or-one law of Hewitt and Savage. For a proof cf. [Fel], second edition, vol. II, p. 124. That the implications are no equivalences can be shown by the following examples in the case $M = \{0, 1\}$, $\mu(\{0\}) = \mu(\{1\}) = \frac{1}{2}$:

$A_1 := \{\mathbf{x} \mid x_{2n} = 0$ for infinitely many $n\}$ is terminal but not invariant.
$A_2 := \{\mathbf{x} \mid x_n = 0$ for some $n \in \mathbb{N}\}$ is permutable but not terminal.
$A_3 := \{(1, 0, 0, \ldots)\}$ satisfies $P(A) = 0 \in \{0, 1\}$ but is not permutable. □

Mostly it is useful to restrict the investigations to those spaces $(M, \mathcal{A}, \mu)$, where not all measurable sets $T \in \mathcal{A}$ are 0- or 1-sets.

DEFINITION 3. *A probability space $(M, \mathcal{A}, \mu)$ is called regular if there is a measurable subset $T \subseteq M$ with $0 < \mu(T) < 1$.*

The spaces of all permutable, terminal or invariant sets are nontrivial examples of probability spaces that are not regular. However, this restriction seems to be harmless with respect to the problem of randomness on "interesting" probability distributions.

In sections 5 to 8 we are going to collect some partial results that will be summarized in section 9.

5 First try: Permutable sets

For the following it is convenient to introduce the sets

$$M_0 := \{x \in M \mid \mu(\{x\}) = 0\}$$

and

$$M_1 := \{x \in M \mid \mu(\{x\}) > 0\}.$$

Note that M_1 is finite or countable, hence measurable if all singletons of M are measurable. Then we have $M_0 = M \setminus M_1$, therefore also M_0 is measurable in this case.

THEOREM 2.

1. *Suppose that $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ contains an $x_{n_0} \in M_0$. Then the set*

$$A := \{\mathbf{y} \mid y_n \neq x_{n_0} \textit{ for all } n \in \mathbb{N}\}$$

is strictly testable, permutable, has measure $P(A) = 1$ and does not contain $\mathbf{x}$.

2. *Suppose that $x_n \in T$ for only finitely many n, say* $\text{card}\{n \mid x_n \in T\} = N$, *for some $T \subseteq M$ with $\mu(T) > 0$. Then the set*

$$B := \{\mathbf{y} \mid \text{card}\{n \in \mathbb{N} \mid y_n \in T\} > N\}$$

is strictly testable, permutable, has measure $P(B) = 1$ and does not contain $\mathbf{x}$.

Proof.

1. It is trivial that A is permutable and $\mathbf{x} \notin A$. To see that $P(A) = 1$ it suffices to mention that $P(A_n) = 1$ for all $n \in \mathbb{N}$ and

$$A = \bigcap_n A_n$$

if we put

$$A_n := \{\mathbf{y} \mid y_n \neq x_{n_0}\}.$$

Finally, a strict test for A is given by

$$t_n(\mathbf{y}) := 1 \quad \text{if} \quad y_k \neq x_{n_0} \qquad \text{for all} \quad k \leq n$$

and

$$t_n(\mathbf{y}) := 0 \qquad \text{otherwise.}$$

2. Again it is trivial that B is permutable and $\mathbf{x} \notin B$. The strong law of large numbers guarantees that $y_n \in T$ happens with probability 1 for infinitely many $n \in \mathbb{N}$, hence $P(B) = 1$. A strict test for B can be defined by

$$t_n(\mathbf{y}) := 1 \quad \text{if} \quad \operatorname{card}\{k \leq n \mid y_k \in T\} > N$$

and

$$t_n(\mathbf{y}) := 0 \qquad \text{otherwise.}$$

□

THEOREM 3. *Suppose that all singletons $\{m\} \subseteq M$ are measurable, $\mu(M_0) = 0$ and $\mathbf{x}$ contains all $m \in M_1$ infinitely many times but no $m \in M_0$. Furthermore let t be a strict test for the permutable one-set A, then $\mathbf{x} \in A$.*

Proof. By Theorem 1.1 we may assume that t is discrete. Note that

$$\mathbf{x} \in B := \bigcap_{m \in M_1} B_m$$

with

$$B_m := \{\mathbf{y} \mid y_n = m \text{ for infinitely many } n \in \mathbb{N}\}.$$

