DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN
The Netherlands

Entropy and counting correct digits

Wieb Bosma, Karma Dajani and Cor Kraaikamp

Report No. 9925 (June 1999)

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NIJMEGEN
Toernooiveld
6525 ED Nijmegen
The Netherlands

Abstract.
Expansions that furnish increasingly good approximations to real numbers are usually related to dynamical systems. Although comparing dynamical systems seems difficult in general, Lochs was able in 1964 to relate the relative speed of approximation of decimal and regular continued fraction expansions (almost everywhere) to the quotient of the entropies of their dynamical systems. He used detailed knowledge of the continued fraction operator. In this paper we show that his result follows from certain general properties that the continued fraction map satisfies. Other expansions, such as the alternating Lüroth series, have similar properties, and this allows us to generalize Lochs' result. Even in cases where the expansion operator satisfies weaker conditions, the same correlation between relative speed of approximation and entropy quotient seems to exist. We were able to prove this in the case of g-adic expansions (by elementary but technical means), and furnish numerical evidence in other cases. If this hypothesis is correct, it provides a means to approximate hitherto unknown values of entropy. We apply this to an example of Bolyai expansions.

1. Introduction

The purpose of this paper is to compare the number of 'digits' determined in one 'expansion' of a real number as a function of the number of 'digits' given in some other 'expansion'. Since one of the aims is to determine conditions on the 'types of expansion' that ensure that such comparison is possible, we have used inverted commas. For the moment the *regular continued fraction* expansion and the *decimal* expansion will serve as examples, the digits then being *partial quotients* and *decimal digits.*

It may seem difficult at first sight to compare the decimal expansion with the continued fraction expansion, since the dynamics of these expansions are very different. However, in the early sixties G. Lochs [5] obtained a surprising and beautiful result that will serve as a prototype for the results we are after. In it, the number $m(n)$ of partial quotients (or continued fraction digits) determined by the first n decimal digits is compared with n. Thus knowledge of the first n decimal digits of the real number x determines completely the value of exactly $m(n)$ partial quotients; to know more partial quotients unambiguously requires more decimal digits. Of course $m(n) = m(x, n)$ will depend on x. But asymptotically the following holds; 'almost all' statements in this paper are with respect to the Lebesgue measure λ.

1.1 Theorem (Lochs). *For almost all x:*

$$\lim_{n\to\infty} \frac{m(n)}{n} = \frac{6\log 2 \log 10}{\pi^2} = 0.97027014\cdots.$$

Roughly speaking, this theorem tells us that usually around 97 partial quotients are determined by 100 decimal digits. By way of example Lochs calculated (on an early main-frame computer [6]) that the first 1000 decimals of π determine 968 partial quotients.

In this paper we will generalize Lochs' result, by showing that it is related to the ratios of the entropies of the maps under consideration and is really a consequence of the Theorem of Shannon-McMillan-Breiman, (see [2, p. 129]), that can be formulated as follows (using cylinder sets defined in the next section).

1.2 Theorem (Shannon-McMillan-Breiman). *If S is an ergodic, measure preserving transformation with invariant measure μ_S and with entropy $h(S)$, then for almost all x:*

$$\lim_{n\to\infty} \frac{-\log \mu_S(B_n(x))}{n} = h(S), \tag{1.3}$$

where $B_n(x)$ is the cylinder of rank n containing x.

With this, the dynamics of two completely different systems can be linked. The essential idea in this generalization is then used to obtain a comparison of g-adic and h-adic expansions. We emphasize here that the result, although superficially the same, is much stronger than the naïve statement that usually n digits in base g approximate x to the same precision as $\log g / \log h$ digits in base h: the latter is merely a statement about the *size* of the interval determined by the first n digits in base g, whereas our result links the relative position of it to that of the interval determined by $m(n)$ digits in base h.

2. Fibred systems

In order to generalize Lochs' result we start with some notations and definitions that place maps like the continued fraction transformation in a more general framework; see also [9], from which the following is taken with a slight modification.

