Turing's unpublished algorithm for normal numbers

Verónica Becher[*] Santiago Figueira[*] Rafael Picchi[*]

Abstract

In an unpublished manuscript Alan Turing gave a computable construction to show that absolutely normal real numbers between 0 and 1 have Lebesgue measure 1; furthermore, he gave an algorithm for computing instances in this set. We complete his manuscript by giving full proofs and correcting minor errors. While doing this, we recreate Turing's ideas as accurately as possible. One of his original lemmas remained unproved but we have replaced it with a weaker lemma that still allows us to maintain Turing's proof idea and obtain his result.

1 Introduction

In this paper we reconstruct Alan Turing's manuscript entitled "A note on normal numbers" which remained unpublished until 1992, when it was included in the "Collected works of Alan Turing" edited by J.L. Britton [15, pp. 117–119, with notes of the editor in pp. 263–265]. The original manuscript is in Turing's archive in King's College, Cambridge, and a scanned version of it is available on the Web from `www.turingarchive.org`.

Our motivation for this work was to explore and make explicit the techniques used by Turing in relation to normal numbers, especially because there are still no known general methods to prove normality of given real numbers nor there are fast algorithms to construct absolutely normal numbers (see [3, 12, 13]).

In his manuscript Turing states two theorems here transcribed as Theorems 1 and 2. The first gives a computable construction to show that almost all real numbers are absolutely normal. A non-constructive proof of this result was given by Émile Borel in 1909 [5]. A constructive, but not effectively based proof was given by Sierpiński in 1917 [14], when computability theory was still undeveloped. Turing's and Sierpiński's constructions not only differ in terms of computability but they are based on different (though equivalent) definitions of absolute normality (see Definition 4). In modern terms, Theorem 1 proves that the set of reals in $(0, 1)$ that are not absolutely normal are included in an effectively null set, and Turing gives an explicit convergence bound for this fact.

[*] Department of Computer Science, FCEyN, University of Buenos Aires, Argentina

We denote with $\mu(A)$ the Lebesgue measure of a set $A \subseteq \mathbb{R}$ and $\mathcal{P}(A)$ is the power set of A.

Theorem 1 (Turing's first theorem)**.** *There is a computable function $c : \mathbb{N} \times \mathbb{N} \to \mathcal{P}((0,1))$ such that*

1. *$c(k,n)$ is a finite union of intervals with rational endpoints;*
2. *$c(k,n+1) \subseteq c(k,n)$;*
3. *$\mu(c(k,n)) > 1 - 1/k$.*

and for each k, $E(k) = \bigcap_n c(k,n)$ has measure $1 - 1/k$ and consists entirely of absolutely normal reals.

The function c is *computable* in the sense that given k and n we can compute $a_1 < b_1 < a_2 < b_2 < \cdots < a_m < b_m$ (m depending on k and n) such that a_i, b_i are rationals in $(0,1)$ and $c(k,n) = (a_1, b_1) \cup \cdots \cup (a_m, b_m)$. [1]

Our proof of Theorem 1 is indeed a completion of Turing's. But one of his original lemmas, a constructive version of the Strong Law of Large Numbers (see Lemma 7), remained unproved. We substituted it with a weaker version (Lemma 8) that still allows to preserve Turing's proof idea and obtain his result.

Turing's second theorem gives an affirmative answer to the then outstanding question of whether there are computable normal numbers.

Theorem 2 (Turing's second theorem)**.** *There is an algorithm that, given $k \in \mathbb{N}$ and an infinite sequence $\theta \in \{0,1\}^\infty$, produces an absolutely normal real number $\alpha \in (0,1)$ in the scale of 2. For a fixed k these numbers α form a set of Lebesgue measure at least $1 - 2/k$, and so that the first n digits of θ determine α to within 2^{-n}.*

The proof of Theorem 2 follows from the observation that there is a computable real outside the effectively null set constructed in Theorem 1, and Turing gives an explicit algorithm to compute such a number. Although Turing's strategy is mainly correct[2], a literal interpretation would not lead to the stated aim. We reinforce Turing's inductive construction with a stronger inductive hypothesis, and provide the missing correctness proof.

Both, Turing's intended algorithm and our reconstruction of it, have an explicit convergence to normality (see Remark 23). The time complexity is double exponential in n, where n is the length of the initial segment of the real number $\alpha \in (0,1)$ output by the algorithm on input n (see Remark 24).

Although nowadays it is known that there are absolutely normal numbers with lower complexity, they are still not *feasible*. A simple exponential complexity bound for computing an absolutely normal number follows from the work of Ambos-Spies, Terwjin and

[1] Turing denotes this set by $E_{c(k,n)}$.

[2] For a different appraisal on this point see in [15] the editor's note number 7 in page 119, elaborated in page 264.

Zheng [2] on reals that are random with respect to polynomial-time martingales (i.e., no polynomial-time computable martingale succeeds on such a real; for a survey see [1]). On the one hand, one can formulate a quadratic-time computable martingale which succeeds on all reals in $[0,1]$ that are not absolutely normal. Therefore, being n^2-computably random already implies being absolutely normal. On the other hand, they show that there exist n^2-computably random sequences in $\mathsf{E} = \mathsf{Deterministic\ Time}(2^{\text{linear in } n})$. Hence, one can conclude that there are absolutely normal numbers in E.

In a strong sense, the problem of giving concrete examples of absolutely normal numbers, raised by Borel in [5] as soon as he introduced the definition of normality, still remains open (see [12]). Existing examples are not fully satisfactory from a definitional perspective in the sense of Borel [7], because they are defined just by some construction, and we know no other singular properties of these numbers; we do not have any symbolic definition other than their construction method (which, as we said before, is still not feasible).

The problem of giving examples of numbers that are *normal to a given scale* has been more successfully tackled. There are fast algorithms to produce particular instances, having suitable analytic formulations in terms of series. For instance, Champernowne's number [8] and its generalization given by Copeland and Erdös [9], the Stoneham and the Korobov classes, and their recent generalization by Bailey and Crandall [3] in connection to pseudorandom generators.

For an account and references to existing work on normal numbers see Kuipers and Niederreiter [13], Harman's book [11] or his more recent article [12].

2 Definitions

Whenever possible, we keep the notation used by Turing. Let t be an integer greater than or equal to 2. The elements in $\{0, \ldots, t-1\}$ are referred to as *digits in the scale of* t. A word in the scale of t is a finite sequence of digits in the scale of t. The set of all words of length r in the scale of t is denoted by $\{0, \ldots, t-1\}^r$. The length of a word w is denoted by $|w|$. The digits of a word w are denoted by $w(i)$ for $0 \le i < |w|$. A word γ *occurs in a word* w *at position* i, $0 \le i < |\gamma|$, if $w(i)\ w(i+1)\ \ldots\ w(i+|\gamma|-1) = \gamma$. A word γ *occurs* in w if it occurs at some position.