By the strong law of large numbers we have $P(B_m) = 1$ for all $m \in M_1$ and thus, since M_1 is finite or countable, $P(B) = P(A \cap B) = 1$. This implies that for every

$$(y_1, \dots, y_n) \in M_1^n$$

there is a $\mathbf{z} \in A \cap B$ such that

$$(y_1, \ldots, y_n) = (z_1, \ldots, z_n).$$

(Note that the set of all $\mathbf{z}$ with this property has positive probability.) Now we are going to derive a contradiction from the assumption $\mathbf{x} \notin A$. We are constructing a sequence $\mathbf{x}'$ in the following way. By $\mathbf{x} \notin A$ there is an $n_1 \in \mathbb{N}$ such that $t_{n_1}(\mathbf{x}) = 0$. Put

$$(x'_1, \ldots, x'_{n_1}) = (x_1, \ldots, x_n).$$

By the observation above there is a $\mathbf{z} \in A \cap B$ such that

$$(x'_1, \ldots, x'_{n_1}) = (z_1, \ldots, z_{n_1}).$$

For sufficiently large $n_2 > n_1$ we therefore get $t_{n_2}(\mathbf{x}') = 1$ if

$$(x'_1, \ldots, x'_{n_2}) = (z_1, \ldots, z_{n_2}).$$

Since $\mathbf{x} \in B$, there is a finite permutation $\pi \in S_k$, k sufficiently large, such that

$$\left(x_{\pi(1)}, \ldots, x_{\pi(n_2)}\right) = (x'_1, \ldots, x'_{n_2}).$$

A is permutable, hence $\mathbf{x} \notin A$ implies $\pi(\mathbf{x}) \notin \pi(A) = A$. Thus we can find an $n_3 > n_2$ such that $t_{n_3}(\mathbf{x}') = 0$ if

$$(x'_1, \ldots, x'_{n_3}) = \left(x_{\pi(1)}, \ldots, x_{\pi(n_3)}\right) = \left(x'_1, \ldots, x'_{n_2}, x_{\pi(n_2+1)}, \ldots, x_{\pi(n_3)}\right).$$

Continuing this construction we get $t_{n_4}(\mathbf{x}') = 1$ etc. Therefore $t_n(\mathbf{x}')$ does not converge, contradicting the condition that t is a strict test. Hence indeed $\mathbf{x} \in A$. □

6 Second try: Terminal sets

THEOREM 4. *Let $A \subseteq X$ be terminal and strictly testable. Then either $A = X$ or $A = \emptyset$.*

Proof. Let t be a strict test for the terminal set A. Let us suppose that the theorem fails, i.e. that there are $\mathbf{x} \in A$ and $\mathbf{y} \in X \setminus A$. We are going to construct the sequence $\mathbf{z}$ in the following way:

Assume that $(z_1, \ldots, z_{n_{2k}})$ is already defined. A is terminal and $\mathbf{x} \in A$, hence

$$\mathbf{x}' := (z_1, \ldots, z_{n_{2k}}, x_{n_{2k}+1}, x_{n_{2k}+2}, \ldots) \in A$$

and

$$t_{n_{2k+1}}(\mathbf{x}') = 1$$

for some $n_{2k+1} > n_{2k}$. Similarly, since $\mathbf{y} \notin A$,

$$\mathbf{y}' := (x'_1, \dots, x'_{n_{2k+1}}, y_{n_{2k+1}+1}, y_{n_{2k+2}+2}, \dots) \notin A$$

and

$$t_{n_{2(k+1)}}(\mathbf{y}') = 0$$

for some $n_{2(k+1)} > n_{2k+1}$. In this way we get the sequence

$$\mathbf{z} = (x_1, \dots, x_{n_1}, y_{n_1+1}, \dots, y_{n_2}, x_{n_2+1}, \dots, x_{n_3}, \dots)$$

such that $t_{n_k}(\mathbf{z}) = 0$ for even k and $t_{n_k}(\mathbf{z}) = 1$ for odd k. Hence $t_n(\mathbf{z})$ does not converge. But this is impossible since t is assumed to be a strict test. The contradiction proves the theorem. □

Theorem 4 says that there are only the trivial sets that are terminal and strictly testable. If we do not require strict testability but only discrete testability, we get an abundance of terminal sets.