2.1 Definitions. Let B be an interval in $\mathbf{R}$, and let $T : B \to B$ be a surjective map. We call the pair (B, T) a *fibred system* if the following conditions are satisfied:

(a) There is a finite or countably infinite set D (called the *digit set*), as well as

(b) a map $k : B \to D$ such that the sets

$$B(i) = k^{-1}\{i\} = \{x \in B : k(x) = i\}$$

form a partition of B. We assume that $\mathcal{P} = \{\mathcal{B}(\rangle) : \rangle \in \mathcal{D}\}$ is a generating partition of intervals; moreover

(c) the restriction of T to any $B(i)$ is an injective continuous map.

The *cylinder of rank n belonging to the digits* $k_1, k_2, \cdots, k_n \in D$ is the set

$$B(k_1, k_2, \cdots, k_n) = B(k_1) \cap T^{-1}B(k_2) \cap \ldots \cap T^{-n+1}B(k_n);$$

by definition B is the cylinder of rank 0. Two cylinder sets B_n, B_n^* of rank n are called *adjacent* if they are contained in the same cylinder set B_{n-1} of rank $n-1$ and their closures have non-empty intersection:

$$\overline{B_n} \cap \overline{B_n^*} \neq \emptyset.$$

For any fibred system (B, T) the map T can be viewed as a *shift map* in the following way. Consider the following correspondence

$$\Psi : x \mapsto (k_1, k_2, \ldots, k_n, \ldots),$$

where $k_i = j$ if and only if $T^{i-1}x \in B(j)$, then

$$T : x \mapsto (k_2, k_3, \ldots, k_n, \ldots).$$

We write $B_n(x)$ for the cylinder of rank n which contains x, i.e.,

$$B_n(x) = B(k_1, k_2, \cdots, k_n), \quad \Longleftrightarrow \quad \Psi(x) = (k_1, k_2, \ldots, k_n, \ldots).$$

A sequence $(k_1, \ldots, k_n) \in D^n$ is called *admissible* if there exists $x \in B$ such that $\Psi(x) = (k_1, \ldots, k_n, \ldots)$.

It will be clear that Ψ associates to any $x \in B$ the 'expansion' $(k_1, k_2, \ldots)$, and that the expansion is determined by T and the digit map k. In order to arrive at the desired type of result for the fibred system (B, T), we need to impose some restrictions on T.

2.2 Definitions. A fibred system (B,T) is called a *number theoretic* fibred system and the map T a *number theoretic fibred map* if T satisfies the additional conditions:

(d) T is ergodic with respect to the Lebesgue measure λ, and

(e) T has an invariant probability measure μ_T which is equivalent to λ, i.e., there exist c_1 and c_2, with $0 < c_1 < c_2 < \infty$, such that

$$c_1 \lambda(E) \le \mu_T(E) \le c_2 \lambda(E),$$

for every Borel set $E \subset B$.

2.3 Examples. The following examples are used throughout the text.

(i) The first example is the *regular continued fraction*. Here $B = [0,1) \subset \mathbf{R}$, the map T is given by $T(x) = \frac{1}{x} - k(x)$, and $k(x) = \lfloor \frac{1}{x} \rfloor$. Hence $B(i) = (\frac{1}{i+1}, \frac{1}{i}]$, for all $i \ge 1$. It is a standard result that the map is ergodic with the Gauss-measure as invariant measure. The entropy of T equals $h(T) = \pi^2/(6\log 2) = 2.373\cdots$.

(ii) The second standard example is the *g-adic expansion*, for an integer $g \ge 2$. Here the map is $T_g(x) = g \cdot x - k(x)$, and $k(x) = \lfloor g \cdot x \rfloor$. Hence $T_g(x)$ is the representative of $g \cdot x \bmod 1$ in $B = [0,1)$, and $B(i) = [\frac{i}{g}, \frac{i+1}{g})$, for $0 \le i < g$. Again this is a number theoretic fibred map, with invariant measure λ. The entropy of T_g equals $h(T_g) = \log g$.

