



## Natural halting probabilities, partial randomness, and zeta functions

Cristian S. Calude<sup>a,\*</sup>, Michael A. Stay<sup>b</sup>

<sup>a</sup>*Department of Computer Science, University of Auckland, New Zealand*

<sup>b</sup>*Department of Mathematics, University of California Riverside, USA*

Received 23 February 2006; revised 12 July 2006

Available online 25 September 2006

---

### Abstract

We introduce the *zeta number*, *natural halting probability*, and *natural complexity* of a Turing machine and we relate them to Chaitin's Omega number, halting probability, and program-size complexity. A classification of Turing machines according to their zeta numbers is proposed: divergent, convergent, and tuatara. We prove the existence of universal convergent and tuatara machines. Various results on (algorithmic) randomness and partial randomness are proved. For example, we show that the zeta number of a universal tuatara machine is c.e. and random. A new type of partial randomness, asymptotic randomness, is introduced. Finally we show that in contrast to classical (algorithmic) randomness—which cannot be naturally characterised in terms of plain complexity—asymptotic randomness admits such a characterisation.

© 2006 Elsevier Inc. All rights reserved.

*Keywords:* Halting probability; Omega number; Zeta function; Zeta number; Turing machine; Tuatara machine; Random real

---

---

\* Corresponding author. Fax: +9 3737 453.

*E-mail addresses:* [cristian@cs.auckland.ac.nz](mailto:cristian@cs.auckland.ac.nz) (C.S. Calude), [mike@math.ucr.edu](mailto:mike@math.ucr.edu) (M.A. Stay).

## 1. Introduction

We introduce the *zeta number*, *natural halting probability*, and *natural complexity* of a Turing machine and we relate them to Chaitin’s Omega number, halting probability, and program-size complexity. A classification of Turing machines according to their zeta numbers is proposed: divergent (zeta number is infinite), convergent (zeta number is finite), and tuatara (zeta number is less than or equal to one). Every self-delimiting Turing machine is tuatara, but the converse is not true. Also, there exist universal convergent and tuatara machines; there is a tuatara machine universal for the class of convergent machines.

The zeta number of a universal self-delimiting Turing machines is c.e. and (algorithmically) random, and for each tuatara machine there effectively exists a self-delimiting Turing machine whose Chaitin halting probability equals its zeta number; if the tuatara machine is universal, then the self-delimiting Turing machine can also be taken to be universal.

For each self-delimiting Turing machine there is a tuatara machine whose zeta number is exactly the Chaitin halting probability of the self-delimiting Turing machine; it is an open problem whether the tuatara machine can be chosen to be a universal self-delimiting Turing machine in the case when the original machine is universal.

Let  $s > 1$  be a computable real,  $T$  a universal Turing machine, and  $K_T$  be the plain complexity induced by  $T$ . In analogy with the notion of Chaitin partially random reals we introduce the notion of a “ $1/s - K$ -random real” (a real  $\alpha = 0.x_1 \cdots x_m \cdots$  such that the prefixes of its binary expansion are  $1/s - K$ -random, i.e.,  $K_T(x_1 \cdots x_m) \geq m/s - c$ , for some  $c \geq 0$  and all  $m \geq 1$ ) as well as the notion of an “asymptotically random real” ( $1/t - K$ -random real, for every computable  $t > s > 1$ ).

The result due to Chaitin and Martin-Löf showing that there is no infinite sequence whose prefixes have all *maximal*  $K$  complexity (also true for  $H$  complexity) is no longer true for asymptotically random reals (Theorem 40 and Theorem 53). However,  $1/s - K$ -randomness is different from Chaitin  $1/s$ -randomness (Proposition 38). Every c.e. random number is asymptotically random (Theorem 48), but the converse implication fails to be true: there exists a self-delimiting Turing machine whose zeta number is asymptotically random, but not random (Theorem 50).

Various examples illustrate the above notions and results. Some open problems conclude the paper.

## 2. Omega and zeta numbers

It is well-known that the Halting Problem, i.e., the problem of deciding whether an arbitrary Turing machine halts or not on a given input, is Turing uncomputable. The probabilistic version of the Halting Problem, first studied by Chaitin [7,8], deals with the halting probability, i.e., the probability that an arbitrary Turing machine halts on a randomly chosen input. Chaitin’s halting probability was studied intensively by various authors (see [23,20,2,11]). Chaitin’s halting probability is not defined for every Turing machine, hence Chaitin and his followers have worked with a sub-class of Turing machines which has equal enumeration power as the class of all Turing machines, namely the self-delimiting Turing machines.

A *self-delimiting Turing machine*  $C$  is a Turing machine which processes binary strings into binary strings and has a *prefix-free* domain; that is, if  $C(x)$  halts (is defined) and  $y$  is either a proper prefix

or a proper extension of  $x$ , then  $C(y)$  is not defined. The domain of  $C$ ,  $\text{dom}(C)$ , is the set of strings on which  $C$  halts (is defined).

**Definition 1** (*Chaitin’s Omega number*). The halting probability (Omega number) of a self-delimiting Turing machine  $C$  is

$$\Omega_{\text{dom}(C)} = \sum_{p \in \text{dom}(C)} 2^{-|p|}.$$

The number  $\Omega_{\text{dom}(C)}$ , usually written  $\Omega_C$ , is a halting probability. Indeed, pick, at random using the Lebesgue measure on  $[0, 1]$ , a real  $\alpha$  in the unit interval and note that the probability that some initial prefix of the binary expansion of  $\alpha$  lies in the prefix-free set  $\text{dom}(C)$  is exactly  $\Omega_C$ .

More formally, let  $\Sigma = \{0, 1\}$  and let  $\Sigma^*, \Sigma^\omega$  be the set of binary strings and infinite binary sequences, respectively. For  $A \subseteq \Sigma^*, A\Sigma^\omega = \{w\mathbf{x} \mid w \in A, \mathbf{x} \in \Sigma^\omega\}$ , the cylinder induced by  $A$ , is the set of sequences having a prefix in  $A$ . The sets  $A\Sigma^\omega$  are the open sets in the natural topology on  $\Sigma^\omega$ . Let  $\mu$  denote the usual product measure on  $\Sigma^\omega$  given by the uniform distribution  $\mu(\{0\}\Sigma^\omega) = \mu(\{1\}\Sigma^\omega) = 2^{-1}$ . For a measurable set  $\mathbf{C}$  of infinite sequences,  $\mu(\mathbf{C})$  is the probability that  $\mathbf{x} \in \mathbf{C}$  when  $\mathbf{x}$  is chosen by a random experiment in which an ‘independent toss of a fair coin’ is used to decide whether  $x_n = 1$ . If  $A$  is prefix-free, then  $\mu(A\Sigma^\omega) = \sum_{w \in A} 2^{-|w|} = \Omega_A$ ; here  $|w|$  is the length of the string  $w$ . We assume everywhere that  $\min \emptyset = \infty$ . For more details see [2,11].

Let  $\alpha = 0.x_1x_2 \cdots x_n \cdots \in [0, 1]$  with  $x_i \in \{0, 1\}$ , and let  $x_1x_2 \cdots x_n \cdots$  be the unending binary expansion of  $\alpha$ . We put  $\alpha[n] = x_1x_2 \cdots x_n$ . If  $y = y_1y_2 \cdots y_n \cdots$ , then  $0.y = \sum_{i=1}^n y_i 2^{-i}$ .

**Definition 2.** The Turing machine  $U$  is *universal* for a class  $\mathfrak{R}$  of Turing machines if for every Turing machine  $C \in \mathfrak{R}$  there exists a fixed constant  $c \geq 0$  (depending upon  $U$  and  $C$ ) such that for every  $x \in \text{dom}(C)$  there is a string  $p_x \in \text{dom}(U)$  with  $|p_x| \leq |x| + c$  and  $U(p_x) = C(x)$ . In case  $U \in \mathfrak{R}$ , we simply say that *the machine*  $U \in \mathfrak{R}$  is *universal*.

A classical result states:

**Theorem 3** [7]. *We can effectively construct a universal self-delimiting Turing machine.*

**Definition 4.** (a) The *plain complexity* of the string  $x \in \Sigma^*$  with respect to a Turing machine  $M$  is  $K_M(x) = \min\{|w| \mid w \in \Sigma^*, M(w) = x\}$ .

(b) The *program-size complexity* of the string  $x \in \Sigma^*$  with respect to a self-delimiting Turing machine  $C$  is  $H_C(x) = \min\{|w| \mid w \in \Sigma^*, C(w) = x\}$ .

**Definition 5.** (a) [22] A real  $\alpha \in (0, 1)$  is *computably enumerable (c.e.)* if it is the limit of an increasing computable sequences of rationals.