THEOREM 5. *Let $(M, \mathcal{A}, \mu)$ be regular and $\mathbf{x} \in X$ arbitrary. Then there is a terminal set A that is discretely testable and satisfies $P(A) = 1$ and $\mathbf{x} \notin A$.*

Proof. Choose $T \subseteq M$ such that

$$\frac{1}{2} \leq p := \mu(T) < 1,$$

and let f denote the characteristic function of T, i.e. $f(x) = 1$ for $x \in T$ and $f(x) = 0$ otherwise. We define

$$t_n(\mathbf{y}) = 1 \quad \text{if} \quad \left(f\left(y_{[\frac{n}{2}]}\right), \dots, f(y_n)\right) \neq \left(f\left(x_{[\frac{n}{2}]}\right), \dots, f(x_n)\right)$$

and

$$t_n(\mathbf{y}) = 0 \qquad \text{otherwise.}$$

By definition t is a discrete test for the set

$$A := X_{t,1} = \{\mathbf{y} \mid t_n(\mathbf{y}) \to 1\}.$$

It is obvious that A is terminal and $\mathbf{x} \notin A$.

The rest we have to show is $P(A)=1$. Using the abbreviations

$$B_n := \{\mathbf{y} \mid \left(f(y_{[\frac{n}{2}]}),\dots,f(y_n)\right) = \left(f(x_{[\frac{n}{2}]}),\dots,f(x_n)\right)\}$$

and

$$B := \bigcap_N \bigcup_{n\geq N} B_n$$

we have $B = X \setminus A$ and $P(B_n) \leq p^{\frac{n}{2}}$, hence

$$P(B) \leq \lim_{N\to\infty} \sum_{n\geq N} P(B_n) \leq \lim_{N\to\infty} p^{\frac{N}{2}} \sum_{n=0}^{\infty} \left(\sqrt{p}\right)^n = 0,$$

implying

$$P(A) = P(X\setminus B) = 1 - P(B) = 1.$$

□

7 Third try: Invariant sets

THEOREM 6. *Let $(M,\mathcal{A},\mu)$ be regular and $\mathbf{x}\in X$ arbitrary. Then there is a discretely refutable set $A\subseteq X$ with $P(A)=1$ which is invariant such that $\mathbf{x}\notin A$.*

Proof. By regularity there is a set $T\subseteq M$ such that $0<\mu(T)<1$. Let f denote its characteristic function and $s_n(\mathbf{y})$ be the maximal integer k such that

$$\left(f(x_i),\dots,f(x_{i+k-1})\right) = \left(f(y_j),\dots,f(y_{j+k-1})\right)$$

for some i, j satisfying $i+k-1\leq n$, $j+k-1\leq n$. Now we put $t_n(\mathbf{y})=0$ if $s_{n-1}(\mathbf{y}) = s_n(\mathbf{y})$ and $t_n(\mathbf{y})=1$ in the other case and $A := X\setminus X_{t,1}$.

A simple reasoning left to the reader shows that

$$B = X_{t,1} = X\setminus A = \bigcup_{k,l\in\mathbb{N}} B_{k,l}$$

with

$$B_{k,l} = \{\mathbf{y} \mid \left(f(y_k), f(y_{k+1}),\dots\right) = \left(f(x_l), f(x_{l+1}),\dots\right)\}.$$

For all k, l we have $P(B_{k,l}) = 0$, hence $P(B)=0$ and $P(A)=1$. Since $\mathbf{x}\in B_{1,1}$ also $\mathbf{x}\notin A$ holds. It is obvious that B and hence A is invariant and, by construction, t is a discrete refutation for A. □

Theorem 7. *Let $N \subseteq M$ be a set of measure $\mu(N) = 0$ and $x_n \in N$ for infinitely many $n \in \mathbb{N}$. Then the set*

$$A := \{\mathbf{y} \mid y_n \in N \text{ for only finitely many } n \in \mathbb{N}\}$$

is invariant, discretely testable and satisfies $P(A) = 1$ and $\mathbf{x} \notin A$.