(iii) A third example is given by the Lüroth expansion. Now $B(i) = [\frac{1}{i+1}, \frac{1}{i})$ for $i \ge 1$. If T is taken as an increasing linear map onto $[0,1)$ on each $B(i)$, the *Lüroth series* expansion is obtained; if T is taken to be decreasing linear onto $[0,1)$ it generates the *alternating Lüroth series* expansion of a number in $[0,1)$. See [1] and [4], where it was shown that the invariant measure is the Lebesgue measure and that the entropy equals

$$h(T) = \sum_{k=1}^{\infty} \frac{\log(k(k+1))}{k(k+1)} = 2.046\cdots.$$

(iv) As a fourth example we take Bolyai's expansion; again $B = [0,1)$, and now $T : [0,1) \to [0,1)$ is defined by

$$T(x) = (x+1)^2 - 1 - \varepsilon_1(x),$$

where

$$\varepsilon_1(x) = \begin{cases} 0 & \text{if } x \in B(0) = [0, \sqrt{2}-1); \\ 1 & \text{if } x \in B(1) = [\sqrt{2}-1, \sqrt{3}-1); \\ 2 & \text{if } x \in B(2) = [\sqrt{3}-1, 1). \end{cases}$$

Setting $\varepsilon_n = \varepsilon_n(x) = \varepsilon_1(T^{n-1}(x))$, $n \geq 1$, one has

$$x = -1 + \sqrt{\varepsilon_1 + \sqrt{\varepsilon_2 + \sqrt{\varepsilon_3 + \cdots}}}.$$

It is shown in [8] that T is a number theoretic fibred map. Neither the invariant measure μ nor the entropy $h(T)$ is known.

We have the following lemma.

2.4 Lemma. *Let T a number theoretic fibred map on B, then for almost all x:*

$$\lim_{n\to\infty} \frac{\log \lambda(B_n(x))}{\log \mu_T(B_n(x))} = 1.$$

Proof. Since

$$c_1\lambda(B_n(x)) \leq \mu_T(B_n(x)) \leq c_2\lambda(B_n(x)),$$

it follows that

$$\frac{-\log c_2 + \log \mu_T(B_n(x))}{\log \mu_T(B_n(x))} \leq \frac{\log \lambda(B_n(x))}{\log \mu_T(B_n(x))} \leq \frac{-\log c_1 + \log \mu_T(B_n(x))}{\log \mu_T(B_n(x))}.$$

Taking limits the desired result follows, since $\mathcal{P}$ is a generating partition and therefore $\lim_{n\to\infty} \mu_T(B_n(x)) = 0$ almost surely.

The following formalizes a property that will turn out to be useful in comparing number theoretic maps.

2.5 Definition. Let T be a number theoretic fibred map, and let $I \subset B$ be an interval. Let $m = m(I) \geq 0$ be the largest integer for which we can find an admissible sequence of digits $a_1, a_2, \ldots, a_m$ such that $I \subset B(a_1, a_2, \ldots, a_m)$. We say that T is *r-regular* for $r \in \mathbf{N}$, if for some constant $L \geq 1$ the following hold:
(i) for every pair B, B^* of adjacent cylinders of rank n:

$$\frac{1}{L} \leq \frac{\lambda(\overline{B_n})}{\lambda(\overline{B_n^*})} \leq L,$$

(ii) for almost every $x \in I$ there exists a positive integer $j \leq r$ such that either $B_{m+j}(x) \subset I$ or $B^*_{m+j} \subset I$ for an adjacent cylinder B^*_{m+j} of $B_{m+j}(x)$.

Thus r-regularity expresses, loosely speaking, that if the expansion T agrees to the first m digits in both endpoints of a given interval I, then for a.e. $x \in I$ there exists a cylinder of rank $m+r$ with the property that it or an adjacent cylinder contains x, and is contained entirely in I; moreover, the size of two adjacent cylinders differs by not more than the constant factor L.

2.6 Examples.

(i) Let $I \subset [0,1)$ be a subinterval of positive length, and let $m = m(I)$ be such that B_m is the smallest RCF-cylinder containing I, i.e., there exists a vector $(a_1, \cdots, a_m) \in \mathbf{N^m}$ such that $B_m = B_m(a_1, \cdots, a_m) \supset I$ and $\lambda(I \cap ([0,1) \backslash B_{m+1}(a_1, \cdots, a_m, a))) > 0$ for all $a \in \mathbf{N}$. One can easily check that $r = 3$ in case of the RCF. Furthermore, it is well-known, see e.g. (4.10) in [2, p. 43], that