(b) ([30,5]) Let  $\varepsilon$  be a computable real and  $U$  a universal self-delimiting Turing machine. A real  $\alpha \in (0, 1)$  is *Chaitin  $\varepsilon$ -random* if there is a constant  $c$  such that for each  $n \geq 1$ ,  $H_U(\alpha[n]) \geq \varepsilon \cdot n - c$ . We say that  $\alpha$  is *Chaitin partially random* if it is *Chaitin  $\varepsilon$ -random* for some computable real  $1 > \varepsilon > 0$ .

(c) [7] A real  $\alpha \in (0, 1)$  is (*algorithmically*) *random* if it is 1-random, i.e., there exists  $c \geq 0$  such that for all  $m \geq 1$ ,  $H_U(\alpha[m]) \geq m - c$ .

The following theorem gives a full characterisation of c.e. and random reals:

**Theorem 6** ([4, 14, 2]). *A real  $\alpha \in (0, 1)$  is c.e. and random iff there exists a universal self-delimiting Turing machine  $U$  such that  $\alpha = \Omega_U$ .*

The definition of Chaitin’s halting probability allows an apparent “ambiguity” as strings with the same length in the domain of the self-delimiting Turing machine contribute equally towards the halting probability.<sup>1</sup> This motivates us to introduce a slightly different “halting probability” in which different strings in the domain of the machine have different contributions to the “halting probability”.

Let  $\mathbf{N} = \{1, 2, \dots\}$  and let  $\text{bin} : \mathbf{N} \rightarrow \Sigma^*$  be the bijection which associates to every  $n \geq 1$  its binary expansion without the leading 1,

$n$	$n_2$	$\text{bin}(n)$	$ \text{bin}(n) $
1	1	$\lambda$	0
2	10	0	1
3	11	1	1
4	100	00	2
$\vdots$	$\vdots$	$\vdots$	$\vdots$

If  $A \subset \Sigma^*$ , then we define  $\Upsilon[A] = \{n \in \mathbf{N} \mid \text{bin}(n) \in A\}$ . In other terms, the binary expansion of  $n$  is  $n_2 = 1\text{bin}(n)$ .

**Definition 7** (*Zeta number of a Turing machine*). The zeta number of the Turing machine  $M$ , denoted  $\zeta_M$ , is

$$\zeta_M = \sum_{n \in \Upsilon[\text{dom}(M)]} \frac{1}{n}.$$

The number  $\zeta_M$  will be shown to be random in the same sense as  $\Omega_M$  in case  $M$  is ‘universal’ (for example, if  $M$  is a universal self-delimiting Turing machine, Theorem 13).

One might ask whether there is also some sense in which  $\zeta_M$  is a halting probability. For many Turing machines,  $\zeta_M$  is not a probability; for example, a total Turing machine  $M$ , i.e.,  $\text{dom}(M) = \Sigma^*$ , has  $\zeta_M = \infty$ .

However, for a universal self-delimiting Turing machine  $M$ ,  $\zeta_M$  is a halting probability. Here is an informal argument. In an alphabet with  $k$  symbols, the probability that the  $k$ -ary expansion of  $n$  appears is proportional to  $k^{-\lfloor \log_k n \rfloor - 1}$ , while the measure assigned to  $n$  in the definition of  $\zeta_M$  is  $k^{-\log_k n}$ . By letting  $k$  approach 1 from above, we can eliminate the roughness in the measure due to the least integer function. Fractional bases  $k$  correspond to strings in base  $\lceil k \rceil$  with restrictions. For instance, using the golden ratio  $\phi = \frac{1+\sqrt{5}}{2} \approx 1.618$  as a base, we get the “Fibonacci” [16] expansion. Here, numbers are represented by binary strings in which consecutive 1 digits are prohibited. As  $k$  approaches 1, the measure of  $n$  approaches  $1/n$ .

<sup>1</sup> The “ambiguity” is apparent because from the first  $n$  bits of  $\Omega_U$  we effectively calculate the strings in  $\text{dom}(U)$  that determine these digits.

**Definition 8** (*Zeta classification of Turing machines*). According to the *zeta number*, Turing machines can be classified into the following three classes:

- *zeta divergent Turing machines*: those machines  $M$  for which  $\zeta_M = \infty$ ,
- *zeta convergent Turing machines*: those machines  $M$  for which  $\zeta_M < \infty$ ,
- *tuatara machines*<sup>2</sup>: those machines  $M$  for which  $\zeta_M \leq 1$ .

**Proposition 9.** *Every self-delimiting Turing machine is a tuatara machine. More precisely, for every self-delimiting Turing machine  $C$ ,  $\zeta_C$  is c.e. and*

$$1 \geq \Omega_C \geq \zeta_C \geq \Omega_C/2 \geq 0.$$

**Proof.** It is easy to see that  $\zeta_C$  is c.e. and

$$\begin{aligned} 1 \geq \Omega_C &\geq \sum_{n \in \Upsilon[\text{dom}(C)]} 2^{-|\text{bin}(n)|} \\ &= \sum_{n \in \Upsilon[\text{dom}(C)]} 2^{-\lfloor \log_2(n) \rfloor} \geq \sum_{n \in \Upsilon[\text{dom}(C)]} 2^{-\log_2(n)} = \zeta_C \\ &\geq \sum_{n \in \Upsilon[\text{dom}(C)]} 2^{-\lfloor \log_2(n) \rfloor - 1} = \sum_{n \in \Upsilon[\text{dom}(C)]} 2^{-|\text{bin}(n)| - 1} \\ &= \Omega_C/2 \geq 0. \quad \square \end{aligned}$$

We continue with the following result [27]:

**Theorem 10.** *Let  $U$  be a universal self-delimiting Turing machine. Then,*

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \log(\#\{p \in \text{dom}(U) \mid |p| \leq n\}) = 1.$$

<sup>2</sup> We chose this name to commemorate the fact that the work was done in New Zealand. Tuatara (“peaks on the back” in Maori) is a reptile (not a lizard) found only in New Zealand. Tuatara is the last remaining member of the ancient group of reptiles *Sphenodontia*, the only survivor of a large group of reptiles that roamed the earth at the time of dinosaurs. Tuatara has not changed its form much in over 225 million years!



Its relatives died out about 60 million years ago. Tuatara has a ‘third eye’; its main role is to soak up ultraviolet rays in the first few months of life. See more in [31].

**Proof.** If  $M$  is a one-to-one (as a partial function), self-delimiting Turing machine, then in view of the universality of  $U$  we have:  $\#\{q \in \text{dom}(M) \mid |q| \leq n - c\} \leq \#\{p \in \text{dom}(U) \mid |p| \leq n\}$ . To obtain the formula in the statement of the theorem we can choose  $M$  such that  $\text{dom}(M) = \mathbb{L}$ , the Lukasiewicz language defined by the equation  $\mathbb{L} = 0 \cup 1 \cdot \mathbb{L}^2$  (see [15]); so, for every odd  $n$  we have

$$\#\{q \in \text{dom}(M) \mid |q| \leq n\} = \sum_{\text{odd } i}^n \frac{1}{2i+1} \binom{2i+1}{i} = \sum_{\text{odd } i}^n C_i,$$

where  $C_i$  is the  $i$ th Catalan number (see [15]).  $\square$

**Fact 11.** *The domain of a universal self-delimiting Turing machine  $U$  cannot be a set of strings such that every element has a length that is an integer power of two.*

**Proof.** The result follows from Theorem 10: otherwise,  $\liminf_{n \rightarrow \infty} \frac{1}{n} \log(\#\{p \in \text{dom}(U) \mid |p| \leq n\}) \leq 1/2$ .  $\square$

**Corollary 12.** *For every universal self-delimiting Turing machine  $U$ ,  $1 > \Omega_U > \zeta_U > \Omega_U/2 > 0$ .*

**Proof.** We have:  $2^{-\lfloor \log_2(n) \rfloor} \geq 2^{-\log_2(n)} > 2^{-\lfloor \log_2(n) \rfloor - 1}$ , where equality holds only when  $n$  is a power of two, so the strict inequalities hold true because of Proposition 9 and Fact 11.  $\square$

**Theorem 13.** *The zeta number  $\zeta_U$  of a universal self-delimiting Turing machine  $U$  is random.*

**Proof.** We define the machine  $C$  as follows: on a string  $w$ ,  $C$  will try to compute  $U(w) = y$ , then continue by enumerating enough elements  $\text{bin}(n_1), \text{bin}(n_2), \dots, \text{bin}(n_k) \in \text{dom}(U)$  such that  $\sum_{i=1}^k 1/n_i > 0.y$  and output  $C(w) = \text{bin}(j)$ , where  $j$  is the minimum positive integer not in the set  $\{n_i \mid 1 \leq i \leq k\}$ . If the computation  $U(w)$  does not halt or the enumeration fails to satisfy the above inequality, then  $C(w)$  is undefined.

