UNIFORM DISTRIBUTION AND ALGORITHMIC RANDOMNESS

JEREMY AVIGAD

Abstract. A seminal theorem due to Weyl [10] states that if (a_n) is any sequence of distinct integers, then, for almost every $x \in \mathbb{R}$, the sequence $(a_n x)$ is uniformly distributed modulo one. In particular, for almost every x in the unit interval, the sequence $(a_n x)$ is uniformly distributed modulo one for every *computable* sequence (a_n) of distinct integers. I show that every Schnorr-random real has this property, and raise questions as to how this property compares with other notions of randomness.

1. Introduction

If x is a real number, let $\{x\}$ denote the fractional part of x, and let λ denote Lebesgue measure. A sequence $(x_n)_{n\in\mathbb{N}}$ of real numbers is said to be *uniformly distributed modulo one* if for every interval $I \subseteq [0, 1]$,

$$\lim_{n\to\infty} \frac{|\{i < n \mid \{x_i\} \in I\}|}{n} = \lambda(I).$$

In words, a sequence is uniformly distributed modulo one if the limiting frequency with which it visits any given interval is what one would expect if the elements of the sequence were chosen at random. A remarkable theorem by Hermann Weyl [10] states the following:

Theorem 1.1. *If (a_n) is any sequence of distinct integers, then for almost every x, $(a_n x)$ is uniformly distributed modulo one.*

This seminal result lies at the intersection of harmonic analysis, number theory, ergodic theory, and computer science; see, for example, [1, 3, 4, 6].

Now let us restrict attention to *computable* sequences (a_n) of distinct integers. Since there are only countably many of these, it follows that almost every $x \in [0, 1]$ has the property that $(a_n x)$ is uniformly distributed modulo one for each such sequence. Let us say that such an x is *UD random*. The goal of this note is to explore the relationship of UD randomness to other senses in which a real number in the unit interval can be said to be "random."

Recall that a real number x is *normal* if, when it expressed in any base $b > 1$, each sequence of k digits occurs with limiting frequency b^{-k}. It is not hard to see that a real number x is normal with respect to base b if and only if the sequence $(b^k x)$ is uniformly distributed modulo one. Thus every UD-random real number is normal to every base. In fact, more is true: if x is UD random and (a_n) is any computable sequence of distinct integers, then the limiting frequency of occurences of any k-digit block at positions $a_0, a_1, a_2, \ldots$ of x is again b^{-k}.

Thus a UD-random number looks random, at least in some ways. In Section 3, however, I conjecture that there are straightforward ways in which a UD-random number can look very *nonrandom*. But first, Section 2 considers the notion of UD randomness from the point of view of algorithmic randomness. In particular, I show that every Schnorr-random element of $[0, 1]$ is UD random, but there are Kurtz-random elements of $[0, 1]$ that are not.

2. UD randomness and algorithmic randomness

The subject of algorithmic randomness [2, 5] aims to characterize different senses in which a real number (or, say, an infinite binary sequence) can be said to be "random." A subset G of $\mathbb{R}$ is said to be *effectively open* if there are computable sequences $(a_i)_{i\in\mathbb{N}}$ and $(b_i)_{i\in\mathbb{N}}$ of rational numbers such that $G = \bigcup_i (a_i, b_i)$. An effectively open subset of the unit interval is obtained by restricting G to $[0, 1]$. A sequence $(G_j)_{j\in\mathbb{N}}$ of effectively open sets is *uniformly effective* if the representing sequences $(a_i^j)_{i\in\mathbb{N}}$ and $(b_i^j)_{i\in\mathbb{N}}$ can be computed by a single algorithm with j as a parameter. A *Martin-Löf test* is a uniformly effective sequence of open sets (G_j) such that for each j, $\lambda(G_j) \leq 2^{-j}$. An element x of $[0, 1]$ *fails* the Martin-Löf test (G_j) if it is an element of $\bigcap_j G_j$, and *passes* the test otherwise. An element x of $[0, 1]$ is *Martin-Löf random* if it passes every Martin-Löf test. In other words, x is Martin-Löf *non*random if it can be covered, effectively, by arbitrarily small open sets, and thus is contained in an effectively presented null G_δ set.