$$\lambda(B_m(a_1, \cdots, a_m)) = \frac{1}{Q_m(Q_m + Q_{m-1})},$$

which yields, together with the well-known recurrence relations for the sequence $(Q_m)_{m \geq 0}$ (see [2, (4.2), p. 41]):

$$Q_{-1} := 0; \quad Q_0 := 1; \quad Q_m = a_m Q_{m-1} + Q_{m-2}, \quad m \geq 1,$$

that

$$\lambda(B_m(a_1, \cdots, a_m)) \leq 3\lambda(B_m(a_1, \cdots, a_m + 1)),$$

for all $(a_1, \cdots, a_m) \in \mathbf{N^m}$, i.e., the continued fraction map T is 3-regular, with $L = 3$.

(ii) The g-adic expansion is not r-regular for any r. Although adjacent rank n cylinders are of the same size, so $L = 1$ can be taken in (2.5)(i), property (2.5)(ii) does not hold. This can be seen by taking x in a very small interval I around $1/g$.

(iii) The *alternating* Lüroth map is 3-regular with $L = 2$.

3. A comparison result.

We would like to compare two expansions. In this Section we will therefore assume that S and T are both number theoretic fibred maps on B. We denote the cylinders of rank n of S by A_n, and those of T by B_n. We let $m(x,n)$ for $x \in B$ and $n \geq 1$ be the number of T-digits determined by n digits with respect to S, so $m(x,n)$ is the largest positive integer m such that $A_n(x) \subset B_m(x)$. Yet another way of putting this, is that both endpoints of $A_n(x)$ agree to exactly the first m digits with respect to T.

We have the following general theorem.

3.1 Theorem Suppose that T is r-regular. *Then for almost all $x \in B$*

$$\lim_{n \to \infty} \frac{m(x,n)}{n} = \frac{h(S)}{h(T)}.$$

Proof. For $x \in B$ let the first n S-digits be given; then by definition of $m = m(x,n)$ one has that $A_n(x) \subset B_m(x)$. Since T is r-regular, there exists $1 \leq j \leq r$ such that

$$B_{m+j} \subset A_n(x) \subset B_m(x),$$

where B_{m+j} is either $B_{m+j}(x)$, or adjacent to it of the same rank. By regularity again, for some L,

$$\frac{1}{L}\lambda(B_{m+r}(x)) \leq \frac{1}{L}\lambda(B_{m+j}(x)) \leq \lambda(B_{m+j}),$$

so

$$\frac{1}{n}\left(-\log L + \log\lambda(B_{m+r}(x))\right) \leq \frac{1}{n}\log\lambda(A_n(x)) \leq \frac{1}{n}\log\lambda(B_m(x)),$$

and the result follows from Lemma (2.4) (applied to both T and S), and the Theorem of Shannon-McMillan-Breiman (1.2).

3.2 Corollary. *Let $m_g^{\text{RCF}}(x,n)$ be the number of partial quotients of x determined by the first n digits of x in its g-adic expansion. Then for almost all x:*

$$\lim_{n\to\infty} \frac{m_g^{\text{RCF}}(x,n)}{n} = \frac{6\log 2\log g}{\pi^2}.$$

This generalizes Lochs' theorem to arbitrary g-adic expansions.

3.3 Corollary. *Let $m_g^{\text{ALE}}(x,n)$ be the number of alternating Lüroth digits of x determined by the first n digits of x in its g-adic expansion. Then for almost all x:*

$$\lim_{n\to\infty} \frac{m_g^{\text{ALE}}(x,n)}{n} = \frac{\log g}{\sum_{k=1}^{\infty}\frac{\log k(k+1)}{k(k+1)}}.$$

Note that the conditions in Theorem (3.1) are not symmetric in S and T. Since the regularity condition is not satisfied by h-adic expansions, Theorem (3.1) is of no use in comparing g-adic and h-adic expansions, nor in proving that the first n regular partial quotients determine usually $\pi^2/(6\log 2\log g)$ digits in base g.

4. Comparing g-adic expansions.

In this Section we compare g-adic and h-adic expansions, for integers $g, h \geq 2$.