First we note that  $C$  is a self-delimiting Turing machine as  $\text{dom}(C) \subset \text{dom}(U)$ . Second, if  $C(w)$  is defined and  $U(w') = U(w)$  with  $|w'| = H_U(U(w))$ , then  $C(w) = C(w')$ , hence

$$H_C(C(w)) \leq |w'| \leq H_U(U(w)). \tag{1}$$

Third, because  $U$  is universal,  $H_U(x) \leq H_C(x) + \text{const}_C$ , for some  $\text{const}_C$  and all strings  $x$ .

Finally,

$$\zeta_U \leq 0.\zeta_U[m+1] + 2^{-m-1}. \tag{2}$$

Given  $U(w) = \zeta_U[m+1]$  we observe that

$$H_U(C(w)) > m. \tag{3}$$

Indeed, if  $C(w) = U(\text{bin}(j))$ , in view of (2), we have

$$1/j = 2^{-\log_2(j)} > 2^{-\lfloor \log_2(j) \rfloor - 1} \geq 2^{-m-1}.$$

Using in order the inequality (3), the universality of  $U$ , and (1) we get the following inequalities:

$$\begin{aligned}
m &< H_U(C(w)) \\
&\leq H_C(C(w)) + \text{const}_C \\
&\leq H_U(U(w)) + \text{const}_C \\
&= H_U(\zeta_U[m+1]) + \text{const}_C,
\end{aligned}$$

proving that  $\zeta_U$  is random.  $\square$

It is clear that  $\Omega_M$  can be defined for every Turing machine, much in the same way as  $\zeta_M$ . Consequently, the zeta classification of Turing machines can be paralleled with:

**Definition 14** (*Omega classification of Turing machines*). According to the Chaitin (*Omega*) halting probability, Turing machines can be classified into the following three classes:

- *Omega divergent Turing machines*: those machines  $M$  for which  $\Omega_M = \infty$ .
- *Omega convergent Turing machines*: those machines  $M$  for which  $\Omega_M < \infty$ .
- *Omega Turing machines*: those machines  $M$  for which  $\Omega_M \leq 1$ .

Every self-delimiting Turing machine is an Omega Turing machine, but the converse implication is false. A natural question arises: do the zeta and Omega classifications coincide?

**Fact 15.** (a) For every Turing machine  $M$ ,  $\zeta_M < \infty$  iff  $\Omega_M < \infty$ , hence the classes of zeta divergent (convergent) Turing machines coincide. (b) If  $\Omega_M \leq 1$ , then  $\zeta_M \leq 1$ , but there exists a tuatara machine  $T$  such that  $\Omega_T > 1$ , hence the class of Omega Turing machines is strictly included in the class of tuatara machines.

**Proof.** The equivalence a) is obvious as well as the fact that for every Turing machine  $M$ ,  $\zeta_M \leq \Omega_M$ . Finally, let  $T$  be the Turing machine defined as follows:  $T(0^i1) = T(10) = 1$ , for all  $i \geq 0$ . It is easy to see that  $\Omega_T = 1 + 1/2 > 1 > \zeta_T$ .  $\square$

**Theorem 16.** For each tuatara machine  $V$  there effectively exists a self-delimiting Turing machine  $C$  such that  $\Omega_C = \zeta_V$ . If  $V$  is tuatara universal, then  $C$  can be taken to be a universal self-delimiting Turing machine.

**Proof.** A real  $\alpha \in [0, 1]$  is c.e. iff there effectively exists a self-delimiting Turing machine  $C$  such that  $\alpha = \Omega_C$  (see Theorem 7.51 in [2]). The first part of the theorem now follows because  $\zeta_V$  is c.e. (see Proposition 9).

The second part of the theorem follows from Theorem 6 and Theorem 13.  $\square$

We can prove directly Theorem 16. To this aim we need the Kraft–Chaitin Lemma, see [2]:

**Lemma 17.** Given a computable enumeration of positive integers  $n_i$  such that  $\sum_i 2^{-n_i} \leq 1$ , we can effectively construct a prefix-free set of binary strings  $\{x_i\}$  such that  $|x_i| = n_i$ .

We can now present a direct proof of Theorem 16: given a computable enumeration of positive integers  $m_i$ , we can write  $1/m_i$  as a possibly infinite sum of reciprocals of powers of 2. We can then lay these out on a grid and enumerate the non-zero elements along each diagonal. For example, given the enumeration  $\{1, 2, 3, 4, 5, 6, \dots\}$  the grid would be as follows:

$$\begin{aligned}
 1/2 &= 1/2 + 0 + 0 + 0 + \dots \\
 1/3 &= 1/4 + 1/16 + 1/64 + 1/256 + \dots \\
 1/4 &= 1/4 + 0 + 0 + 0 + \dots \\
 1/5 &= 1/8 + 1/16 + 1/128 + 1/256 + \dots \\
 1/6 &= 1/8 + 1/32 + 1/128 + 1/512 + \dots \\
 &\dots
 \end{aligned}$$

The diagonal enumeration, taking diagonals from lower left to upper right, would be  $\{1/2, 1/4, 1/4, 1/16, 1/8, 1/64, 1/8, 1/16, 1/256, \dots\}$ . Since this enumeration is also computable, we can apply Lemma 17 to get a c.e. prefix-free set  $S$ .

Let  $\{m_i\}$  be an enumeration of  $\text{dom}(W)$  and derive  $S$  as above. Define  $\text{dom}(V)=S$ , hence  $\zeta_W = \Omega_V$ .  $\square$

We have seen that every self-delimiting Turing machine is a tuatara machine (Proposition 9), but the converse is not true (Fact 15, b)). Another example follows.

**Example 18.** Given a self-delimiting Turing machine  $C$  we construct a new machine  $\Pi_C$  (which we call a product machine), such that

$$\begin{aligned}
 \text{dom}(\Pi_C) &= \{p_1 p_2 \dots p_n \mid \text{bin}^{-1}(p_1) \leq \text{bin}^{-1}(p_2) \leq \dots \leq \text{bin}^{-1}(p_n), \\
 &p_i \in \text{dom}(C), 1 \leq i \leq n\},
 \end{aligned}$$

and

$$\Pi_C(p_1 p_2 \dots p_n) = C(p_1) C(p_2) \dots C(p_n).$$

Clearly,  $\Pi_C$  is not self-delimiting, but

$$0 < \Omega_{\Pi_C} = \prod_{p \in \text{dom}(C)} \frac{1}{1 - 2^{-|p|}} \leq 1.$$

**Comment.** The zeta number can be easily extended to Turing machines working on an arbitrary finite alphabet: we simply replace the computable bijection  $\text{bin}$  with the quasi-lexicographical enumeration of strings over the given alphabet (see more in [2]). Because the strings in the domain of the Turing machine do not use any of the new symbols, the new bijection maps them to a much smaller subset of the natural numbers, and every binary Turing machine becomes convergent/tuatara when thought of in the class of, say, ternary/quaternary machines.

Next we answer in the affirmative the following question: is every Omega number also a zeta number? To answer, we need two simple lemmata.

**Lemma 19.** *If  $M, M' \geq 2$  are integers, and  $q > 0$  is a rational such that  $1/M \leq q < 1/(M - 1)$  and  $1/M' \leq q - 1/M < 1/(M' - 1)$ , then  $M < M'$ .*

**Lemma 20.** *Fix an integer  $N \geq 2$ . Then, every rational can be effectively written as a finite sum of distinct unit fractions whose denominators are all greater than or equal to  $N$ .*



**Proof.** Let  $H_{i,j} = \frac{1}{i} + \frac{1}{i+1} + \dots + \frac{1}{j}$ , for  $i \leq j$ . Fix the rational  $q$ . As for every  $i \geq 1$ ,  $\lim_{j \rightarrow \infty} H_{i,j} = \infty$ , given  $N \geq 2$  we can effectively find an integer  $k \geq 0$  (depending on  $N$ ) such that

$$H_{N, N+k} \leq q < H_{N, N+k+1}.$$

Put  $q' = q - H_{N, N+k}$  and note that

$$0 \leq q' < \frac{1}{N + k + 1}. \tag{4}$$

We apply now the greedy algorithm for representing  $q'$  as an Egyptian fraction (i.e., as sum of distinct unit fractions, see [12]) and we show that the denominators of all unit fractions will be larger or equal to  $N$ . First we get an integer  $M' \geq 2$  such that

$$\frac{1}{M'} \leq q' < \frac{1}{M' - 1}, \tag{5}$$

and we note that in view of (4) and (5) we have  $M' > N + k + 1$ . We continue with the greedy algorithm

$$\frac{1}{M''} \leq q' - \frac{1}{M'} < \frac{1}{M'' - 1}, \tag{6}$$

and we apply Lemma 19 to (5) and (6) to deduce that  $M'' > M'$ . The algorithm eventually stops because the greedy algorithm always stops over the rationals as the numerator decreases at each step (it must eventually reach 1, at which point what remains is a unit fraction, and the algorithm terminates).  $\square$