One can weaken or strengthen this notion of randomness by restricting or enlarging the class of tests. For example, a *Schnorr test* is a Martin-Löf test with the additional property that for each j, the measure of G_j is computable. This has the effect that when enumerating the inervals that cover the nonrandoms with a set of size at most 2^{-j}, one knows how much of the set is yet to come. An element x of $[0, 1]$ is Schnorr random if it passes every Schnorr test.

An element x of $[0, 1]$ is *Kurtz random* if it is contained in an effectively closed set of measure 0, that is, the complement of an effectively open set of measure 1. Clearly every Martin-Löf-random element of $[0, 1]$ is Schnorr random, and every Schnorr-random element is Kurtz random. The notion of Kurtz randomness is fairly weak. For example, it is not hard to show that if a real x is weakly 1-generic then it is Kurtz random but its binary digits fail to satisfy the law of large of numbers. (See [2, Section 2.24] for the definition of weak 1-genericity, and Section 8.11 for the facts just mentioned.) In particular, this shows that there are Kurtz-random reals that are not UD random.

The main result of this section is this:

Theorem 2.1. *Every Schnorr-random element of* $[0, 1]$ *is UD random.*

The proof of this lemma follows conventional proofs of Weyl's theorem, for example, as presented in [1, 3, 4]; we need only make certain constructive and quantitative aspects of the argument explicit. To start with, we need an equivalent formulation of uniform distribution modulo one.

Lemma 2.2. *For any sequence* (x_n) *of real numbers, the following are equivalent:*

(1) (x_n) *is uniformly distributed modulo 1.*
(2) *For any continuous function* $f : [0,1] \to \mathbb{R}$*,*

$$\int f \, d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{j<n} f(x_j \bmod 1).$$

(3) *For any* $h \neq 0$*,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j<n} e^{2\pi i h x_j} = 0.$$

The third property is known as "the Weyl criterion." Roughly, the second follows from the first using Riemann sums to approximate the integral; the third follows immediately from the second; and the first follows from the third, using trigonometric sums to approximate the characteristic function of an interval. For details, see any of the references mentioned above.

It will be convenient to adopt the Vinogradov convention of writing $e(x)$ for $e^{2\pi i x}$, thus freeing up the symbol i to be used as an index. Note that $\overline{e(x)} = e(-x)$ and $e(x+y) = e(x)e(y)$ for every x and y. Theorem 2.1 is an immediate consequence of the next lemma, since if x is a Schnorr-random element of $[0,1]$ and $(a_j)_{j\in\mathbb{N}}$ is any computable sequence of distinct integers, the lemma implies that $(a_j x)_{j\in\mathbb{N}}$ is uniformly distributed modulo one.

Lemma 2.3. *For each* i*, let* $(a^i_j)_{j\in\mathbb{N}}$ *be a sequence of distinct integers, such that* a^i_j *is uniformly computable from* i *and* j*. Then there is a Schnorr test* S *such that for every* x *not in* S *and every integer* i*,* $(a^i_j x)_{j\in\mathbb{N}}$ *is uniformly distributed modulo one.*

Proof. For the moment, fix a sequence (a_i) of distinct integers, and for any x in $[0,1]$ write

$$S_n(x) = \frac{1}{n} \sum_{j<n} e(a_j x).$$

Then we have

$$|S_n(x)|^2 = S_n(x)\overline{S_n(x)} = \frac{1}{n^2} \sum_{j,k<n} e((a_j - a_k)x),$$

and hence

$$\int_0^1 |S_n(x)|^2 \, dx = \frac{1}{n^2} \sum_{j,k<n} \int_0^1 e((a_j - a_k)x) \, dx = \frac{1}{n},$$

since each term in the sum has integral 0 except when $j = k$. (A more perspicuous way of carrying out this calculation is to notice that the left hand side is the square $\|S_n\|_2^2$ of the norm of S_n in the Hilbert space $L^2([0,1])$, and the functions $e(a_0 x), e(a_1 x), \ldots$ form an orthornormal set.) Markov's inequality implies that for every $\varepsilon > 0$ we have

$$\mu(\{x \mid |S_n(x)| > \varepsilon\}) = \mu(\{x \mid |S_n(x)|^2 > \varepsilon^2\}) \leq \frac{1}{n\varepsilon^2}.$$

In other words, for n large, $S_n(x)$ is small for most x.