Given the first n digits in the g-adic expansion of an irrational number $x \in [0,1)$, there exists a unique positive integer $\ell = \ell(n)$, such that

$$h^{-(\ell(n)+1)} \leq g^{-n} \leq h^{-\ell(n)}. \tag{4.1}$$

Thus the measure $\lambda(A_n)$ of a g-cylinder of rank n is comparable to that of an h-cylinder of rank $\ell(n)$.

It follows that

$$\limsup_{n\to\infty} \frac{\ell(n)}{n} \leq \frac{\log g}{\log h} \leq \liminf_{n\to\infty} \frac{\ell(n)+1}{n},$$

i.e., for almost all x

$$\lim_{n\to\infty} \frac{\ell(n)}{n} = \frac{\log g}{\log h}, \tag{4.2}$$

which is the ratio of the entropies of the maps T_g and T_h introduced in (2.3)(ii). Of course, one expects the following to hold:

$$\lim_{n\to\infty} \frac{m_g^{(h)}(x,n)}{n} = \frac{\log g}{\log h} \qquad \text{almost everywhere}, \tag{4.3}$$

which is the analog of Lochs' result for the maps T_g and T_h; here $m_g^{(h)}(x,n)$ is defined as before: it is the largest positive integer m such that $A_n(x)$ is contained in the h-adic cylinder $B_m(x)$. One should realize however, that $\ell(n)$ has in general no obvious relation to $m(n)$: (4.2) is merely a statement about the relative speed with which g-adic and h-adic cylinders shrink. However, a 'Lochs-type result' can still be obtained in this situation, as we will now show.

Let $x \in [0,1)$ be a generic number for $S = T_g$ for which we are given the first n digits $t_1, t_2, \ldots, t_n$ of its g-adic expansion. These digits define the g-adic cylinder

$$A_n(x) = A_n(t_1, t_2, \ldots, t_n).$$

Let $m(n) = m_g^{(h)}(n)$ and $\ell(n)$ be as defined above; note that

$$m(n) \le \ell(n), \quad \text{for all} \;\; n \ge 1. \tag{4.4}$$

The sequence $(m(n))_{n=1}^{\infty}$ is non-decreasing, but may remain constant some time; this means that it 'hangs' for a while, so $m(n+t) = \cdots = m(n+1) = m(n)$, after which it 'jumps' to a larger value, so $m(n+t+1) > m(n+t)$. Let $(n_k)_{k\ge1}$ be the subsequence of n for which $m(n)$ 'jumps', that is, for which $m(n_k) > m(n_k - 1)$.

4.5 Lemma.

$$\lim_{k\to\infty} \frac{m_g^{(h)}(n_k)}{n_k} = \frac{\log g}{\log h} \quad \text{for almost all} \;\; x.$$

Proof. By definition, $B_{m(n_k)}(x) \supset A_{n_k}(x)$. The cylinder $B_{m(n_k)}(x)$ consists of h cylinders of rank $m(n_k)+1$, and $A_{n_k}(x)$ intersects (at least) two of these, since otherwise some $B_{m(n_k)+1} \supset A_{n_k}(x)$, contradicting maximality of $m(n_k)$, and thus it contains an endpoint e of some $B_{m(n_k)+1}$ lying in the interior of $B_{m(n_k)}(x)$.

On the other hand, by definition of n_k, we know that $m(n_k) > m(n_k - 1)$, so $B_{m(n_k)}(x) \not\supset A_{n_k-1}(x)$, and therefore $A_{n_k-1}(x)$ contains an endpoint f of $B_{m(n_k)}(x)$ as well as e. Now e and f are at least $\lambda(B_{m(n_k)+1}) = h^{-(m(n_k)+1)}$ apart. Therefore

$$h^{-(m(n_k)+1)} \le \lambda(A_{n_k-1}(x)) = g \cdot \lambda(A_{n_k}(x)) = g \cdot g^{-n_k} \le g \cdot h^{-\ell(n_k)},$$

by (4.1). Hence

$$h^{-(m(n_k)+1)} \le g \cdot h^{-\ell(n_k)},$$

which, in combination with (4.4) implies

$$\ell(n_k) - 1 - \frac{\log g}{\log h} \le m(n_k) \le \ell(n_k), \quad k \ge 1. \tag{4.6}$$

But from (4.6) the Lemma follows immediately.