**Theorem 21.** *For each self-delimiting Turing machine  $C$  there effectively exists a tuatara machine  $V$  such that  $\zeta_V = \Omega_C$ .*

**Proof.** We start with the expansion of  $\Omega_C = \sum_{i \geq 1} 2^{-|x_i|}$ , where  $x_1, x_2, \dots$  is a c.e. enumeration of  $\text{dom}(C)$  and we use Lemma 20 to produce a c.e. enumeration of non-negative distinct integers  $n_1, n_2, \dots$  from the representations as sum of distinct unit fractions of the terms  $2^{-|x_1|}, 2^{-|x_2|}, \dots$ , and finally we define  $V(\text{bin}(n_i)) = \text{bin}(n_i)$ .  $\square$

Actually, we can describe a more precise simulation of a self-delimiting Turing machine with a tuatara machine. Let  $HW(p)$  be the Hamming weight of the string  $p$ , i.e., the number of 1 bits in  $p$ .

**Theorem 22.** *Given a self-delimiting Turing machine  $C$  we can effectively construct a tuatara machine  $V$  such that  $\zeta_V = \Omega_C$ . Furthermore,  $\text{dom}(V) \supset \text{dom}(C)$ , and to each string  $p \in \text{dom}(C)$  we have  $HW(p) + 1$  strings in  $\text{dom}(V)$ ,  $p$  among them.*

**Proof.** We define the domain of the tuatara machine  $V$  to be

$$\text{dom}(V) = \bigcup_{p \in \text{dom}(C)} X(p),$$

where  $X(p)$  is the set  $\{p\} \cup \{p0^i \mid p_i = 1\}$  and  $p_i$  is the  $i$ th bit of  $p$ , numbering from the left and starting with  $i = 1$ . We note that for each  $p \in \text{dom}(V)$  with  $p_i = 1$  we have  $\text{bin}^{-1}(p0^i) = 2^i \cdot \text{bin}^{-1}(p)$ , so for every  $p \in \text{dom}(V)$  we have:

$$\sum_{x \in X(p)} \frac{1}{\text{bin}^{-1}(x)} = \frac{1}{\text{bin}^{-1}(p)} + \sum_{i=1}^{|p|} \frac{p_i}{2^i \text{bin}^{-1}(p)} = \frac{\text{bin}^{-1}(p)}{2^{|p|} \text{bin}^{-1}(p)} = 2^{-|p|}.$$

Consequently, the contribution of  $2^{-|p|}$  to  $\Omega_V$  is matched by the sum of distinct unit fractions  $\sum_{x \in X(p)} \frac{1}{\text{bin}^{-1}(x)}$ , for each  $p \in \text{dom}(C)$ , so  $\zeta_V = \Omega_C$ .<sup>3</sup> Furthermore,  $X(p)$  has  $HW(p)$  elements, and for distinct strings  $p, q \in \text{dom}(V)$ , the sets  $X(p)$  and  $X(q)$  are disjoint, hence the unit fractions derived are mutually distinct.  $\square$

**Scholium 23.** Given a universal self-delimiting Turing machine  $U$  we can effectively construct a tuatara machine  $W$  universal for all self-delimiting Turing machines such that  $\zeta_W = \Omega_U$ .

**Proof.** In case  $U = C$  is a universal self-delimiting Turing machine, the construction in the proof of Scholium 22 gives a tuatara machine  $W$  which is universal (but not self-delimiting) for the class of self-delimiting Turing machines.  $\square$

Next we turn our attention to universal convergent/tuatara machines.

**Theorem 24.** *The sets of convergent machines and tuatara machines are c.e.*

**Proof.** If  $M \geq 1$  is an integer and  $C[M] = \{T \mid T \text{ is a Turing machine with } \zeta_T \leq M\}$ , then the set of convergent machines is  $\cup_{M \geq 1} C[M]$  and  $C[1]$  is the set of tuatara machines. Standard proofs (see [2]) show that both sets are c.e.  $\square$

**Theorem 25.** *Let  $(C_i)_{i \geq 1}$  be an enumeration of tuatara machines. We define  $W(0^i 1x) = C_i(x)$ , for all  $x \in \Sigma^*$ . Then,  $W$  is a universal tuatara machine.*

**Proof.** First note that  $0^i 1 \text{bin}(n) = \text{bin}(2^{i+1+\lceil \log_2(n) \rceil} + n)$ ,  $n \geq 1$ . Now  $W$  acts as follows:

$$W(\text{bin}(2^{i+1+\lceil \log_2(n) \rceil} + n)) = C_i(\text{bin}(n)). \tag{7}$$

The machine  $W$  is universal because it can simulate any other tuatara machine with a constant prefix, and it is tuatara because:

$$\begin{aligned} \zeta_W &= \sum_{k \in \Upsilon[\text{dom}(W)]} \frac{1}{k} \\ &= \sum_{i \geq 1} \sum_{n \in \Upsilon[\text{dom}(C_i)]} \frac{1}{2^{i+1+\lceil \log_2(n) \rceil} + n} \end{aligned}$$

<sup>3</sup> For example,  $X(1011) = \{1011, 10110, 1011000, 10110000\}$ ;  $\frac{1}{\text{bin}^{-1}(1011)} = 1/27$ , and  $1/27 + 1/54 + 1/216 + 1/432 = 1/16$ .

$$\begin{aligned}
&= \sum_{i \geq 1} \sum_{n \in \Upsilon[\text{dom}(C_i)]} \frac{1}{2^{i+1} 2^{\lfloor \log_2(n) \rfloor} + 2^{\log_2(n)}} \\
&\leq \sum_{i \geq 1} \sum_{n \in \Upsilon[\text{dom}(C_i)]} \frac{1}{(2^i + 1) 2^{\log_2(n)}} \\
&= \sum_{i \geq 1} \frac{1}{2^i + 1} \cdot \sum_{n \in \Upsilon[\text{dom}(C_i)]} \frac{1}{2^{\log_2(n)}} \\
&= \sum_{i \geq 0} \frac{1}{2^i + 1} \cdot \sum_{n \in \Upsilon[\text{dom}(C_i)]} \frac{1}{n} \\
&\leq \sum_{i \geq 0} \frac{1}{2^i + 1} \cdot \zeta_{C_i} \leq 1. \quad \square
\end{aligned}$$

**Comment.** The same argument as in Theorem 25 shows that each  $C[M] = \{T \mid T \text{ is a Turing machine with } \zeta_T \leq M\}$  has a universal machine.

**Theorem 26.** *There exists a universal convergent machine; furthermore, this machine can be chosen to be tuatara.*

**Proof.** If  $(C_i^M)_{i \geq 1}$  is an enumeration of  $C[M]$ , then we define  $W(0^{J(i,M)} 1x) = C_i^M(x)$ , for all  $x \in \Sigma^*$ ; here  $J(i, M) = 2^i(2M + 1) - 1$ . In view of (7) and

$$\begin{aligned}
\zeta_W &= \sum_{k \in \Upsilon[\text{dom}(W)]} \frac{1}{k} \\
&= \sum_{i \geq 1, M \geq 1} \sum_{n \in \Upsilon[\text{dom}(C_i^M)]} \frac{1}{2^{J(i,M)+1+\lfloor \log_2(n) \rfloor} + n} \\
&\leq \sum_{i \geq 1, M \geq 1} \sum_{n \in \Upsilon[\text{dom}(C_i^M)]} \frac{1}{(2^{J(i,M)} + 1) 2^{\log_2(n)}} \\
&= \sum_{i \geq 1, M \geq 1} \frac{1}{2^{J(i,M)} + 1} \cdot \zeta_{C_i^M} \\
&\leq \sum_{i \geq 1, M \geq 1} \frac{M}{2^{2^i(2M+1)-1} + 1} \\
&\leq \sum_{i \geq 1, M \geq 1} \frac{1}{2^{2^i} + 1} \cdot \frac{M}{2^{2M} + 1} < 1/2.
\end{aligned}$$

It follows that  $W$  is tuatara and universal for the class of convergent machines.  $\square$

### 3. Natural complexity

Many properties can be elegantly expressed in terms of complexity. For example,  $U$  is a universal self-delimiting Turing machine iff for every self-delimiting Turing machine  $C$  there exists a fixed constant  $c$ , depending on  $U$  and  $C$ , such that for every string  $x \in \Sigma^*$ ,  $H_U(x) \leq H_C(x) + c$ . In this spirit we present a complexity-theoretic proof of the randomness of the zeta number of a universal tuatara machine. We need first the following definition:

**Definition 27.** [3] The natural complexity of the string  $x \in \Sigma^*$  (with respect to the tuatara machine  $V$ ) is  $\nabla_V(x) = \min\{n \geq 1 \mid V(\text{bin}(n)) = x\}$ .