Proof. A discrete test t for A is given by $t_n(\mathbf{y}) := 1$ if $y_n \notin N$ and $t_n(\mathbf{y}) := 0$ for $y_n \in N$. It is obvious that A is invariant and does not contain $\mathbf{x}$. Hence our last step is to prove $P(A) = 1$. This follows from

$$B := X \setminus A = \bigcap_n \bigcup_{k \geq n} B_k$$

with $B_k = \{\mathbf{y} \mid y_k \in N\}$ and therefore

$$P(B) \leq P\Big(\bigcup_{k \in \mathbb{N}} B_k\Big) \leq \sum_{k \in \mathbb{N}} P(B_k) = \sum_{k \in \mathbb{N}} \mu(N) = 0 .$$

□

8 Addition: Baire properties

Theorem 8. *Let M be considered as a discrete topological space and X be equipped with the product topology. If $A \subseteq X$, $A \neq X$, is terminal and discretely testable, then it is of first category.*

Proof. If t is a discrete test for A, then

$$A = \{\mathbf{y} \mid t_n(\mathbf{y}) \to 1\} = \bigcup_{N \in \mathbb{N}} \bigcap_{n \geq N} A_n$$

with

$$A_n = \{\mathbf{x} \mid t_n(\mathbf{x}) = 1\} .$$

Since each A_n is closed, the same is true for all

$$B_N := \bigcap_{n \geq N} A_n$$

and the proof is complete if we show that the sets B_N, $N \in \mathbb{N}$, have empty interior. To do this, suppose that $\mathbf{x}$ is an inner point of B_N. By definition of product topology there is an $n_0 \in \mathbb{N}$ such that

$$B_N \supseteq B' := \{\mathbf{y} \mid y_1 = x_1, \ldots, y_{n_0} = x_{n_0}\} .$$

On the other hand take any $\mathbf{y} \in X \setminus A$. A is terminal, hence

$$\mathbf{y}' := (x_1, \ldots, x_{n_0}, y_{n_0+1}, y_{n_0+2}, \ldots) \in X \setminus A ,$$

implying $t_{n_1}(\mathbf{y}') = 0$ for some $n_1 \geq N$, $\mathbf{y}' \notin A_{n_1}$. But this means $\mathbf{y}' \notin B_N$, contradicting $\mathbf{y}' \in B' \subseteq B_N$. □

COROLLARY 1. *Let again M be a discrete topological space and $A \subseteq X$ nonvoid, terminal and discretely refutable. Then A is of second category.*

Proof. The set $X \setminus A$ is discretely testable and also terminal, hence by Theorem 8 of first category. The topological space of the sequences on a discrete space is a complete metric space, hence, by Baire's Theorem, A is of second category. □

The condition in Theorem 8 cannot be weakened in general. This is expressed by

THEOREM 9. *Let M be a complete metric space and $\emptyset \neq T \neq M$ an open set. Consider*

$$A := \{\mathbf{x} \mid \ x_n \in T \ \textit{for infinitely many} \ n \in \mathbb{N}\}$$

and

$$B := \{\mathbf{x} \mid \ x_n \in T \ \textit{for some} \ n \in \mathbb{N}\}\,.$$

1. *A is a nontrivial example of a set which is terminal, discretely refutable and residual, i.e. its complement is meager. Hence, by Baire's Theorem, A is of second category.*
2. *B is a nontrivial example of a set which is permutable, strictly testable, open and dense, hence residual and of second category.*

Proof.

1.

$$t_n(\mathbf{x}) := f(x_n)$$

(f the characteristic function of $M \setminus T$) a discrete refutation. It is trivial that $A \neq X$ and A is terminal. To show that A is of second category we consider the closed sets

$$C_N := \bigcap_{n \geq N} \{\mathbf{x} \mid \ x_n \notin T\}\,.$$

Since

$$X \setminus A = C := \bigcup_{N \in \mathbb{N}} C_N$$

and since each C_N has empty interior, C is of first category.

2. It is obvious that B is permutable and $B \neq X$. A strict test for B can be defined by $t_n(\mathbf{x}) = 1$ if $x_k \in T$ for some $k \leq n$ and $t_n(\mathbf{x}) = 0$ otherwise. Finally

$$B = \bigcup_{k \in \mathbb{N}} \left(\prod_{n=1}^{k-1} M \times T \times \prod_{n=k+1}^{\infty} M \right)$$

is open and dense. □

9 Coda: Summary and final remarks

We are going to summarize the results of sections 5, 6 and 7. This can be done by the following table, where $(M, \mathcal{A}, \mu)$ is assumed to be regular. The abbreviations s., d., t. and r. are standing for strictly, discretely, testable and refutable.