Next we would like to show that (4.3) follows from (4.5); this is easy in case $h = 2$.

4.7 Corollary. *With notations as above, for any $g \in \mathbf{N}_{\ge 2}$:*

$$\lim_{n\to\infty} \frac{m_g^{(2)}(x,n)}{n} = \frac{\log g}{\log 2} \qquad \text{for almost all } \ x.$$

Proof. If $h = 2$ the mid-point ξ of $B_{m(n)}$ lies somewhere in A_n. Now $m(n+1) = m(n)$ if this mid-point ξ is located in $A_{n+1}(t_1, , t_2, \ldots, t_n, t_{n+1})$, that is, if it is located in the same g-adic cylinder of order $n+1$ as x. Notice that this happens with probability $\frac{1}{g}$, and that the randomness here is determined by the g-adic digit t_{n+1}. To be more precise, let H denote the event that we will 'hang' at time n, i.e., the event that $m(n) = m(n+1)$, and let D be a random variable with realizations $t \in \{0, 1, \cdots g-1\}$, defined by $\xi \in A_{n+1}(t_1, \ldots, t_n, D)$. Now

$$\mathrm{P}(H) = \sum_{i=0}^{g-1} \mathrm{P}(H|D = i) \cdot \mathrm{P}(D = i),$$

and from $\mathrm{P}(H|D = i) = 1/g$ for $0 \le i \le g-1$ it then follows that $\mathrm{P}(H) = 1/g$; due to the discrete uniform distribution of the digit t_{n+1} of x we do not have to know the probabilities $\mathrm{P}(D = i)$.

Thus we see that for each $k \ge 1$ (and with $n_0 = 0$) the random variable

$$v_k = n_k - n_{k-1}$$

is geometrically distributed with parameter $p = 1/g$. Furthermore, the v_k's are independent, and, since $n_k \ge k$

$$1 \le \frac{n_{k+1}}{n_k} = \frac{n_k + n_{k+1} - n_k}{n_k} = 1 + \frac{1}{n_k} v_{k+1} \le 1 + \frac{1}{k} v_{k+1}.$$

Because $v_1, v_2, \ldots$ are independently identically distributed with finite expectation it follows from the Lemma of Borel-Cantelli that $\lim_{k\to\infty} \frac{v_{k+1}}{k} = 0$ (a.e.), and therefore

$$\lim_{k\to\infty} \frac{n_{k+1}}{n_k} = 1 \quad \text{(a.e.)}.$$

Given any $n \geq 1$, there exists $k = k(n) \geq 0$ such that $n_k < n \leq n_{k+1}$. Since $m(n_k) \leq m(n) \leq m(n_{k+1})$ one has

$$\frac{m(n_{k(n)})}{n_{k(n)+1}} \leq \frac{m(n)}{n} \leq \frac{m(n_{k(n)+1})}{n_{k(n)}},$$

But then the result follows from (4.5) with $h = 2$.

If $h \geq 3$ the situation is more complicated; there might be more than one midpoint ξ_i 'hitting' A_n. In case only one ξ_i 'hits' A_n (as is always the case when $h = 2$), we speak of a type 1 situation. In case there is more than one 'hit' we speak of a type 2 situation. Now change the sequence of jump-times $(n_k)_{k\geq 1}$ by adding those n for which we are in a type 2 situation, and remove all n_k's for which we are in a type 1 situation. Denote this sequence by $(n_k^*)_{k\geq 1}$. If this sequence is finite, we were originally in a type 1 situation for n sufficiently large. In case $(n_k^*)_{k\geq 1}$ is an infinite sequence, notice that for every n for which there exists a k such that $n = n_k^*$ one has 1-regularity, i.e.,

$$B_{m(n_k^*)+1} \subset A_{n_k^*}(x) \subset B_{m(n_k^*)}(x)\,,$$

where $B_{m(n_k^*)+1}$ is either $B_{m(n_k^*)+1}(x)$ or an adjacent interval. Following the proof of Theorem (3.1) one has

$$\lim_{k\to\infty} \frac{m(n_k^*)}{n_k^*} = \frac{\log g}{\log h} \quad \text{(a.e.)}.$$