**Fact 28** [3]. (a) A tuatara machine  $W$  is universal iff for every tuatara machine  $V$  there exists a constant  $\varepsilon$  (depending upon  $W$  and  $V$ ) such that  $\nabla_W(x) \leq \varepsilon \cdot \nabla_V(x)$ , for all strings  $x \in \Sigma^*$ .

(b) A real  $\alpha \in (0, 1)$  is random iff there exist a universal tuatara machine  $W$  and an  $\varepsilon > 0$  such that for all  $n \geq 1$ ,  $2^{-n} \cdot \nabla_W(\alpha[n]) \geq \varepsilon$ .

**Comment.** The natural complexity of a string  $x$  is the position in the enumeration given by  $\text{bin}$  of the ‘elegant’ program for  $x$ , denoted  $x^* = \text{bin}(\nabla_W(x))$ . The following facts follow from the definition:

- for each string  $x$ ,  $W(\text{bin}(\nabla_W(x))) = x$ ,
- for every  $j \geq 1$ , if  $W(\text{bin}(j)) = x$ , then  $\nabla_W(x) \leq j$ ,
- for each string  $x$ ,  $x^*$  is the minimal (according to the quasi-lexicographical order) input for  $W$  producing  $x$ .

**Example 29.** For the tuatara machine constructed in the proof of Theorem 25 we have:  $\nabla_W(x) \leq 2^{i+1} \cdot \nabla_{C_i}(x)$ .

**Theorem 30.** The zeta number  $\zeta_W$  of a universal convergent (tuatara) machine  $W$  is random.

**Proof.** The proof follows the same steps as the proof of Theorem 13. We define the tuatara machine  $D$  acting as follows: on a string  $w$ ,  $D$  will try to compute  $W(x) = y$ , then continue by enumerating enough elements  $\text{bin}(n_1), \text{bin}(n_2), \dots, \text{bin}(n_k) \in \text{dom}(W)$  such that  $\sum_{i=1}^k 1/n_i > 0.y$  and output  $C(w) = \text{bin}(j)$ , where  $j$  is the minimum positive integer not in the set  $\{n_i \mid 1 \leq i \leq k\}$ . If the computation  $W(x)$  does not halt or the enumeration fails to satisfy the above inequality, then  $D(x)$  is undefined.

If  $D(x)$  is defined, then  $W(x)$  is also defined, so  $D$  is tuatara. More,  $W(x) = W(x^*)$ , where  $x^* = \text{bin}(\nabla_W(x))$ . It follows that  $D(x) = D(x^*)$ , hence

$$\nabla_D(D(x)) \leq \text{bin}^{-1}(x^*) = \nabla_W(W(x)). \quad (8)$$

By universality of  $W$  we get a constant  $\varepsilon_D > 0$  such that for all strings  $x$ ,

$$\nabla_W(x) \leq \varepsilon_D \cdot \nabla_D(x). \quad (9)$$

Next we show that if  $W(x) = \zeta_W[m]$ , then

$$\nabla_W(D(x)) > 2^m. \quad (10)$$

Indeed, from  $\nabla_W(D(x)) \leq 2^m$  it follows that if  $W(\text{bin}(j)) = x$ , then  $1/j$  contributes towards  $\zeta_W$ , so it has to be no larger than  $2^m$ .

Using in order the inequalities (10), (9), and (8) we get the following inequalities:

$$\begin{aligned} 2^m &< \nabla_W(D(x)) \\ &\leq \varepsilon_D \cdot \nabla_D(D(x)) \\ &\leq \varepsilon_D \cdot \nabla_W(W(x)) \\ &= \varepsilon_D \cdot \nabla_W(\zeta[m]), \end{aligned}$$

proving that  $\zeta_U$  is random.  $\square$

Chaitin considered LISP program-size complexity [10] and found that the number of characters required in a program to produce the first  $n$  bits of LISP’s halting probability was asymptotic to  $n/\log_2$  (number of characters). This is the first use we know of where an author has considered the asymptotic randomness of a string and the idea that the lower bound on the complexity of prefixes of a binary sequence might be proportional to a constant less than one times the length of the prefix. In this case, the constant comes from considering characters rather than bits.

Staiger [24,25], Tadaki [30], and Calude, Terwijn, and Staiger [5] have studied the degree of randomness of sequences or reals by measuring their “degree of compression”. Tadaki [30] studied the partial randomness of a generalisation of Chaitin’s halting probability. The lower bound on the complexity of successive prefixes of a random sequence is a line with slope 1. The lower bound for the prefixes of a partially random sequence is a line with slope  $< 1$ .

More precisely, following [30] (see also [5]), for every  $s > 0$  and universal self-delimiting Turing machine  $U$  we define the real:<sup>4</sup>

$$\Omega_U(s) = \sum_{p \in \text{dom}(U)} 2^{-s|p|}.$$

If  $0 < s < 1$ , then  $\Omega_U(s) = \infty$ .

**Theorem 31** [30]. *For every computable  $s > 1$ , the number  $\Omega_U(s)$  is Chaitin  $1/s$ -random, that is, there exists a constant  $c > 0$  such that for all  $m \geq 1$  we have:*

$$H_U(\Omega_U(s)[m]) \geq m/s - c.$$

An earlier result in algorithmic information theory states that there is no infinite sequence whose prefixes have all maximal  $K_T$  complexity (see more in [2]). To state this result more precisely we fix a universal Turing machine  $T$  and denote by  $K_T$  the induced plain complexity.

**Theorem 32** ([6,18]). *For every  $c > 0$  the set  $\{\alpha \in (0,1) \mid K_T(\alpha[m]) \geq m - c, \text{ for all } m \geq 1\}$  is empty.*

Theorem 3 has given rise to alternate definitions of random sequences with respect to the plain complexity [11] and a characterisation of random reals: *the real  $\alpha \in (0, 1)$  is random iff  $K_T(\alpha[m]) \geq$*

<sup>4</sup> Tadaki’s original notation was  $\Omega_U^D$ , where  $D = 1/s$ .

$m - K_T(\text{bin}(m)) - c$ , for all  $m \geq 1$ , [19]. We are not going to pursue this line here, but instead we will study the validity of Theorem 3 for partial randomness.

Given a computable  $s > 1$ , we will investigate reals  $\alpha^s \in (0, 1)$  for which there is constant  $c > 0$  such that for every  $m \geq 1$

$$K_T(\alpha^s[m]) \geq m/s - c. \tag{11}$$

A real  $\alpha^s$  satisfying the inequality (11) will be called “ $1/s - K$ -random”.

Random reals satisfy (11). We will investigate some other examples of “ $1/s - K$ -random” reals.

**Example 33.** Let  $T$  be a universal Turing machine and define  $M(xx) = T(x)$ , for every string  $x$ . The zeta number of  $M$  is  $1/2 - K$ -random.

One particularly simple self-delimiting Turing machine is Barker’s language Iota [1]. The simplest way to define Iota is in terms of Church’s  $\lambda$ -calculus: the universal basis  $\{S = \lambda xyz.xz(yz), K = \lambda xy.yx\}$  suffices to produce every lambda term, but for universality it is not necessary to have two combinators. There are one-combinator bases, known as *universal combinators*. Iota is a very simple universal combinator,  $\lambda f.fSK$ , denoted 0. To make Iota unambiguous, there is a prefix operator, 1, for application.

The construction is essentially a very stripped-down version of LISP with only one atom, 0; since the atom takes a single input, we can represent the open parenthesis with 1, and we note that closing parentheses are unnecessary.