With $\lfloor \alpha \rfloor$ and $\lceil \alpha \rceil$ we denote the floor and ceiling of a real α. For each real number α we consider the unique fractional expansion in the scale of t of the form

$$\alpha = \lfloor \alpha \rfloor + \sum_{n=1}^{\infty} a_n t^{-n}$$

where the integers a_n are in $\{0, \ldots, t-1\}$, and $a_n < t-1$ infinitely many times. this last condition over a_n is introduced to ensure a unique representation of very rational number.

We use $\#A$ for the number of elements of a set A.

Definition 3. *Let α be any real in $(0,1)$. We denote by $S(\alpha, t, \gamma, R)$ the number of occurrences of the word γ in the first R digits after the fractional point in the expansion of α*

written the scale of t:

$$S(\alpha, t, \gamma, R) = \#\{i : \alpha(i)\ \alpha(i+1)\ \ldots\ \alpha(i+|\gamma|-1) = \gamma\}.$$

Turing uses the following definition of normality given by Borel in [5, 6] as a characterizing property of absolutely normal numbers.

Definition 4. *α is* normal *in the scale of t if for every word γ in the scale of t,*

$$\lim_{R\to\infty} \frac{S(\alpha, t, \gamma, R)}{R} = \frac{1}{t^{|\gamma|}}.$$

α is absolutely normal *if it is normal to every scale $t \geq 2$.*

In [5] Borel defines normality of real numbers as follows: $\alpha \in \mathbb{R}$ is *simply normal* in the scale of t if for every digit $d \in \{0, \ldots, t-1\}$,

$$\lim_{R\to\infty} \frac{S(\alpha, t, d, R)}{R} = \frac{1}{t}.$$

α is *absolutely normal* if it is simply normal to every scale $t \geq 2$. Based just on digits instead of words, this definition of absolute normality seems weaker than that in Definition 4. A nice proof of their equivalence can be read in Harman's book [11, Theorem 1.3, pp. 5–7].

Throughout this paper we will consistently use the following convention:

Convention 5. *$R \in \mathbb{N}$ will be used for denoting the length of prefixes after the fractional point; n will be a natural number, generally between 0 and R; $t \in \mathbb{N}$, $t \geq 2$ will denote a scale; γ will denote a word in the scale of t; r will be the length of γ; $\varepsilon \in \mathbb{R}$ will denote a (small) real used to bound certain deviations from expected values.*

3 Turing's first theorem

Given $k \in \mathbb{N}$ large enough, Turing gives a uniform method to construct a set $E(k)$ of points in $(0,1)$ that are absolutely normal such that $\mu(E(k)) = 1 - 1/k$. $E(k)$ is an infinite countable intersection of certain recursively defined sets of intervals $c(k,n)$ containing the reals that are *candidates* to be absolutely normal. Given k and n, a real α is in the set $c(k,n)$ if the initial segment of the fractional expansion of α of length R expressed in each scale up to T, every word with length up to L occurs the expected number of times plus or minus εR, where R, T, L, and ε are computable functions of k and n (see Definitions 17 and 20). The sets $c(k,n)$ are defined as a finite boolean combination of intervals with rational endpoints, and they are tailored to have Lebesgue measure equal to $1 - 1/k + 1/(k+n)$.

In Turing's manuscript, the proof that the sets $c(k,n)$ have this desired measure depends on an unproved constructive version of the Strong Law of Large Numbers [3], here transcribed

[3] There is a footnote in Turing's manuscript but no text for this footnote.

as Turing's Unproved Lemma 7. This lemma gives an upper bound for the number of words of a given length for which a certain word occurs too often or too seldom. We have not been able to prove Turing's bound verbatim, but in Lemma 8 we provide an alternative bound, less sharp than Turing's but still allowing for the same construction. From his lemma Turing derives some bounds on the Lebesgue measure of some auxiliary sets of real numbers, necessary for his construction. In Propositions 14 and 16 we give our version of them.

3.1 Unproved Turing's Lemma

Definition 6. *Let t, γ and r as in Convention 5.*

1. *$S(w, \gamma)$ is the number of occurrences of γ in w;*
2. *$P(t, \gamma, n, R) = \{w \in \{0, \dots t-1\}^R \colon S(w, \gamma) = n\}$;*
3. *$N(t, \gamma, n, R) = \# P(t, \gamma, n, R)$.*

The symbolic expression of the function N is not a simple one because of the possible "overlapping" of different occurrences of γ when $|\gamma| > 1$; for instance, the word $\gamma = 00$ occurs once in 1100, twice in 1000 and three times in 0000. However, in any scale t, the symbolic expression for the function N considering the exact number of occurrences of a given *digit* is simple: The number of words of length R in the scale of t with exactly n occurrences of the *digit* d in *assigned* places is $(t-1)^{R-n}$. Hence, the number of words of length R in the scale of t with exactly n occurrences of the digit d in *some* place is

$$N(t, d, n, R) = \binom{R}{n} (t-1)^{R-n} \tag{1}$$

and of course

$$\sum_{0 \le n \le R} N(n, d, n, R) = t^R. \tag{2}$$

Unproved Turing's Lemma 7. *Let t, γ and r be as in Convention 5, and let $\delta \in \mathbb{R}$ be such that $\delta \frac{t^r}{R} < 0.3$. Then,*

$$\sum_{|n - R/t^r| > \delta} N(t, \gamma, n, R) < 2t^R e^{-\frac{\delta^2 t^r}{4R}}.$$

The rest of the present section is devoted to prove Lemma 8, our substitution of the unproved Turing's Lemma 7. The auxiliary results, Lemmas 9 and 12, appear below in this section.

Lemma 8. *Let t, γ and r be as in Convention 5 and let ε such that $6/\lfloor R/r \rfloor \le \varepsilon \le 1/t^r$. Then,*

$$\sum_{|n - R/t^r| \ge \varepsilon R} N(t, \gamma, n, R) < 2t^{R+2r-2} r \; e^{-\frac{t^r \varepsilon^2 R}{6r}}.$$

Proof. We shall fix a bijection between words of length r in the scale of t with digits in the scale of t^r corresponding to the change of scale. We write $(\gamma)_{t^r}$ to denote the digit d corresponding to γ in the scale of t^r.