Now fix a doubly-indexed sequence (a^i_j) as in the statement of the lemma, and for each i write

$$S^i_n(x) = \frac{1}{n} \sum_{j<n} e(a^i_j x).$$

Without loss of generality we can assume that for every $h > 1$ and i the sequence $(a^i_j h)_{j\in\mathbb{N}}$ already appears as one of the sequences $(a^{i'}_j)_{j\in\mathbb{N}}$ for some i', since we can replace the original sequence of sequences with a sequence that includes all such multiples. By the Weyl criterion, then, it suffices to find a Schnorr test S such that for each x not in S we have $\lim_{n\to\infty} S^i_n(x) = 0$.

A simple calculation shows that for any m such that $n^2 \leq m < (n+1)^2$ we have $|S_m(x)| \leq |S_{n^2}(x)| + 2/\sqrt{n}$; in other words, between n^2 and $(n+1)^2$, the averages do not change that much. This reduces our task to finding a Schnorr test S such that for each x not in S we have $\lim_{n\to\infty} S^i_{n^2}(x) = 0$.

For each i, rational $\varepsilon > 0$, and n define

$$A_{i,\varepsilon,m} = \{x \mid \exists n \geq m \; |S^i_{n^2}(x)| > \varepsilon\}.$$

By the calculation above, we have

$$\mu(A_{i,\varepsilon,m}) \leq \sum_{n\geq m} \mu(\{x \mid |S^i_{n^2}(x)| > \varepsilon\}) \leq \sum_{n\geq m} \frac{1}{n^2\varepsilon^2},$$

which decreases to 0 as m approaches infinity. Choose an enumeration (i_j, ε_j) of all pairs (i, ε), and for each k and j choose $m_{j,k}$ large enough so that $\mu(A_{i_j,\varepsilon_j,m_{j,k}}) < 2^{-(j+k+1)}$. For each k, let

$$G_k = \bigcup_j A_{i_j,\varepsilon_j,m_{j,k}}.$$

Then for each k, $\mu(G_k) \leq \sum_j 2^{-(j+k+1)} = 2^{-k}$. Moreover, the measure of G_k is uniformly computable in k, since for each v the measure of $\bigcup_{j\leq v} A_{i_j,\varepsilon_j,m_{j,k}}$ is computable and the measure of $\bigcup_{j>v} A_{i_j,\varepsilon_j,m_{j,k}}$ is at most $\sum_{j>v} 2^{-(j+k+1)} = 2^{-(k+v)}$. Thus the sequence (G_k) is a Schnorr test, and need only confirm that it meets the specification above.

So suppose x passes this test, that is, for some k, x is not in G_k. Given any i and $\varepsilon > 0$, choose j such that $(i, \varepsilon) = (i_j, \varepsilon_j)$. Since x is not in G_k, it is not in $A_{i_j,\varepsilon_j,m_{j,k}}$. This means that for every $n \geq m_{j,k}$, we have $|S^i_{m^2}(x)| \leq \varepsilon$. Since i and $\varepsilon > 0$ were arbitrary, we have that $S^i_{m^2}(x)$ approaches 0 for every i, as required. □

The lemma also has the following nice consequence, to the effect that we can actually *compute* an element x that is UD random with respect to a computable list of computable sequences.

Theorem 2.4. *For every uniformly computable sequence of sequences $(a^i_j)_{j\in\mathbb{N}}$ of distinct integers there is a computable element x of $[0, 1]$ that is UD random for these sequences; that is, such that $(a^i_j x)_{j\in\mathbb{N}}$ is uniformly distributed modulo one for each i.*