$\cap$	invariant		terminal		permutable		1-set
s.t.	X	$\Longleftarrow$	X (Th. 4)		U[1.)] (Th. 2, 3)		$\emptyset$[2.)]
d. t.	$\emptyset$[3.)] (Th. 7)		$\emptyset$ (Th. 5)	$\Longrightarrow$	$\emptyset$	$\Longrightarrow$	$\emptyset$
d. r.	$\emptyset$ (Th. 6)	$\Longrightarrow$	$\emptyset$	$\Longrightarrow$	$\emptyset$	$\Longrightarrow$	$\emptyset$

[1.)] Under the further condition that all singletons are measurable. U denotes the set of all sequences $\mathbf{x}$ containing exactly those $m \in M$ with $\mu(\{m\}) > 0$, each of them infinitely many times.

[2.)] To see this consider any measurable $T \subseteq M$ with $0 < \mu(T) < 1$ and define for arbitrary $\mathbf{x} = (x_n)_{n \in \mathbb{N}} \in X$ $t_n(y_n) = 0$ if $f(x_k) = f(y_k)$ for all $k = 1, \ldots, n$ (f characteristic function of T) and $t_n(y_n) = 1$ otherwise. It is clear that t is a strict test for a 1-set A with $\mathbf{x} \notin A$.

[3.)] Under the further condition that all singletons are 0-sets. In this case for an arbitrary sequence $\mathbf{x}$ the set $N = \{x_n \mid n \in \mathbb{N}\}$ satisfies the condition of Theorem 7. S. Shelah has pointed out that also in the discrete case no sequence passes all invariant discrete tests for 1-sets, cf. the forthcoming paper [G2]. This paper will also characterize the different notions of testability in terms of the hierarchy of Borel sets (open sets, closed sets, F_σ-sets, G_σ-sets etc.) and in particular will show that the set of uniformly distributed sequences is testable but not refutable. Corollary 1 to Theorem 8 shows that, for any nontrivial invariant and discretely testable 1-set, the set $X \setminus A$ has to be of second category, hence uncountable.

To explain how to read the above table we treat an example. The letter U in the line "strictly testable" and the row "permutable" means that the intersection of all strictly testable and permutable 1-sets is just the set U defined in the first footnote. The references to this statement are Theorems 2 and 3. Implication arrows indicate that Proposition 3 suffices to deduce one statement expressed in the table from another one.

Final remark. The only nontrivial entry in the above table is the set U in the case of strictly testable and permutable sets. But even this set does not look useful to define *pseudorandom sequences*. One of the simplest examples to show this is the very "deterministic" sequence

$$\mathbf{x} = (0,1,0,1,\ldots) \in U$$

on the space $M = \{0,1\}$ with $\mu(\{0\}) = \mu(\{1\}) = \frac{1}{2}$. $\mathbf{x}$ is not even 2-u.d. (cf. section 2).

Thus the results of this paper seem to indicate that definitions of *pseudorandomness* that are given in more general terms than special distribution properties (as for instance those presented in section 2) are either too weak or too restrictive. This situation would also remain if the intersections were taken only over the countable set of those tests that can be constructed explicitly such that in the table one had instead of $\emptyset$ some nontrivial 1-sets. This fact gets clear by inspection of the proofs showing that the resulting intersections could contain only sequences that cannot be constructed explicitly, i.e. that are useless for computational matters (cf. section 1).

REFERENCES

[Cha1] CHAITIN, G. J.: *Algorithmic Information Theory*, Cambridge University Press, 1987, 1988, 1990.

[Cha2] CHAITIN, G. J.: *Information, Randomness and Incompleteness – Papers on Algorithmic Information Theory*, World Scientific, Singapore, 1987, 1990.

[Cha3] CHAITIN, G. J.: *Incompleteness theorems for random reals*, Adv. in Appl. Math. **8** (1987), 119–146.

[Chu] CHURCH, A.: *On the concept of a random sequence*, Bull. Amer. Math. Soc. **46** (1940), 130–135.

[D] DÖRGE, K.: *Zu der von R. v. Mises gegebenen Begründung der Wahrscheinlichkeitstheorie*, Math. Z. **32** (1930), 232–258.

[Dr-Ti-Wi] DRMOTA, M.—TICHY, R. F.—WINKLER, R.: *Completely uniformly distributed sequences of matrices.* In: Number-Theoretic Analysis. Lecture Notes in Math. 1452, Springer, Berlin-Heidelberg-New York, 1990, pp. 43–57.