Lemma 9 ensures that, for any *digit* p in the scale of t, whenever $6/R \leq \varepsilon \leq 1/t$,

$$\sum_{n \geq R/t + \varepsilon R} N(t, p, n, R) < t^R e^{-t\varepsilon^2 R/6}. \tag{3}$$

The idea is to use (3) with $\tilde{t} = t^r$, $\tilde{R} = R/r$, and the digit $d = (\gamma)_{\tilde{t}}$. By the second part of Lemma 12, which relates sums of $N(t, \gamma, n, R)$ and sums of $N(t^r, d, n, \lfloor R/r \rfloor)$, we have

$$\sum_{n \geq R/t^r + \varepsilon R} N(t, \gamma, n, R) \leq t^{r-1} r \sum_{n \geq \tilde{R}/\tilde{t} + \varepsilon \tilde{R}} N(\tilde{t}, d, n, \lfloor \tilde{R} \rfloor).$$

Since $\lfloor \tilde{R} \rfloor = \tilde{R} - x/r$ for some $x \in \{0, \ldots, r-1\}$ and since $\lfloor \tilde{R} \rfloor \leq \tilde{R}$, applying (3) we obtain

$$\begin{aligned} \sum_{n \geq R/t^r + \varepsilon R} N(t, \gamma, n, R) &\leq t^{r-1} r \sum_{n \geq \lfloor \tilde{R} \rfloor / \tilde{t} + \varepsilon \lfloor \tilde{R} \rfloor} N(\tilde{t}, d, n, \lfloor \tilde{R} \rfloor) \\ &\leq t^{r-1} r \; \tilde{t}^{\lfloor \tilde{R} \rfloor} \; e^{-\tilde{t}\varepsilon^2 \lfloor \tilde{R} \rfloor / 6} \\ &= t^{r-1} r \; \tilde{t}^{\tilde{R} - x/r} e^{-\varepsilon^2 \tilde{t} (\tilde{R} - x/r)/6} \\ &= t^{R+r-1} r \; e^{\frac{-\varepsilon^2 t^r R}{6r}} \; e^{\frac{\varepsilon^2 t^r x}{6r}} t^{-x} \\ &\leq t^{R+r-1} r \; e^{-\frac{\varepsilon^2 t^r R}{6r}} \end{aligned} \tag{4}$$

To check the last inequality observe that, since $\varepsilon \leq 1/t^r$, the expression $e^{\varepsilon^2 t^r x/(6r)} t^{-x}$ is at most 1 (indeed, $\varepsilon/(6r) \leq \ln t$ because ε is at most $1/2$ and $6r \ln t$ is at least 4).

The other sum is trickier. Lemma 9 ensures that for any digit p in the scale of t, whenever $6/R \leq \varepsilon \leq 1/t$,

$$\sum_{n \leq R/t - \varepsilon R} N(t, p, n, R) < t^R e^{-t\varepsilon^2 R/6}. \tag{5}$$

By the first part of Lemma 12 and the definitions of d, $\tilde{t}$ and $\tilde{R}$ used above, we know

$$\begin{aligned} \sum_{n \leq R/t^r - \varepsilon R} N(t, \gamma, n, R) &\leq t^{r-1} r \sum_{n \leq \tilde{R}/\tilde{t} - \varepsilon \tilde{R}} N(\tilde{t}, d, n, \lfloor \tilde{R} \rfloor) \\ &\leq t^{r-1} r \sum_{n \leq \tilde{R}/\tilde{t} - \varepsilon \tilde{R}} N(\tilde{t}, d, n, \lceil \tilde{R} \rceil). \end{aligned} \tag{6}$$

Let $R = \lfloor \tilde{R} \rfloor r + x$ where $x \in \{0, \ldots, r-1\}$. If $x \neq 0$, since $\lceil \tilde{R} \rceil = \tilde{R} + (r-x)/r$ there is $y \in \{1, \ldots, r-1\}$ such that $\lceil \tilde{R} \rceil = \tilde{R} + y/r$, and if $x = 0$ then $y = 0$ also satisfies the condition. Thus

$$\frac{\lceil \tilde{R} \rceil}{\tilde{t}} - \varepsilon \lceil \tilde{R} \rceil = \frac{\tilde{R}}{\tilde{t}} + \frac{y}{\tilde{t} r} - \varepsilon \tilde{R} - \frac{\varepsilon y}{r} \geq \frac{\tilde{R}}{\tilde{t}} - \varepsilon \tilde{R}, \tag{7}$$

where the last inequality holds because $y/(\tilde{t}r) \geq \varepsilon y/r$ when $\varepsilon \leq 1/t^r$. From (6), using (7) and (5), we get

$$\begin{aligned}
\sum_{n \leq R/t^r - \varepsilon R} N(t,\gamma,n,R) &\leq t^{r-1} r \sum_{n \leq \lceil\tilde{R}\rceil/\tilde{t} - \varepsilon\lceil\tilde{R}\rceil} N(\tilde{t},d,n,\lceil\tilde{R}\rceil) \\
&\leq t^{r-1}\tilde{t}^{\lceil\tilde{R}\rceil} r\ e^{-\tilde{t}\varepsilon^2\lceil\tilde{R}\rceil/6} \\
&= t^{R+r-1} r\ e^{-t^r\varepsilon^2\lceil\tilde{R}\rceil/6}\ t^y \\
&\leq t^{R+2r-2} r\ e^{-\frac{t^r\varepsilon^2 R}{6r}}. \qquad (8)
\end{aligned}$$

The last inequality follows from the fact that $t^y \leq t^{r-1}$ and $\lceil\tilde{R}\rceil \geq \tilde{R}$. Joining (4) and (8), we obtain the desired upper bound. □

Lemma 9 (adapted from Harman [11, page 5, Lemma 1.1]). *Let d be a digit in the scale of t, $t \geq 2$. Assuming $R > 6t$ and with ε such that $6/R \leq \varepsilon \leq 1/t$, both*

$$\sum_{n \geq R/t + \varepsilon R} N(t,d,n,R) \quad \textit{and} \quad \sum_{n \leq R/t - \varepsilon R} N(t,d,n,R) \quad \textit{are at most } t^R e^{-t\varepsilon^2 R/6}.$$

Proof. Since t, d and R are fixed, we write $N(n)$ for $N(t,d,n,R)$. Recalling from (1) the symbolic expression for $N(n)$,

$$\frac{N(n)}{N(n+1)} = \frac{(n+1)(t-1)}{R-n}. \qquad (9)$$

For all $n \leq R/t$ we have $N(n) > N(n-1)$ and for all $n > R/t$, $N(n) \leq N(n-1)$. It is not difficult to see that the quotients in (9) increase as n increases.

Let $a = R/t - \varepsilon R$ and $b = R/t + \varepsilon R$. The strategy is to "shift" the first sum to the right by $m = \lfloor \varepsilon R/2 \rfloor$ positions, and the second sum to the left by $m+1$ positions.