**Example 34.** The zeta number of Iota is at least Chaitin  $1/193$ -random.

**Proof.** The traditional representation of  $F$  and  $T$  in combinatorial logic is  $K = F$  and  $KI = T$ . In Iota, these are represented by the strings 1010100 and 10100, respectively. We can encode bit strings as lists  $\langle \text{head}, \text{tail} \rangle$ , where the pairing operator  $\langle -, - \rangle$  is the lambda-calculus term  $P = \lambda xyz.zxy$ . In Iota, this operator is encoded by the 184-bit string

```
P=11101010100111010101001101010010101010011101010100110101001010100101
  0100111010101001101010010101010011101010100110101001101010100
  1001110101010011010100101010010011101010100110101001010100100
```

It has the property that  $1F11Px y = x$ , while  $1T11Px y = y$ , so we can extract the head and the tail of the list. We can distinguish a list from a Boolean value, so by terminating the list with  $F$ , we can know when we have read the whole string. Each bit in the list requires two applications (one to apply  $P$  to the head, and another to apply the result to the tail), the pairing operator itself, and a Boolean value. The longest this can possibly be is  $2+184+7 = 193$  bits. We can write a program that will read a bit string  $x = F^i T p$ , where  $p$  is any string of bits, and return  $C_i(p)$ , where  $C_i$  is the  $i$ th self-delimiting Turing machine in an enumeration of the set. The zeta number will be at least Chaitin  $1/193$ -random, because it takes no more than  $193n + c$  bits to output  $n$  bits of  $\Omega_U$  for any universal self-delimiting Turing machine  $U$ . Therefore the zeta number of Iota itself is at least Chaitin  $1/193$ -random.  $\square$

**Comment.** A sharper result will be presented in Example 44. There are much better encodings available in Iota than the naive one above. We conjecture that, in fact, Iota’s zeta number is “more random” (see our list of open questions at the end of this paper).

We continue with a more general construction. Tadaki’s generalization of Chaitin’s halting probability (see [30]) is a zeta function. Zeta functions appear as partition functions and in expectation values in statistical systems, and the parameter  $s$  corresponds to an inverse temperature. The partition function for a statistical system  $X$  has the form

$$Z(s) = \sum_{x \in X} e^{-sH(x)},$$

where  $H$  is the energy (Hamiltonian) of the state  $x$ , and  $s$  is inversely proportional to the temperature of the system.<sup>5</sup> An observable is a function  $\kappa : X \rightarrow \mathbf{R}$ . The average value of the observable for a system at equilibrium is

$$\langle \kappa \rangle (s) = \frac{\sum_{x \in X} \kappa(x) e^{-sH(x)}}{Z(s)}.$$

The partition function acts like a normalization constant.

Taking  $X$  to be the set of programs, we let the “energy” of a program be its length. The partition function becomes

$$Z(s) = \sum_{p \in X} e^{-s|p|}.$$

We can recover the base 2 if we let  $s = s' \ln 2$ :

$$Z(s' \ln 2) = \sum_{p \in X} 2^{-s'|p|}.$$

Taking  $X$  to be prefix-free guarantees that the partition function converges at  $s' = 1$  by the Kraft-Chaitin Lemma; however, the function converges for any subset of  $\Sigma^*$  when  $s' > 1$ . When  $X = \Sigma^*$ , the set of all binary strings,  $Z(s' \ln 2) = 1/(1 - 2^{-s'+1})$ .

We now define our observable to be the halting function of the Turing machine  $T$ :  $\kappa_T(p) = 1$  if  $T$  halts on  $p$ , 0 otherwise. The probability that a program will halt is then

$$\langle \kappa_T \rangle (s) = \frac{\sum_{p \in X} \kappa_T(p) 2^{-s'|p|}}{Z(s)} = \frac{\Omega_T(s')}{Z(s)}.$$

All of this passes over nicely to the zeta number. We let  $X = \mathbf{N}$  and let the “energy” of  $n \in \mathbf{N}$  be  $\ln n$ . The partition function becomes

$$Z(s) = \sum_{n \in \mathbf{N}} e^{-s \ln n} = \sum_{n \in \mathbf{N}} n^{-s} = \zeta(s),$$

<sup>5</sup> The notation  $H$ —used only in this motivational part—does not denote the program-size complexity, although it is not too far away from it.

the Riemann zeta function.

We define the “zeta function of  $T$ ” to be

$$\zeta_T(s) = \sum_{n \in \mathbb{N}} \kappa_T(\text{bin}(n)) \cdot n^{-s} = \sum_{n \in \Upsilon[\text{dom}(T)]} n^{-s},$$

and the probability that a program will halt on  $T$  is

$$\langle \kappa_T \rangle(s) = \frac{\zeta_T(s)}{\zeta(s)}.$$

Given a Turing machine  $M$  (which may or may not be self-delimiting), we define “the halting probability of  $M$  at  $s$ ” to be

$$\langle \kappa_M \rangle(s) = \left( \sum_{p \in \text{dom}(M)} 2^{-s|p|} \right) / \left( \sum_{q \in \mathbb{N}} 2^{-s|\text{bin}(q)|} \right) = (1 - 2^{-s+1}) \sum_{p \in \text{dom}(M)} 2^{-s|p|}.$$

**Fact 35.** For real  $s > 1$  and universal  $T$ ,  $0 < \langle \kappa_T \rangle(s) < 1$ .

**Proof.** Since  $T$  is a universal Turing machine, then there must be some integer  $q$  such that  $\text{bin}(q) \notin \text{dom}(T)$ . Therefore the numerator, which sums only over those  $q$  such that  $\text{bin}(q) \in \text{dom}(T)$ , is smaller than the denominator, which sums over all positive natural  $q$ . Since there must be at least one program that halts, the numerator is positive.  $\square$

**Theorem 36.** For every computable real  $s > 1$  and universal  $T$ ,  $\langle \kappa_T \rangle(s)$  is  $1/s - K$ -random.

**Proof.** Given the first  $m + \lfloor \log_2(1 - 2^{-s+1}) \rfloor$  bits of  $\langle \kappa_T \rangle(s)$ , we can compute the halting status of all programs  $p \in \text{dom}(T)$  such that  $|p| < m/s$ . Then, there is a computable function  $\Psi$  that, given  $\langle \kappa_T \rangle(s)[m + \lfloor \log_2(1 - 2^{-s+1}) \rfloor]$ , produces a string not in the output of those programs, hence

$$K_T(\langle \kappa_T \rangle(s)[m]) \geq m/s - (c_\Psi + \lfloor \log_2(1 - 2^{-s+1}) \rfloor). \quad \square$$

**Comment.** (a) The number  $\langle \kappa_T \rangle(s)$  is a halting probability (see [2]). One particularly nice value is  $s = 2$  where  $\sum_{n>0} 2^{-2\lfloor \log_2(n) \rfloor} = 2$ . With reference to Example 33, if  $T = U$  is self-delimiting, then  $\Omega_M = \Omega_M(1)$  is Chaitin  $1/2$ -random

$$\Omega_M(1) = \sum_{x \in \text{dom}(M)} 2^{-|x|} = \sum_{x \in \text{dom}(U)} 2^{-2|x|} = \Omega_U(2) = 2\langle \kappa_U \rangle(2).$$

(b) Theorem 36 shows a property true for partial random reals, but not for random reals, cf. [2]. An opposite phenomenon was described in [5]. The following characterisation of random reals is no longer true for partial random reals: *A real  $\alpha \in (0, 1)$  is random iff there exist a constant  $c \geq 0$  and an infinite computable set  $M \subseteq \mathbb{N}$  such that  $H_U(\alpha[n]) \geq n - c$ , for each  $n \in M$ .*

Obviously, if  $\alpha$  is  $1/s - K$ -random, then it is also Chaitin  $1/s$ -random.



**Corollary 37.** *If  $U$  is a universal self-delimiting Turing machine, then for every computable real  $s > 1$ ,  $\langle \kappa_U \rangle(s)$  is Chaitin  $1/s$ -random.*

Note that in this case,  $\langle \kappa_U \rangle(s)$  is just a computable factor times  $\Omega_U(s)$ .