Proof. The lemma implies that there is a Schnorr test S such that every x outside of S, x has the desired property. In particular, there is an effectively open set G with computable measure less than $1/2$, such that any x outside of G has the desired property. Then it is not hard to compute an element x that does not land in G, as follows. First, divide $[0, 1]$ into the two intervals $[0, 1/2]$ and $[1/2, 1]$, and compute enough of G to be able to select one of the intervals on which the measure of G is at most $3/8$. Then divide that interval in two, and compute enough of G to be able to select one on which the measure of G is at most $7/32$. Continue so that at the nth state, one has a closed interval of size at most $1/2^n$ on which the measure of G is at most $1/2^n - 1/2^{2n+1}$. Let x be the limit of these nested intervals. □

3. How random is UD random?

The previous section showed that a reasonable randomness hypothesis on an element x of $[0, 1]$ is enough to ensure that x is UD random. It is natural to wonder, conversely, what randomness properties are implied by UD randomness? In Section 1 we observed that every UD-random real is normal to every base. One can ask:

- Is every UD-random real Schnorr random?
- Is every UD-random real Kurtz random?

I suspect that the answer to the second question, and hence the first question, is "no," in a very stronge sense: I suspect that there is a single Kurtz test (i.e. a null effectively closed set) C, such that for any countable collection of sequences (a^i_j) there is an element x of C such that $(a^i_j x)_{j\in\mathbb{N}}$ is uniformly distributed modulo one for every i. The conjectures below suggest particular Kurtz tests that might have this property. I also speculate on a way of constructing such a test that might be easier than proving any of the conjectures.

The failure of Kurtz randomness is striking, since it can be expressed in terms of observable properties. That is, the statement that a real number x is in an effectively closed set C is a universal property, which means that if x is *not* in C, then one can verify this fact by carrying out a finite computation on finitely many bits of x. Thus, if the conjecture above holds, there are real numbers x that are UD random, but, at any finite level of accuracy, can be seen to satisfy a distinctly non-random property.

For example, for almost every x, the binary digits of x will satisfy the law of the iterated logarithm. Thus one can ask:

- Are there a UD-random x, a rational $\varepsilon > 0$, and a m such that for every $n \geq m$, if u_n is excess of 1's over 0's in the first n digits of x, we have $u_n \leq (1 - \varepsilon)\sqrt{2n \log(\log n)}$?

Even more strikingly:

- Is there a UD-random x such that every initial segment of the binary representation of x has at least as many 0's as 1's?

I suspect that the answer is "yes," even when UD randomness is defined with respect to any countable sequence (a^i_j) of sequences.

With sequences (a_n) of distinct integers, it does not seem that the requirement that the sequences $(a_n x)$ be uniformly distributed modulo one can place serious requirements on the relationship between bits of x that are not close by.

- If (b_j) is a sufficiently fast-growing sequence of integers, is there necessarily a UD-random x with the property that for every j, bits b_{2j} and b_{2j+1} of the binary representation of x are equal?

Notice that for such an x, the sequence $x_{b_0} x_{b_1} x_{b_2} x_{b_3} \ldots$ fails to be normal.

It is interesting to compare the notion of UD randomness to Church stochasticity. Roughly speaking, a real number x is Church stochastic if the zeroes in any subsequence of digits of the binary representation of x obtained by a computable "selection procedure" has a limiting density of one half. The notion is, in a sense, orthogonal to UD randomness: by a suitable choice of the sequence (a_n), a test for UD randomness can sample bits in any order; but, to compensate, a computable selection procedure is allowed to see the previous bits of x before deciding whether or not to select the next bit. Ville's theorem [8] (see also [2, Section 6.5]) shows that there are real numbers x that are Church stochastic (in fact, with respect to any countable collection of selection procedures) but have the property that every initial segment of the binary representation of x has at least as many 1's as 0's. Wang [9] has shown that there are numbers x that are Schnorr random but not Church stochastic (see also [2, Section 8.4]). By Theorem 2.1, this implies that UD randomness does not imply Church stochasticity. One can ask about the converse direction:

- If there a real number x that is Church stochastic but not UD random?