Let us compute the stated upper bound for the first sum. For any n

$$N(n) = \frac{N(n)}{N(n+1)} \cdot \frac{N(n+1)}{N(n+2)} \cdot \ldots \cdot \frac{N(n+m-1)}{N(n+m)} \cdot N(n+m) \qquad (10)$$

and for each i such that

$$i \leq \lfloor a \rfloor + m - 1 \qquad (11)$$

we have

$$\begin{aligned}
\frac{N(i)}{N(i+1)} &\leq \frac{N(\lfloor a \rfloor + m - 1)}{N(\lfloor a \rfloor + m)} \\
&= \frac{(\lfloor a \rfloor + m)(t-1)}{R - \lfloor a \rfloor - m + 1} \\
&< \frac{(R/t - \varepsilon R/2)(t-1)}{R - R/t + \varepsilon R/2} \\
&= 1 - \frac{\varepsilon t/2}{1 - 1/t + \varepsilon/2}.
\end{aligned}$$

Since $\varepsilon \leq 1/t$, we conclude

$$\frac{N(i)}{N(i+1)} < 1 - \varepsilon t/2 < e^{-t\varepsilon/2}. \tag{12}$$

If $n \leq a$ then $n \leq \lfloor a \rfloor$ and hence $i = n + m - 1$ satisfies condition (11). Since the greatest quotient among the ones which appear in equation (10) is the last one, we can apply (12) to each factor in (10) to obtain

$$\begin{aligned} N(n) &< e^{-t\varepsilon m/2}\, N(n+m) \\ &\leq e^{-t\varepsilon(\varepsilon R/2-1)/2}\, N(n+m) \\ &= e^{-t\varepsilon^2 R/4 + t\varepsilon/2}\, N(n+m) \\ &\leq e^{-t\varepsilon^2 R/6}\, N(n+m) \end{aligned} \tag{13}$$

where we use the definition of m, and in the last inequality (13) we have $\varepsilon^2 tR/6 \leq \varepsilon^2 tR/4 - \varepsilon t/2$, since $\varepsilon R \geq 6$. Hence by (2) we have

$$\sum_{n \leq a} N(n) < e^{-t\varepsilon^2 R/6} \sum_{n \leq a} N(n+m) \leq t^R e^{-t\varepsilon^2 R/6}.$$

To bound the second sum, we use the same strategy, but now we shift the sum to the left by $m+1$ positions. For any n,

$$N(n) = \frac{N(n)}{N(n-1)} \cdot \frac{N(n-1)}{N(n-2)} \cdot \ldots \cdot \frac{N(n-m)}{N(n-m-1)} \cdot N(n-m-1) \tag{14}$$

(with these ratios increasing as $n - i$ decreases), and for each i such that

$$i \geq \lceil b \rceil - m \tag{15}$$

we have

$$\begin{aligned} \frac{N(i)}{N(i-1)} &\leq \frac{N(\lceil b \rceil - m)}{N(\lceil b \rceil - m - 1)} \\ &= \frac{R - \lceil b \rceil + m + 1}{(\lceil b \rceil - m)(t-1)} \\ &\leq \frac{R - R/t - \varepsilon R/2 + 1}{(R/t + \varepsilon R/2)(t-1)} \\ &< 1 - \varepsilon t/3. \end{aligned}$$

The last inequality is just equivalent to $\varepsilon t - 2/t - \varepsilon < 1 - \frac{6}{\varepsilon t R}$ and since $\varepsilon t \leq 1$ and $\varepsilon > 0$ it is sufficient to prove that $1 - 2/t < 1 - \frac{6}{\varepsilon t R}$, which clearly holds for $\varepsilon > 3/R$. Therefore,

$$\frac{N(i)}{N(i-1)} < e^{-t\varepsilon/3}. \tag{16}$$

If $n \geq b$, then $n \geq \lceil b \rceil$ and hence $i = n - m$ satisfies condition (15). Since the greatest quotient among those which appear in equation (14) is the last one, we can apply (16) to each factor in (14) to obtain, as in (13)

$$N(n) < e^{-t\varepsilon(m+1)/3} N(n-m-1) \leq e^{-\frac{t\varepsilon^2 R}{6}} N(n-m-1)$$

and from this and (2),

$$\sum_{n \geq b} N(n) < t^R e^{-t\varepsilon^2 R/6}.$$

This completes the proof. □

Definition 10. *Let t, γ and r be as in Convention 5. For $j \in \{0, \ldots, r-1\}$ we define*

1. *$S_j(w, \gamma)$ as the number of occurrences of γ in w at positions of the form $r \cdot q + j$ (i.e. congruent to j modulo r);*

2. *$P_j(t, \gamma, n, R) = \{w \in \{0 \ldots t-1\}^R \colon S_j(w, \gamma) = n\}$.*

Lemma 11. *Let t, γ and r as in Convention 5 and let $w \in P(t, \gamma, n, R)$. There is $j \in \{0, \ldots, r-1\}$ such that $w \in P_j(t, \gamma, m, R)$ for some $m \leq n/r$ and there is $j \in \{0, \ldots, r-1\}$ such that $w \in P_j(t, \gamma, m, R)$ for some $m \geq n/r$.*

Proof. Suppose $w \in P(t, \gamma, n, R)$, i.e., γ has n occurrences in w. For each $j \in \{0, \ldots, r-1\}$ let $n_j \geq 0$ be the number of occurrences of γ in w at positions congruent to j modulo r. Then, $w \in P_j(t, \gamma, n_j, R)$, and clearly $\sum_{0 \leq n_j \leq r-1} n_j = n$. This equality implies that $n_j \leq n/r$ for some j, and not all n_js can be strictly smaller than n/r. □

The next lemma relates sums of $N(t, \gamma, n, R)$ and sums of $N(t^r, d, n, \lfloor R/r \rfloor)$, where $d = (\gamma)_{t^r}$.

Lemma 12. *Let t, γ and r be as in Convention 5 and let d be the digit corresponding to the word γ in the scale of t^r. Then,*

$$\begin{aligned}
\sum_{n \leq a} N(t, \gamma, n, R) &\leq t^{r-1} r \sum_{m \leq a/r} N(t^r, d, m, \lfloor R/r \rfloor), \text{ and} \\
\sum_{n \geq a} N(t, \gamma, n, R) &\leq t^{r-1} r \sum_{m \geq a/r} N(t^r, d, m, \lfloor R/r \rfloor).
\end{aligned}$$

Proof. For any $j = 0, \ldots, r-1$ we define a map f_j which transforms words of length R written in the scale of t into words of length $\lfloor R/r \rfloor$ written in the scale of t^r as follows: let $w \in \{0, \ldots, t-1\}^R$ and let $k = \lfloor R/r \rfloor$. We split w into $k-1$ blocks of length r

$$\begin{aligned}
b_1 &= w(j) \ldots w(j+r-1); \\
b_2 &= w(j+r) \ldots w(j+2r-1); \\
&\vdots \\
b_{k-1} &= w(j+r(k-2)) \ldots w(j+(k-1)r-1);
\end{aligned}$$

and complete the last segment $w(j+r(k-1))\dots w(l)$, where $l = \min(j+rk-1, R-1)$, into a block b_k of length r by adding a word u in the scale of t in such a way that if $l < R-1$ (i.e., u is not the empty word), then b_k is different from γ (for example, take u to be the least such word in lexicographic order).