Furthermore, the converse implication is false:

**Proposition 38.** *There exists a Chaitin  $1/2$ -random real which is not  $1/2 - K$ -random.*

**Proof.** Let  $K = K_T$ , where  $T$  is a universal Turing machine, and let  $\alpha = 0.x_1x_2 \dots x_n \dots$  be Chaitin  $1$ -random. On one hand, the real  $\alpha$  is not  $1 - K$ -random, cf. [2]. On the other hand, the number  $\beta = 0.0x_10x_2 \dots 0x_n \dots$  is Chaitin  $1/2$ -random; if  $\beta$  were  $1/2 - K$ -random, then  $\alpha$  would be  $1 - K$ -random, a contradiction. Indeed, for all  $n \geq 1, K_T(0x_10x_2 \dots 0x_n) \leq K_{F \circ T}(0x_10x_2 \dots 0x_n) + c' \leq K_T(x_1x_2 \dots x_n) + c'$ , where  $F \circ T(y) = 0x_10x_2 \dots 0x_n$  whenever  $T(y) = x_1x_2 \dots x_n$ .  $\square$

**Lemma 39.** *Let  $\alpha \in (0, 1)$ . If there exist two integers  $c, N \geq 0$  and a real  $a \in (0, 1]$  such that for all  $m \geq N$  we have  $K_T(\alpha[m]) \geq a \cdot m - c$ , then we can find a constant  $b \geq 0$  such that  $K_T(\alpha[m]) \geq a \cdot m - b$ , for all  $m \geq 1$ .*

**Proof.** Put  $b = \max_{1 \leq i \leq N} \max\{0, a \cdot i - K_T(\alpha[i])\} + c$ .  $\square$

We now define a new form of partial randomness by requiring that the real is as close as we wish to being (partially) random, without necessarily being random. Following [21,26] we define the lower asymptotic complexity

$$\underline{k}(\alpha) = \liminf_{n \rightarrow \infty} \frac{1}{n} K(\alpha_1 \dots \alpha_n).$$

Following [27] we have:

**Theorem 40.** *Let  $s \geq 1$  be computable. Then, for a real  $\alpha$ , the following statements are equivalent:*

- (1) *We have:  $\underline{k}(\alpha) \geq 1/s$ .*
- (2) *For every computable real  $t > s > 1$ ,  $\alpha$  is  $1/t - K$ -random.*
- (3) *For every computable real  $t > s > 1$ ,  $\alpha$  is Chaitin  $1/t$ -random.*

**Proof.** Conditions (2) and (3) are equivalent because of the asymptotics. The equivalence with (1) can be verified by elementary calculus.  $\square$

**Definition 41.** Let  $s > 1$  be computable. We say that a real number  $\alpha \in (0, 1)$  is asymptotically  $1/s$ -random if one of the equivalent conditions in Theorem 40 is satisfied. If  $s = 1$ , then  $\alpha$  is called asymptotically random.

The notion of asymptotic  $1/s$ -randomness induces a strict hierarchy on  $s > 1$ . We need the following result (for the definition of the Hausdorff dimension see Falconer [13]):

**Theorem 42.** [21] *Let  $\dim_H$  be the Hausdorff dimension and let  $s > 1$  be computable. Then:*

$$\dim_H(\{\alpha \in [0, 1] \mid \underline{k}(\alpha) \leq 1/s\}) = \dim_H(\{\alpha \in [0, 1] \mid \underline{k}(\alpha) = 1/s\}) = 1/s.$$

In view of Theorem 42 we will refer only to asymptotic  $1/s$ -randomness (without mentioning  $K$  or  $H$ ). Consequently, using Theorems 40 and 42 we get:

**Corollary 43.** *The notion of asymptotic  $1/s$ -randomness real induces a strict hierarchy for  $s > 1$ .*

**Example 44.** The zeta number of Iota is at least  $1/194 - K$ -random and at least asymptotically  $1/193$ -random.

**Proof.** Since we know where the encoded bit string ends, Iota can simulate an arbitrary universal Turing machine, not just a self-delimiting one. For any  $s > 1$  we can print  $m$  bits of  $\zeta_U(s)$  with at most  $193m + c$  bits. So the zeta number of Iota is at least  $1/194 - K$ -random and at least asymptotically  $1/193$ -random.  $\square$

Given an arbitrary Turing machine  $M$ , we define “the natural halting probability at  $s$ ” to be

$$\langle \kappa_M^n \rangle(s) = \left( \sum_{q \in \Upsilon[\text{dom}(M)]} q^{-s} \right) / \left( \sum_{q \in \mathbb{N}} q^{-s} \right) = \zeta_M(s) / \zeta(s),$$

where we have added a superscript to  $\kappa$  to distinguish it from the Tadaki-Chaitin case.

Next, we can define the set

$$P = \{p_i \mid \text{bin}(i) \in \text{dom}(M)\},$$

where  $p_i$  is the  $i$ th prime in increasing order, and the set

$$S = \{n \mid \text{all prime factors of } n \text{ are in } P\}.$$

The set  $\text{bin}(S)$  is the domain of a Turing machine  $R(M)$  (*prime product machine*) that performs the following steps on an input  $x \in \Sigma^*$ :

1. Compute  $n = \text{bin}^{-1}(x)$ .
2. Compute the prime factors  $p_i$  of  $n$ .
3. For each  $p_i$ , simulate  $M(\text{bin}(i))$ .
4. Output the empty string.

Then,

$$\zeta_{R(M)}(s) = \sum_{n \in S} n^{-s} = \prod_{p \in P} 1/(1 - p^{-s}).$$

The definition of the Omega number works nicely for string concatenation (see the product machine, Example 18); the formula for the zeta number for this machine is complicated. In contrast, the zeta number works well for integer multiplication as in the case of the prime product machine; the formula for the Omega number for this machine is complicated.

**Theorem 45.** *For every universal Turing machine  $T$  and computable  $s > 1$ ,  $\langle \kappa_{R(T)}^n \rangle(s)$  is asymptotically  $1/s$ -random.*

**Proof.** The Prime Number Theorem implies that for  $i > 5$ ,  $i \log(i) < p_i$ . Fix a computable real  $s > 1$ . Given  $\lfloor ms \rfloor + 1$  bits of  $\langle \kappa_{R(T)}^n \rangle(s)$ , we can compute the halting status of all programs  $\text{bin}(i)$  such that  $p_i < 2^m$ . Consequently,

$$\begin{aligned} i \log(i) &< 2^m, \\ i &< 2^m / W(2^m), \\ |\text{bin}(i)| = \lfloor \log_2(i) \rfloor &< m - \log_2(W(2^m)) = W(2^m) / \ln(2). \end{aligned}$$

Here,  $W$  is the Lambert  $W$ -function, the inverse function of  $f(x) = xe^x$ , [32]; it has the series expansion  $W(x) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} x^n$ . Therefore, given  $\lfloor ms \rfloor + 1$  bits, we can compute the halting status of all programs whose lengths are each less than  $W(2^m) / \ln(2)$ . Since

$$\lim_{m \rightarrow \infty} \frac{W(2^m)}{m \ln(2)} = 1, \tag{12}$$

the result follows. Indeed, in view of (12), for each  $i > 5$  and  $\varepsilon > 0$  there exists a bound  $N_\varepsilon$  such that  $K_T(\langle \kappa_{R(T)}^n \rangle(s)[m]) \geq (1 - \varepsilon)m$ , for every  $m \geq N_\varepsilon$ , hence in view of Lemma 39, we can find a constant  $c_s \geq 0$  such that for all  $m \geq 1$ ,  $K_T(\langle \kappa_{R(T)}^n \rangle(s)[m]) \geq m/s - c_s$ .  $\square$

**Corollary 46.** *If  $U$  is a universal self-delimiting machine and  $s > 1$  is computable, then  $\langle \kappa_{R(U)}^n \rangle(s)$  is asymptotically  $1/s$ -random.*

**Lemma 47.** *Let  $\alpha \in (0, 1)$ . If there exist three integers  $c, a, N \geq 0$  such that for all  $m \geq N$  we have  $K_T(\alpha[m + c \lfloor \log_2 m \rfloor]) \geq m - a$ , then for every computable  $s > 1$  we can find a constant  $b \geq 0$  such that  $K_T(\alpha[m]) \geq m/s - b$ , for all  $m \geq 1$ .*

**Proof.** For each computable  $s > 1$  we can find a constant  $d \geq 0$  such that for all  $m \geq 1$ ,  $K_T(\alpha[m + c \lfloor \log_2 m \rfloor]) \geq \frac{1}{s}(m + c \lfloor \log_2 m \rfloor) - d$ , so the required inequality follows from Lemma 39.  $\square$

**Theorem 48.** *If  $U$  is a universal self-delimiting Turing machine, then  $\Omega_U$  is asymptotically random.*

**Proof.** Since  $\Omega_U$  is random, there exists a constant  $c \geq 0$  such that for all  $m \geq 1$   $H_U(\Omega_U[m]) \geq m - c$ . On the other hand, there exists  $a \geq 0$  such that for all  $m \geq 1$  we have  $K_T(\Omega_U[m + a \lfloor \log_2 m \rfloor]) \geq H_U(\Omega_U[m]) \geq m - c$ , hence in view of Lemma 47, for every computable  $s > 1$  there exists an integer  $b \geq 0$  such that for all  $m \geq 1$  we have:  $K_T(\Omega_U[m]) \geq m/s - b$ . This shows that  $\Omega_U$  is asymptotically random.  $\square$

**Corollary 49.** *If  $U$  is a universal self-delimiting machine, then  $\zeta_U$  is asymptotically random.*

**Proof.** Use Theorem 48 and Scholium 23.  $\square$

The converse implication fails to be true:

**Theorem 50.** *There is a self-delimiting Turing machine  $V$  such that  $\zeta_V$  is asymptotically random, but not random.*

**Proof.** Let  $\bar{p}$  be a self-delimiting version of the string  $p$  such that  $|\bar{p}| \approx |p| + 2 \log_2 |p|$  (see for example [2]). Let  $(C_i)$  be a c.e. enumeration of all self-delimiting Turing machines and define  $V(0^i \bar{p}) =$

$C_i(p)$ . Clearly, there is a constant  $c \geq 0$  such that for all  $m \geq 1$ ,  $K_T(\zeta_V[m + 2\lceil \log_2 m \rceil]) \geq m - c$ , so in view of Lemma 47,  $\zeta_V$  is asymptotically random. However,  $V$  is not universal, so  $\zeta_V$  is not random.  $\square$

**Comment.** A different proof for Theorem 50 can be obtained using a non-sparse dilution, cf. Example 3.18 in [24] or Theorem 4.3 in [17].