[Dr-Wi] DRMOTA, M.—WINKLER, R.: *$s(N)$-uniform distribution modulo* 1, J. Number Theory, (To appear).

[Fe1] FELLER, W.: *An Introduction to Probability Theory and its Applications*, John Wiley and Sons Inc., New York-London-Sydney-Toronto, 1966.

[Fe2] FELLER, W.: *Über die Existenz sogenannter Kollektive*, Fund. Math. **32** (1939), 87–96.

[Fl-Ki-Ti1] FLAJOLET, P.—KIRSCHENHOFER, P.—TICHY, R. F.: *Discrepancy of Sequences in Infinite Strings. Colloq. Math. Soc. Janos Bolyai*, North-Holland, Amsterdam-New York, 1986.

[Fl-Ki-Ti2] FLAJOLET, P.—KIRSCHENHOFER, P.—TICHY, R. F.: *Deviations from uniformity in random strings*, Probab. Theory Related Fields **80** (1988), 139–150.

[G1] GOLDSTERN, M.: *An application of Shoenfield's absoluteness theorem to the theory of uniform distribution*, Monatsh. Math., (To appear).

[G2] GOLDSTERN, M.: *Two more remarks on pseudorandom sequences*, (Preprint).

[Gra] GRABNER, P.: *Block distribution in random strings*, Ann. Inst. Fourier (Grenoble), (To appear).

[Gri] GRILL, K.: *A note on randomness*, Statist. Probab. Letters **14** (1992), 229–233.

[H] HLAWKA, E.: *Theorie der Gleichverteilung*, Bibl. Inst., Mannheim-Wien-Zürich, 1979.

[Ki-Ti] KIRSCHENHOFER, P.—TICHY, R. F.: *Some distribution properties of 0 - 1 -sequences*, Manuscripta Math. **54** (1985), 205–219.

[Kn] KNUTH, D. E.: *The Art of Computer Programming. Vol. II*, Addison-Wesley, Reading Mass., 1981.

[Ko1] KOLMOGOROFF, A. N.: *Grundbegriffe der Wahrscheinlichkeitsrechnung. Ergeb. Math. Grenzgeb.* (2), Springer, Berlin-New York, 1933.

[Ko2] KOLMOGOROFF, A. N.: *Drei Zugänge zur Definition des Begriffs "Informationsgehalt".* (Russian), Problemy Peredachi Informatsii **1** (1965), 3–11.

[Ko3] KOLMOGOROFF, A. N.: *On tables of random numbers*, Sankhyā Ser. A **25** (1963), 369–376.

[Ku-N] KUIPERS, L.—NIEDERREITER, H.: *Uniform Distribution of Sequences*, Wiley, New York, 1974.

[vL] van LAMBALGEN, M.: *Von Mises' definition of random sequences reconsidered*, J. Symbolic Logic **52** (1987), 725–755.

[L] LEVIN, L. A.: *On the notion of a random sequence*, Soviet Math. Dokl. **14** (1973), 1414–1416.

[Le1] LEVINE, M. B.: *On the uniform distribution of the sequence* $\{\alpha\lambda^x\}$. (Russian), Mat. Sb. **98** (1975), 207–222, (Translation: Math. USSR-Sb. **27** (1975), 183–197).

[Le2] LEVINE, M. B.: *On the completely uniform distribution of fractional parts of the exponentional function.* (Russian, English Summary), Trudy Sem. Im. Petrovsk. **7** (1981), 245–256.

[Lev-Sh-So-Tu] LEVITAN, YU. L.—SHUKHMAN, B. V.—SOBOL, I. M.—TURCHANINOV, V. I.: *Quasirandom Sequence Generators*, Keldysh Inst. of Appl. Math., Russian Acad. of Sciences, 1992.

[M1] MARTIN-LÖF, P.: *The definition of random sequences*, Inform. Control **9** (1966), 602–619.

[M2] MARTIN-LÖF, P.: *On the notion of randomness.* In: Intuitionism Proof Theory. Proc. Summer Conf. Buffalo N.Y. 1968, 1970, pp. 73–78.

[vM1] von MISES, R.: *Grundlagen der Wahrscheinlichkeitstheorie*, Math. Z. **5** (1919), 52–99.