$$b_k = w(j + r(k-1)) \dots w(l)\ u.$$

The blocks b_i get transformed into single digits in the scale of t^r; set

$$f_j(w) = (b_1)_{t^r} \dots (b_k)_{t^r}.$$

Let us now compute the cardinality of $f_j^{-1}[f_j(w)]$. If v is another word of length R in the scale of t, then $f_j(v) = f_j(w)$ if and only if

$$v(j) \dots v(j + kr - 1) = b_1 \dots b_k.$$

Thus, v may differ from w in at most the positions $0, \dots, j-1$ and $j+kr, \dots, R-1$; their number is $R - kr - 1 \leq r - 1$. Hence, there are at most t^{r-1} words in $f_j^{-1}[f_j(w)]$. This implies that

$$\#P_j(t, \gamma, m, R) \leq t^{r-1} N(t^r, d, m, k).$$

Suppose w has exactly n occurrences of γ. By the first part of Lemma 11 we know that for all n there is $j \in \{0, \dots, r-1\}$ and $m \leq \lfloor n/r \rfloor$ such that $w \in P_j(t, \gamma, m, R)$. Therefore,

$$\begin{aligned}
\bigcup_{n \leq a} P(t, \gamma, n, R) &\subseteq \bigcup_{0 \leq j < r} \ \bigcup_{m \leq a/r} P_j(t, \gamma, m, R) \\
\sum_{n \leq a} N(t, \gamma, n, R) &= \# \bigcup_{0 \leq n \leq a} P(t, \gamma, n, R) \\
&\leq \sum_{j<r} \sum_{m \leq a/r} \#P_j(t, \gamma, m, R) \\
&\leq t^{r-1} r \sum_{m \leq a/r} N(t^r, d, m, \lfloor R/r \rfloor).
\end{aligned}$$

This completes the proof of the first part. The second part is similar, applying the last assertion in Lemma 11. □

3.2 The sets $c(k, n)$

Definition 13. *We denote by $B(\varepsilon, \gamma, t, R)$ the set of reals $\alpha \in (0,1)$ such that*

$$|S(\alpha, t, \gamma, R) - R/t^r| < \varepsilon R.$$

Proposition 14. *Let t, γ and r be as in Convention 5 and let ε and R be such that $6/\lfloor R/r \rfloor \leq \varepsilon \leq 1/t^r$. Then*

$$\mu\left(B(\varepsilon, \gamma, t, R)\right) > 1 - 2t^{2r-2} r\ e^{-\frac{t^r \varepsilon^2 R}{6r}}.$$

Proof. Let $\overline{B}(\varepsilon,\gamma,t,R) = (0,1)\setminus B(\varepsilon,\gamma,t,R)$. Observe that if a real $\alpha\in(0,1)$ belongs to $\overline{B}(\varepsilon,\gamma,t,R)$ then every real $\beta\in(0,1)$ such that the first R digits of α (written in the scale of t) coincide with the first R digits of β (written in the scale of t) also belongs to $\overline{B}(\varepsilon,\gamma,t,R)$, which means that the interval

$$[0.\alpha\restriction R\;000\ldots,\;\; 0.\alpha\restriction R\;(t-1)(t-1)(t-1)\ldots]\,.$$

of measure t^{-R} is included in $\overline{B}(\varepsilon,\gamma,t,R)$. Here $\alpha\restriction R$ denotes the first R digits of the fractional expansion of α in the scale of t. Then

$$\overline{B}(\varepsilon,\gamma,t,R) = \bigcup_{|n-R/t^r|\geq\varepsilon R}\;\bigcup_{w\in P(t,\gamma,n,R)} [0.w000\ldots,0.w(t-1)(t-1)(t-1)\ldots]$$

Since the intervals in the right hand side are disjoint for different words w, we have:

$$\mu\left(\overline{B}(\varepsilon,\gamma,t,R)\right) = t^{-R}\sum_{|n-R/t^r|\geq\varepsilon R} N(t,\gamma,n,R) < 2\;t^{2r-2}r\;e^{-\frac{t^r\varepsilon^2R}{6r}}. \qquad (17)$$

For the last equation apply Lemma 8. The proof is completed by taking complements. □

Definition 15. *For any ε, T, L and R, let*

$$A(\varepsilon,T,L,R) = \bigcap_{2\leq t\leq T}\;\bigcap_{1\leq r\leq L}\;\bigcap_{\gamma\in\{0,\ldots,t-1\}^r} B(\varepsilon,\gamma,t,R).$$

Since each $B(\varepsilon,\gamma,t,R)$ is a finite union of intervals with rational endpoints, then so is each of the sets $A(\varepsilon,T,L,R)$.

Proposition 16. *For any ε, T, L and R, such that $6/\lfloor R/L\rfloor\leq\varepsilon\leq 1/T^L$,*

$$\mu\left(A(\varepsilon,T,L,R)\right) > 1 - 2\;L\;T^{3L-1}\;e^{-\frac{\varepsilon^2R}{3L}}.$$

Proof. Let $\overline{A}$ and $\overline{B}$ denote the complements of the sets A, B, respectively, in the interval $(0,1)$.

$$\mu\left(\overline{A}(\varepsilon,T,L,R)\right) \leq \sum_{2\leq t\leq T}\;\sum_{1\leq r\leq L}\;\sum_{\gamma\in\{0,\ldots,t-1\}^r}\mu\left(\overline{B}(\varepsilon,\gamma,t,R)\right).$$

Observe that in the third summand there are t^r many γs and that

$$\sum_{2\leq t\leq T}\;\sum_{1\leq r\leq L} t^r = \sum_{2\leq t\leq T}\frac{t^{L+1}-1}{t-1}\;\leq\;T^{L+1}.$$

The upper bound for $\mu\left(\overline{B}(\varepsilon,\gamma,t,R)\right)$ in (17) yields the following uniform upper bound in terms of the present parameters ε,T,R,L:

$$\mu\left(\overline{B}(\varepsilon,\gamma,t,R)\right) < 2\;T^{2L-2}L\;e^{-\frac{2\varepsilon^2R}{3L}}$$

valid for all $2 \leq t \leq T$, $1 \leq r \leq L$ and $\gamma \in \{0, \dots, t-1\}^r$. Indeed, from $1 \leq r \leq L$, we get $2r/L \leq 2 \leq t^r$; hence, $\varepsilon^2 R/(3L) \leq \varepsilon^2 R t^r/(6r)$, which gives

$$\mu\left(\overline{B}(\varepsilon, \gamma, t, R)\right) < 2\ t^{2r-2} r\ e^{-\frac{t^r \varepsilon^2 R}{6r}} \ < \ 2\ T^{2L-2}\ e^{-\frac{\varepsilon^2 R}{3L}}.$$

Hence we obtain,

$$\mu\left(\overline{A}(\varepsilon, T, L, R)\right) < 2\ L\ T^{3L-1}\ e^{-\frac{\varepsilon^2 R}{3L}}.$$

The proof is completed by taking complements. □

We now define $A(\varepsilon, T, L, R)$ for specific values of its parameters.