**Corollary 51.** *There is a self-delimiting Turing machine  $V$  such that  $\zeta_V$  is asymptotically random, but not random.*

**Comment.** If  $x_1x_2 \dots$  is a random sequence, then the sequence

$$x_10^{\lfloor \log_2 1 \rfloor} x_20^{\lfloor \log_2 2 \rfloor - \lfloor \log_2 1 \rfloor} \dots x_n0^{\lfloor \log_2 n \rfloor - \lfloor \log_2 (n-1) \rfloor} \dots$$

is not random, but asymptotically random.

**Lemma 52.** *For every pair of computable reals  $r, t > 1$  and integer  $c \geq 1$  there exists a computable real  $s > 1$  such that for every  $m \geq 1$  we have:  $(\frac{1}{s} - \frac{1}{r}) \cdot m \geq \frac{c}{t} \cdot \log_2 m$ .*

**Proof.** Take  $\frac{1}{s} = \frac{c}{t} + \frac{1}{r}$ .  $\square$

Theorem 3 proves that there is no infinite sequence whose prefixes have all maximal  $K_T$  complexity. A similar result can be proved for program-size complexity  $H_U$ . However, this result will be false for asymptotic randomness.

**Theorem 53.** *There exists a real  $\alpha \in (0, 1)$  such that for every pair of computable reals  $r, t > 1$  and integer  $c \geq 1$ , there exists an integer  $b \geq 1$  such that for every  $m \geq 1$ ,*

$$H_U(\alpha[m]) \geq \frac{1}{r} \cdot m + \frac{c}{t} \cdot \log_2 m - b.$$

So,  $H_U(\alpha[m])$  is as close as we want, but never equal, to  $\max_{|x|=m} H_U(x) - O(1)$ .

**Proof.** Take an asymptotically random real  $\alpha$ , consequently, for every computable  $s > 1$  there is a constant  $a \geq 0$  such that  $H_U(\alpha[m]) \geq \frac{1}{s} \cdot m - a$ , for all  $m \geq 1$ , and then use Lemma 52.  $\square$

#### 4. Open problems

Many interesting questions remain unsolved. For example, can the machine  $V$  in Scholium 23 be taken to be universal self-delimiting or universal tuatara?

The zeta number of Iota is at least  $1/194 - K$ -random and at least asymptotically  $1/193$ -random (Example 44); we conjecture that natural halting probability of Iota is asymptotically random, but not random.

Let  $U^K$  is a universal self-delimiting machine with an oracle to the Halting Problem, and  $\Omega^K = \Omega_{U^K}$ ;  $\Omega^K(2)$  is Chaitin  $1/2 - 2$ -random. Is  $\Omega^K(2)$  random or asymptotically random?

## Acknowledgments

We thank Greg Chaitin, André Nies, and the anonymous referees for many useful comments and references, Nick Hay for interesting questions, Rich Schröppel for the idea on which the proof of Lemma 20 is based, and Ludwig Staiger for suggesting Theorems 10 and 40.

## References

- [1] C. Barker, Iota and Jot: the simplest languages?, 2001, <http://ling.ucsd.edu/~barker/Iota/>.
- [2] C.S. Calude, *Information and Randomness: An Algorithmic Perspective*, Second ed., Revised and Extended, Springer-Verlag, Berlin, 2002.
- [3] C.S. Calude, M.A. Stay, From Heisenberg to Gödel via Chaitin, *Intl. J. Theor. Phys.* 44 (7) (2005) 1053–1065.
- [4] C.S. Calude, P. Hertling, B. Khoussainov, Y. Wang, Recursively enumerable reals and Chaitin  $\Omega$  numbers, *Theoret. Comput. Sci.* 255 (2001) 125–149.
- [5] C.S. Calude, L. Staiger, S.A. Terwijn, On partial randomness, *Ann. Appl. Pure Logic* 138 (2006) 20–30.
- [6] G.J. Chaitin, On the length of programs for computing finite binary sequences, *J. Assoc. Comput. Mach.* 13 (1966) 547–569 (Reprinted in: [9], 219–244).
- [7] G.J. Chaitin, A theory of program size formally identical to information theory, *J. Assoc. Comput. Mach.* 22 (1975) 329–340 (Reprinted in: [9], 113–128).
- [8] G.J. Chaitin, *Algorithmic Information Theory*, Cambridge University Press, Cambridge, 1987 (3rd printing 1990).
- [9] G.J. Chaitin, *Information, Randomness and Incompleteness*, *Papers on Algorithmic Information Theory*, Second ed., World Scientific, Singapore, 1990.
- [10] G.J. Chaitin, Lisp program-size complexity II–IV, *Appl. Math. Comput.* 52 (1992) 103–126, 172–139, 141–147.
- [11] R. Downey, D. Hirschfeldt, *Algorithmic Randomness and Complexity*, Springer, Heidelberg, to appear.
- [12] D. Epstein, Egyptian Fractions, <http://www.ics.uci.edu/~eppstein/numth/egypt/>.
- [13] K. Falconer, *Fractal Geometry. Mathematical Foundations and Applications*, Wiley & Sons, New York, 1990.
- [14] A. Kučera, T.A. Slaman, Randomness and recursive enumerability, *SIAM J. Comput.* 31 (2001) 199–211.
- [15] W. Kuich, On the entropy of context-free languages, *Information and Control* 16 (2) (1970) 173–200.
- [16] G. Lekkerkerker, Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci, *Simon Stevin* 29 (1952) 190–195.
- [17] J.H. Lutz, The dimensions of individual strings and sequences, *Inform. Comput.* 187 (2003) 49–79.
- [18] P. Martin-Löf, Complexity oscillations in infinite binary sequences, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 19 (1971) 225–230.
- [19] J. Miller, L. Yu, Contrasting plain and prefix-free complexity, ASL Meeting “Special Session on Computability and Randomness”, Stanford, March 2005, <http://www.cs.auckland.ac.nz/~nies/>.
- [20] G. Rozenberg, A. Salomaa, *Cornerstones of Undecidability*, Prentice-Hall, Englewood Cliffs, 1994.
- [21] B.Yu. Ryabko, Coding combinatorial sources and Hausdorff dimension, *Soviet Math. Dokl.* 30 (1) (1984) 219–222.
- [22] R.I. Soare, Recursion theory and Dedekind cuts, *Trans. Am. Math. Soc.* 140 (1969) 271–294.
- [23] R.M. Solovay, Draft of a paper (or series of papers) on Chaitin’s work done for the most part during the period of Sept.-Dec. 1974, unpublished manuscript, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, May 1975, 215 pp.
- [24] L. Staiger, Kolmogorov complexity and Hausdorff dimension, *Inform. Comput.* 103 (1993) 159–194.
- [25] L. Staiger, A tight upper bound on Kolmogorov complexity and uniformly optimal prediction, *Theory Comput. Syst.* 31 (1998) 215–229.
- [26] L. Staiger, The Kolmogorov complexity of real numbers, *Theoret. Comput. Sci.* 284 (2002) 455–466.
- [27] L. Staiger, Private communication to C. Calude, 21 April 2006.
- [28] M.A. Stay, Very simple Chaitin machines for concrete AIT, *Fundamenta Informaticae* 68 (3) (2005) 231–247.
- [29] M.A. Stay, *Truth and Light: Physical Algorithmic Randomness*, MSc thesis, University of Auckland, NZ, 2005.

- [30] K. Tadaki, A generalization of Chaitin's halting probability  $\Omega$  and halting self-similar sets, *Hokkaido Math. J.* 31 (2002) 219–253.
- [31] <http://www.kcc.org.nz/animals/tuatara.asp>.
- [32] E.W. Weisstein, Lambert  $W$ -Function, *MathWorld—A Wolfram Web Resource*, <http://mathworld.wolfram.com/LambertW-Function.html>.