I will close by sketching a strategy that might be used to show that there are reals that are UD random but not Kurtz random. In Section 2, we saw that we can define sets

$$G = \bigcup_j A_{i_j, \varepsilon_j, m_j}$$

where G has arbitrarily small measure and any x outside of G is UD random. Here,

$$A_{i,\varepsilon,m} = \{x \mid \exists n \geq m \; |S^i_{n^2}(x)| > \varepsilon\}$$

is the set defined in the proof of Lemma 2.3. Suppose we can produce an effective decreasing sequence $C_0 \supset C_1 \supset C_2 \supset \ldots$ of closed sets whose measure decreases to 0, such that for every u, $C_u \setminus G$ is nonempty. Then $\bigcap_u C_u$ is a Kurtz null set, and, by compactness, it will contain an element in the complement of G.

The problem, of course, is that for large u, C_u is very small, while G has some fixed size. Thus, if we are not careful, C_u may be entirely contained in G. The proof of Lemma 2.3 showed that G is small by showing that the sets $A_{i,\varepsilon,m}$ are small; and it obtained this from the fact that the $L_2([0,1])$-norms of the functions $S^i_n(x)$ are sufficiently small. What we need is to be able to show that for any u, $G \bigcap C_u$ is small *relative to* C_u; that is, the fact that C_u doesn't have an inordinately large proportion of elements of G. That can be achieved by showing that the $L_2(C_u)$-norm of the restriction of $S^i_n(x)$ to C_u is small relative to C_u, which is to say, that the integrals $\int_{C_u} |S^i_n(x)|^2$ are sufficiently small relative to $\lambda(C_u)$. In other words, we

want to know that an integral that is small on $[0, 1]$ is correspondingly small on C_u, which is to say, C_u looks like a "random" subset $[0, 1]$ with regard to this property.

One way to ensure this is to contrive things so that the functions $x \mapsto e(hx)$ have small integrals over C_u, which is to say, C_u doesn't have any structure that is easily picked out by the periodic functions $e(hx)$. We can pull the problem back to asking for finite sets of natural numbers $A \subset \{0, \ldots, n-1\}$ with such a property by dividing the unit interval into n pieces and letting

$$C_u = \left\{ x \in \left[\frac{i}{n}, \frac{i+1}{n} \right] \mid i \in A \right\}.$$

Making C_u small amounts to making A small, and making the norms of the functions $e(hx)$ small over C_u corresponds to ensuring that that A is *pseudorandom*, in the sense of additive combinatorics [7, Chapter 4]. In fact, "most" subsets A of n will have the latter property (see [7, Theorem 4.16] for a precise statement), and one can obtain an explicit such A, with density roughly $1/2$, by taking n to be prime and taking A to be the quadratic residues modulo n (see [7, Theorem 4.14]).

In sum, if A is sufficiently pseudorandom and C_u is defined as above, one can show that the norm of S_n^i restricted to C_u is small relative to the measure of C_u. (If n is large, this will hold because most of the functions $x \mapsto e(a_j x)$ occuring in S_n^i will have high frequency with respect to the intervals in C_u; when n is small, pseudorandomness takes over.) As a result, we can ensure that C_u has a nontrivial intersection with G.

The problem is that this only produces a *single* set C_u, and we need an infinite descending sequence of sets C_u with the property just described, whose measures decrease to 0. I do not know how to produce such a sequence.

References

[1] Bernard Chazelle. *The discrepancy method*. Cambridge University Press, Cambridge, 2000.

[2] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity*.

[3] Glyn Harman. *Metric number theory*. Oxford University Press, New York, 1998.

[4] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Wiley-Interscience, New York, 1974.

[5] André Nies. *Computability and randomness*. Oxford University Press, Oxford, 2009.

[6] Joseph M. Rosenblatt and Máté Wierdl. Pointwise ergodic theorems via harmonic analysis. In K. Petersen and I. Salama, eds., *Ergodic theory and its connections with harmonic analysis*. Cambridge University Press, Cambridge, 1995.

[7] Terence Tao and Van Vu. *Additive combinatorics*. Cambridge University Press, Cambridge, 2006.

[8] J. Ville. *Étude critique de la notion de collectif*. Gauthier-Villars, Paris, 1939.

[9] Y. Wang. *Randomness and complexity*. PhD thesis, Heidelberg, 1996.

[10] Hermann Weyl. Über die gleichverteilung von zahlen mod. eins. *Mathematische Annalen*, 77(3):313–352, 1916.

Departments of Philosophy and Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213