Definition 17. *Let $A_k = A(\varepsilon, T, L, R)$ for $R = k$, $L = \sqrt{\ln k}/4$, $T = e^L$ and $\varepsilon = 1/T^L$.*

Proposition 18. *There is k_0 such that for all $k \geq k_0$, $\mu(A_k) \geq 1 - \frac{1}{k(k-1)}$.*

Proof. Let R, T, L and ε be the functions of k given in Definition 17. Observe that $T^L = \sqrt[16]{k}$. Since $\varepsilon \geq 6/\lfloor R/L \rfloor$ for all $k \geq 2$, the hypothesis of Proposition 16 is satisfied. We now prove that

$$2LT^{3L-1}\ e^{-\frac{\varepsilon^2 R}{3L}} \leq \frac{1}{k(k-1)}$$

for large enough k. It suffices to prove $T^{3L}k^2 \leq e^{\frac{\varepsilon^2 R}{3L}}$ because $2L \leq T$. This is equivalent to

$$1/\varepsilon^2 \cdot (9L^2 \ln T + 6L \ln k) \leq k.$$

Since $1/\varepsilon^2 = T^{2L} = \sqrt[8]{k}$, $9L^2 \ln T = (9/64)(\ln k)^{3/2}$ and $6L \ln k = (3/2)(\ln k)^{3/2}$, (17) reduces to

$$(105/64)\sqrt[8]{k}(\ln k)^{3/2} \leq k$$

which can be proved to hold for any $k \geq 1$. □

Remark 19. Observe that the assignment of Definition 17 gives initial values of T smaller than 2 and initial values of L smaller than 1. This implies that the initial intersections in

$$A_k = A(\varepsilon, T, L, R) = \bigcap_{2 \leq t \leq T} \ \bigcap_{1 \leq r \leq L} \ \bigcap_{\gamma \in \{0, \dots, t-1\}^r} B(\varepsilon, \gamma, t, R)$$

will have an empty range. However, as k increases, these variables will take greater and greater values.

One can give different assignments for $L = L(k)$, $T = T(k)$ and $\varepsilon = \varepsilon(k)$, where $\lim_k L(k) = \infty$, $\lim_k T(k) = \infty$ and $\lim_k \varepsilon(k) = 0$ and such that $L \geq 1, T \geq 2$ and Proposition 18 is verified for suitable large k.

From now on let k_0 be the value determined in Proposition 18 (or Remark 19).

Turing defines $c(k, n)$ as intersections of finitely many A_ks and he restricts these sets so that they have measure exactly $1 - 1/k + 1/(k+n)$.

Definition 20. *The computable function* $c : \mathbb{N} \times \mathbb{N} \to \mathcal{P}((0,1))$*, is defined as follows. For any* $k \geq k_0$ *let* $c(k,0) = (0,1)$ *and*

$$c(k, n+1) = A_{k+n+1} \cap c(k,n) \cap (\beta_n, 1)$$

where $(\beta_n, 1)$ *is an interval so that* $\mu\,(c(k, n+1)) = 1 - 1/k + 1/(k+n+1)$*.*

Remark 21. It is worth noting that some interval $(\beta_n, 1)$ as above always exists and it is unique. This is because

$$\mu\,(A_{k+n+1} \cap c(k,n)) \geq 1 - 1/k + 1/(k+n+1).$$

Since $c(k,n)$ and A_{k+n+1} are finite unions of intervals with rational endpoints, their respective measures are effectively computable; β_n is rational and it can be determined effectively. Hence $c(k,n)$ may be represented by a finite union of disjoint intervals $(a_1, b_1) \cup \cdots \cup (a_m, b_m)$ such that $a_i, b_i \in \mathbb{Q} \cap (0,1)$, $a_i < b_i < a_{i+1}$ and such that $(a_1, b_1, a_2, b_2, \ldots, a_m, b_m)$ is computable from k and n.

3.3 Proof of Turing's first theorem

Proof of Theorem 1. We first prove that the set $\bigcap_{k \geq k_0} A_k$ contains only absolutely normal numbers. Assume $\alpha \in \bigcap_{k \geq k_0} A_k$ and α is not normal to the scale of t. This means that

$$\lim_{R \to \infty} \frac{S(\alpha, t, \gamma, R)}{R} \neq \frac{1}{t^r}$$

for some word γ of length r in the scale of t. Hence there is $\delta > 0$ and there are infinitely many Rs such that

$$|S(\alpha, t, \gamma, R) - R/t^r| > R\delta. \tag{18}$$

Let $T(k)$, $L(k)$ and $\varepsilon(k)$ be the assignments of Definition 17 or Remark 19. Now fix $k_1 \geq k_0$ large enough such that $T(k_1) \geq t$, $L(k_1) \geq r$ and $\varepsilon(k_1) \leq \delta$. This is always possible because $T(k) \to \infty$, $L(k) \to \infty$ $\varepsilon(k) \to 0$ when $k \to \infty$. For any $k \geq k_1$, $\alpha \in A_k$, and by Definition 15, $\alpha \in B(\varepsilon(k), \gamma, t, k)$. By Definition 13 we have

$$|S(\alpha, t, \gamma, k) - k/t^r| < k\varepsilon(k) \leq k\delta.$$

for any $k \geq k_1$. Now, any $R \geq k_1$ satisfying (18) leads to a contradiction.

Clearly, conditions 1, 1 and 1 of Theorem 1 follow from the definition of $c(k,n)$ (see Definition 20). Since $E(k) \subseteq \bigcap_{i \geq k} A_i$, by the argument given above, we conclude that if $k \geq k_0$, any real number in $E(k)$ is absolutely normal. By condition 1 and the fact that $\mu\,(c(k,n)) = 1 - 1/k + 1/(k+n)$, we get

$$\mu\,(E(k)) = \lim_{n \to \infty} \mu\,(c(k,n)) \;=\; 1 - 1/k.$$

This completes the proof. □

4 Turing's second theorem

The idea of Turing's algorithm is to recursively select for each integer $n > 0$ an interval I_n with dyadic rational endpoints such that

1. $I_{n+1} \subset I_n$
2. $\mu(I_n) = 2^{-(n+1)}$
3. $\mu(E(k) \cap I_n) > 0$.

The intersection of these intervals, $\bigcap_n I_n$, contains exactly one number which must be absolutely normal. The correctness of the algorithm relies on the fact that at each stage n the measure $\mu(E(k) \cap I_n)$ is big enough to allow to proceed with the stage $n+1$ and never run out of measure (i.e. keeping $\mu(E(k) \cap I_n) > 0$). If we can do this forever, that is for all n, then we have an effective procedure to determine every digit of an absolutely normal number.