[vM2] von MISES, R.: *Wahrscheinlichkeit, Statistik und Wahrheit*, Springer, Wien, 1951.

[vM3] von MISES, R.: *Mathematical Theory of Probability and Statistics*, Acad. Press, New York-London, 1964.

[N1] NIEDERREITER, H.: *Quasi-Monte Carlo methods and pseudorandom numbers*, Bull. Amer. Math. Soc. **84** (1978), 957–1041.

[N2] NIEDERREITER, H.: *Pseudozufallszahlen und die Theorie der Gleichverteilung*, Sitzungsber. Österreich. Akad. Wiss. Math.-Natur. Kl. Abt. II **195** (1986), 109–138.

[N3] NIEDERREITER, H.: *Point sets and sequences with small discrepancy*, Monatsh. Math. **104** (1987), 273–337.

[N4] NIEDERREITER, H.: *Pseudorandom numbers generated from shift register sequences.* In: Number-Theoretic Analysis. Lecture Notes in Math. 1452 (E. Hlawka, R. F. Tichy, eds.), Springer, Berlin-Heidelberg-New York, 1990, pp. 165–177.

[N5] NIEDERREITER, H.: *Random Number Generation and Quasi-Monte Carlo Methods*, Society for industrial and applied mathematics, Philadelphia, Pennsylvania, 1992.

[N-Ti] NIEDERREITER, H.—TICHY, R. F.: *Solution of a problem of Knuth on complete uniform distribution of sequences*, Mathematika **23** (1985), 26–32.

[P] POPPER, K.: *Logik der Forschung, Achte, weitere verbesserte und vermehrte Auflage, J. C. B. Mohr (Paul Siebeck), Tübingen,* 1984.

[S1] SCHNORR, C. P.: *Zufälligkeit und Wahrscheinlichkeit, Eine algorithmische Begründung der Wahrscheinlichkeitstheorie. Lecture Notes in Math.* 218, Springer, Berlin-Heidelberg-New York, 1971.

[S2] SCHNORR, C. P.: *Process complexity and effective random tests*, J. Comput. System Sci. **7** (1973), 376–388.

[Sol] SOBOL, I. M.: *Quasi-Monte Carlo methods*, Progress in Nuclear Energy **24** (1990), 55–61.

[So2] SOBOL, I. M.: *Die Monte-Carlo-Methode*, Deutscher Verlag der Wissenschaften, Berlin, 1991.

[Sv] SVOZIL, K.: *The mathematical foundations of physical randomness and indeterminisness.* In: Jahrb. Kurt-Gödel-Ges., Wien, 1988, pp. 53–85.

[Ta] TAUSWORTHE, R. C.: *Random numbers generated by linear recurrence modulo two*, Math. Comp. **19** (1965), 201–209.

[Tel] TEZUKA, S.: *On the discrepancy of GFSR pseudorandom numbers*, J. Assoc. Comput. Math. **34** (1987), 939–949.

[Te2] TEZUKA, S.: *On optimal GFSR pseudorandom number generators*, Math. Comp. **50** (1988), 531–533.

[Til] TICHY, R. F.: *Ein metrischer Satz über vollständig gleichverteilte Folgen*, Acta Arith. **48** (1987), 197–207.

[Ti2] TICHY, R. F.: *Zur Analyse und Anwendung von Zufallszahlen.* In: Jahrb. Kurt-Gödel-Ges., Wien, 1990, pp. 109–116.

[V] VILLE, J.: *Etude critique de la notion de collectif*, Gauthier-Villars, Paris, 1939.

[Wal] WALD, A.: *Sur la notion de collectif dans le calcul des probabilités*, C. R. Acad. Sci. Paris **202** (1936), 180–183.

[Wa2] WALD, A.: *Die Widerspruchsfreiheit des Kollektivbegriffs in der Wahrscheinlichkeitsrechnung*, Ergebnisse eines math. Koll. **8** (1937), 38–72.

[Wi] WINKLER, R.: *Some constructive examples in uniform distribution on finite sets and normal numbers*, Anz. Österreich. Akad. Wiss. Math.-Natur. Kl. **126** (1989), 1–8.

Received September 14, 1992
Revised September 26, 1993

Institut für Algebra und Diskrete Mathematik
TU Wien
Wiedner Hauptstrasse 8-10/118
A-1040 Wien
Austria