A literal reading of the algorithm that appears in Turing's manuscript does not give a correct algorithm. We reconstruct it by introducing suitable changes, but keeping the strategy. Turing uses exactly the same sets $c(k, n)$ that appear in Theorem 1 (see Definition 20), where $\mu(c(k, n)) = 1 - 1/k + 1/(k + n + 1)$. We refine them to have Lebesgue measure $1 - 1/k + 1/k2^{2n+1}$ (see Definition 22). This modification respects the strategy since for each k,

$$\lim_{n\to\infty} 1 - 1/k + 1/(k + n + 1) = \lim_{n\to\infty} 1 - 1/k + 1/k2^{2n+1} = 1 - 1/k,$$

and it still holds that $E(k) = \bigcap_{n\geq 0} c(k, n)$.

Let k_0 be as determined in Proposition 18 (or Remark 19).

Definition 22. *We redefine the computable function $c : \mathbb{N} \times \mathbb{N} \to \mathcal{P}((0, 1))$, as follows. For any $k \geq k_0$ let $c(k, 0) = (0, 1)$ and for $n > 0$*

$$c(k, n) = A_{k2^{2n+1}} \cap c(k, n - 1) \cap (\beta_n, 1);$$

where $(\beta_n, 1)$ is an interval so that $\mu(c(k, n)) = 1 - 1/k + 1/k2^{2n+1}$.

The reader may verify that it is always possible to find such a β_n for $k \geq k_0$, because $\mu(A_{k2^{2n+1}}) \geq 1 - 1/(k2^{2n+1})(k2^{2n+1} - 1)$, hence $\mu(A_{k2^{2n+1}} \cap c(k, n - 1)) > 1 - 1/k + 1/k2^{2n+1}$.

Proof of Theorem 2. The following algorithm constructs a real α in $(0, 1)$ in the scale of 2. It depends on an infinite sequence $\theta \in \{0, 1\}^\infty$ used as oracle to possibly determine some digits of α, and on a fixed parameter k large enough ($k \geq k_0$ and $k \geq 4$).

```
Start with $I_{-1} = (0,1)$.

At stage $n \geq 0$:

  – Split the interval $I_{n-1}$ into two halves $I_n^0$ and $I_n^1$. That is, say $I_{n-1} = (a_{n-1}, b_{n-1})$, then let
      $I_n^0 = \left(a_{n-1}, \frac{a_{n-1}+b_{n-1}}{2}\right)$ and $I_n^1 = \left(\frac{a_{n-1}+b_{n-1}}{2}, b_{n-1}\right)$.
  – If $\mu\left(c(k,n) \cap I_n^0\right) > 1/k2^{2n}$ and $\mu\left(c(k,n) \cap I_n^1\right) > 1/k2^{2n}$ then
      * Let $\alpha(n) = \theta(n)$.
      * Let $I_n = \begin{cases} I_n^0 & \text{if } \theta(n) = 0; \\ I_n^1 & \text{otherwise.} \end{cases}$
  – Else, if $\mu\left(c(k,n) \cap I_n^1\right) \leq 1/k2^{2n}$ then
      * Let $I_n = I_n^0$.
      * Let $\alpha(n) = 0$.
  – Else
      * Let $I_n = I_n^1$.
      * Let $\alpha(n) = 1$.
```

At each stage n, I_n is either the left half of I_{n-1} (denoted I_n^0) or the right half of it (denoted I_n^1). As we mentioned in Remark 21, $c(k,n)$ is computable. Therefore we can compute its measure, and also compute the measures of both $c(k,n) \cap I_n^0$ and $c(k,n) \cap I_n^1$. All these measures are rational numbers in $(0,1)$.

The above algorithm defines $\alpha = \bigcap_n I_n$ bit by bit, i.e. at stage n the n-th bit of α is defined. To prove that α is absolutely normal, we show $\alpha \in E(k) = \bigcap_n c(k,n)$. We prove, by induction on n, that for every $n \geq 0$,

$$\mu\left(c(k,n) \cap I_n\right) > 1/k2^{2n}. \tag{19}$$

For $n = 0$, observe that by Definition 22, $c(k,0) = (0,1)$ and then

$$\mu\left(c(k,0) \cap I_0\right) = 1/2 > 1/k.$$

For the induction, assume (19) holds. Since $c(k,n+1) \subseteq c(k,n)$ we have

$$\begin{aligned} c(k,n+1) \cap I_n &= (c(k,n) \cap I_n) \ \setminus \ ((c(k,n) \setminus c(k,n+1)) \cap I_n) \\ \mu\left(c(k,n+1) \cap I_n\right) &= \mu\left(c(k,n) \cap I_n\right) - \mu\left((c(k,n) \setminus c(k,n+1)) \cap I_n\right) \\ &\geq \mu\left(c(k,n) \cap I_n\right) - \mu\left(c(k,n) \setminus c(k,n+1)\right). \end{aligned} \tag{20}$$

Using (19) and that $\mu\left(c(k,n) \setminus c(k,n+1)\right) = 1/k2^{2n+1} - 1/k2^{2(n+1)+1}$, from (20) we obtain

$$\mu\left(c(k,n+1) \cap I_n\right) > 1/k2^{2n} - (1/k2^{2n+1} - 1/k2^{2n+3}) > 2/k2^{2(n+1)}.$$

It is impossible that both $\mu\left(c(k,n+1) \cap I^0_{n+1}\right)$ and $\mu\left(c(k,n+1) \cap I^1_{n+1}\right)$ be less than or equal to $1/k2^{2(n+1)}$. It follows that at least one of the sets $c(k,n+1) \cap I^i_{n+1}$, $i \in \{0,1\}$, has measure greater than $1/k2^{2(n+1)}$. The algorithm picks as I_{n+1} the set I^i_{n+1} which fulfills this condition, with the oracle used to decide in case both sets verify it. Hence, at every stage n, $c(k,n) \cap I_n$ is non-empty, so there are absolutely normal numbers in it; furthermore, by construction all reals in $c(k,n) \cap I_n$ have a fractional expansion starting with $\alpha(0)\ \alpha(1) \ldots \alpha(n)$.

We now prove that, for a fixed k, these real numbers α form a set of Lebesgue measure at least $1 - 1/k$. Consider the inductively defined set $M(k,n+1)$ consisting of all possible intervals $\left(\frac{m}{2^{n+1}}, \frac{m+1}{2^{n+1}}\right)$, with $m = 0, 1, \ldots 2^{n+1} - 1$, as we allow the first $n+1$ digits of θ to run through all possibilities. I.e., having deleted those intervals that would be discarded by the algorithm up to stage n. Notice that the algorithm discards the interval $\left(\frac{m}{2^{n+1}}, \frac{m+1}{2^{n+1}}\right)$ when $\mu\left(c(k,n) \cap \left(\frac{m}{2^{n+1}}, \frac{m+1}{2^{n+1}}\right)\right) \leq 1/k2^{2n}$.

Let $M : \mathbb{N} \times \mathbb{N} \to \mathcal{P}((0,1))$. $M(k,0) = (0,1)$, and for $n \geq 0$,

$$M(k,n+1) = \bigcup_{\substack{I_m \subseteq M(k,n) \\ \mu\left(c(k,n) \cap I_m\right) > 1/k2^{2n}}} I_m$$

where $I_m = \left(\frac{m}{2^{n+1}}, \frac{m+1}{2^{n+1}}\right)$, for $m = 0, 1, \ldots, 2^{n+1} - 1$.

Then, $\mu\left(E(k) \cap M(k,n+1)\right) = \mu\left(E(k) \cap M(k,n)\right) -$

$$-\sum_{m=0}^{2^n-1} \mu\left(E(k) \cap (M(k,n) \setminus M(k,n+1)) \cap \left(\tfrac{m}{2^n}, \tfrac{m+1}{2^n}\right)\right).$$

Since it is impossible that both halves of $\left(\frac{m}{2^n}, \frac{m+1}{2^n}\right)$ are included in $M(k,n+1)$, we have $\mu\left(E(k) \cap (M(k,n) \setminus M(k,n+1)) \cap \left(\frac{m}{2^n}, \frac{m+1}{2^n}\right)\right) \leq 1/k2^{2n}$, so that

$$\begin{aligned}
\mu\left(E(k) \cap M(k,n+1)\right) &\geq \mu\left(E(k) \cap M(k,n)\right) - 1/k2^n \\
&\geq \mu\left(E(k) \cap M(k,n-1)\right) - 1/k2^{n-1} - 1/k2^n \\
&\vdots \\
&\geq \mu\left(E(k) \cap M(k,1)\right) - 1/k\sum_{i=1}^{n} 1/2^n \\
&> \mu\left(E(k)\right) - 1/k = 1 - 2/k.
\end{aligned}$$

where the last inequality follows because $c(k,0) = (0,1)$ and $k > 2$, so $M(k,1) = (0, \frac{1}{2}) \cup (\frac{1}{2}, 1)$; henceforth, $E(k) \cap M(k,1) = E(k)$. We conclude $\mu\left(E(k) \cap \bigcap_n M(k,n)\right) \geq 1 - 2/k$. This completes the proof. □

Remark 23 (Convergence to normality)**.** The algorithm outputs the real $\alpha \in \bigcap_{n \geq 0} c(k,n)$. By Definitions 22 and 17, $c(k,n) \subseteq A_{k2^{2n+1}}$ and

$$A_{k2^{2n+1}} = \bigcap_{2 \leq t \leq T} \bigcap_{1 \leq r \leq L} \bigcap_{\gamma \in \{0,\ldots t-1\}^r} \{\alpha \in (0,1) : |S(\alpha, t, \gamma, R) - R/t^r| < \varepsilon R\}$$

with , $R = k2^{2n+1}$, $L = \sqrt{\ln R}/4$, $T = e^L$, $\varepsilon = T^{-L}$. This gives an explicit convergence to absolute normality of α: for each initial segment of α of length $R = k2^{2n+1}$ expressed in each scale up to $T = e^L$ all words of length up to $L = \sqrt{\ln R}/4$ occur with the expected frequency plus or minus e^{-L^2}.

Remark 24 (Complexity of the algorithm)**.** The algorithmic complexity of computing the n-th digit of α comes exclusively from the computation of $\mu\left(c(k,n) \cap I_n^i\right)$, $i = 0,1$. The naive way to obtain this is by constructing $c(k,n) = A_{k2^{2n+1}} \cap c(k,n-1) \cap (\beta_n, 1)$ and leads to a double exponential time algorithm. In Turing's manuscript there are no properties that would allow for a faster computation, like exploring the relation between the sets $A_{k2^{2n+1}}$ and $A_{k2^{2n+2}}$.

Remark 25 (Absolutely normal reals in every Turing degree)**.** It follows from the algorithm that, for a fixed $k \in \mathbb{N}$, by taking particular sequences $\theta \in \{0,1\}^\infty$ one obtains particular absolutely normal numbers, computable in θ. In [10] we use a variation of Turing's algorithm that queries the oracle infinitely many times in a controlled way: the algorithm intercalates the oracle digits in fixed positions of the absolutely normal number being constructed. One obtains absolutely normal numbers in each Turing degree (in fact, in each 1-degree). This result can be based either in the reconstruction of Turing's idea presented here, or in our algorithm [4] inspired by Sierpiński's work [14].

Acknowledgements: We thank the referees for their comments, including the results on polynomial time martingales, and other useful remarks. We acknowledge a suggestion of Glyn Harman (personal communication) that helped with the proof of Lemma 8. We are indebted to Max Dickmann for vivid discussions since we started with this work, and for all his numerous and always careful observations.

References

[1] Klaus Ambos-Spies and Elvira Mayordomo. Resource-bounded measure and randomness. In *Complexity, Logic, and Recursion Theory*, pages 1–47. A. Sorbi, Ed., Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, 1997.

[2] Klaus Ambos-Spies, Sebastiaan Terwijn, and Xizhong Zheng. Resource bounded randomness and weakly complete problems. *Theoretical Computer Science*, 172:195–207, 1997.

[3] David H. Bailey and Richard E. Crandall. Random generators and normal numbers. *Experimental Mathematics*, 11(4):527–546, 2004.

[4] Verónica Becher and Santiago Figueira. An example of a computable absolutely normal number. *Theoretical Computer Science*, 270:947–958, 2002.

[5] Émile Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rendiconti del Circolo Matematico di Palermo*, 27:247–271, 1909.

[6] Émile Borel. *Leçons sur la thèorie des fonctions.* Gauthier Villars, 2nd edition, 1914.

[7] Émile Borel. La définition en mathématiques. In François Le Lionnais, editor, *Les grands courants de la pensèe mathématique.* Hermann, 1998.

[8] David G. Champernowne. The construction of decimals in the scale of ten. *Journal of the London Mathematical Society*, 8:254–260, 1933.

[9] Arthur H. Copeland and Paul Erdös. Note on normal numbers. *Bulletin American Mathematical Society*, 52:857–860, 1946.

[10] Santiago Figueira. *Aspects of Randomness.* PhD thesis, Universidad de Buenos Aires, May 2006.

[11] Glyn Harman. *Metric Number Theory*, volume 18 of *London Mathematical Society Monographs.* Oxford University Press, 1998.

[12] Glyn Harman. One hundred years of normal numbers. In M. A. Bennett, B.C. Brendt, N. Boston, H.G. Diamond, A.J. Hildebrand, and W. Philipp, editors, *Millennial Conference on Number Theory*, volume 2 of *Number Theory for the Millennium*, pages 149–166. A. K. Peters, 2002.

[13] Lauwerens Kuipers and Harald Niederreiter. *Uniform distribution of sequences.* Wiley Interscience, New York, 1974.

[14] Wacław Sierpiński. Démonstration élémentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'un tel nombre. *Bulletin de la Société Mathématique de France*, 45:127–132, 1917.

[15] Alan M. Turing. A note on normal numbers. In J.L. Britton, editor, *Collected Works of A.M. Turing: Pure Mathematics*, pages 117–119. North Holland, Amsterdam, 1992.