Let $(\hat{n}_k)_{k\geq 1}$ be the sequence we get by merging $(n_k)_{k\geq 1}$ and $(n_k^*)_{k\geq 1}$. Notice that

$$\lim_{k\to\infty} \frac{m(\hat{n}_k)}{\hat{n}_k} = \frac{\log g}{\log h} \quad \text{(a.e.)}. \tag{4.8}$$

If $(\tilde{n}_k)_{k\geq 1}$ is $\mathbf{N} \setminus (\hat{\mathbf{n}}_\mathbf{k})_{\mathbf{k}\geq \mathbf{1}}$, we are left to show that

$$\lim_{k\to\infty} \frac{m(\tilde{n}_k)}{\tilde{n}_k} = \frac{\log g}{\log h} \quad \text{(a.e.)}.$$

Let $n = \tilde{n}_j$ for some $j \geq 1$. Then there exist unique $k = k(j)$ and $h = h(j)$ such that $n_{k+1} = \hat{n}_{h+1}$, $n_k \leq \hat{n}_h$ and

$$\hat{n}_h < n < n_{k+1}.$$

Since $n_{k+1} - \hat{n}_h = \hat{n}_{h+1} - \hat{n}_h = v_{h+1}$ is geometrically distributed with parameter $1/g$, and since $v_1, v_2, \ldots$ are independent with finite expectation, it follows as in the case $h = 2$, that

$$\frac{\hat{n}_{h+1}}{\hat{n}_h} = 1 + \frac{1}{\hat{n}_h} v_{h+1} \leq 1 + \frac{1}{h} v_{h+1} \to 1 \quad \text{as} \quad j \to \infty \quad \text{(a.e.)}.$$

But then (4.3) follows from (4.8) and from

$$\frac{m(\hat{n}_h)}{\hat{n}_{h+1}} \leq \frac{m(n)}{n} \leq \frac{m(\hat{n}_{h+1})}{\hat{n}_h}.$$

We have proved the following theorem.

4.9 Theorem. *Let $g, h \in \mathbf{N}_{\geq 2}$. Then for almost all x*

$$\lim_{n\to\infty} \frac{m_g^{(h)}(x,n)}{n} = \frac{\log g}{\log h}.$$

5. Remarks on generalizations.

At first sight one might think that the proof of Theorem (4.9) can easily be extended to more general number theoretic fibered maps. However, closer examination of the proof of Theorem (4.9) shows that there are two points that make generalizations of the proof hard, if not impossible. The first is the observation, that the 'hanging time' v_k (in a type 1 situation) is geometrically distributed, and that the v_k's are independently identically distributed. Recall that this observation follows from the fact that the digits given by the g-adic map S_g are independently identically distributed and have a discrete uniform distribution. As soon as this last property no longer holds (e.g., if $S = T_\beta$, the expansion with respect to some non-integer $\beta > 1$), we need to know $\mathrm{P}(D = i)$ for each $i \in \{0, 1, \cdots \lfloor\beta\rfloor - 1\}$, and this might be difficult.

Even if we assume S to be the g-adic map, another problem arises when T is not the h-adic map, but, for example, $T = T_\beta$. In that case all the ingredients of the proof of Theorem (4.9) seem to work with the exception that at the 'jump-times' $(n_k)_k$ we might not be able to show that

$$\ell(n_k) - C \leq m(n_k) \leq \ell(n_k), \tag{5.1}$$

with $C \leq 1$ some fixed constant. Notice that from (5.1) it would follow that

$$\lim_{k\to\infty} \frac{m(n_k)}{n_k} = \frac{\log g}{\log \beta}.$$

For one class of $\beta \in (1, 2)$ one can show that (5.1) still holds. These β's are the so-called 'pseudo-golden mean' numbers; $\beta > 1$ is a 'pseudo-golden mean' number if β is the positive root of $X^k - X^{k-1} - \cdots - X - 1 = 0$, for some $k \in \mathbf{N}, \mathbf{k} \geq \mathbf{2}$ (in case $k = 2$ one has that $\beta = \frac{1}{2}(1 + \sqrt{5})$, which is the golden mean). These 'pseudo-golden mean' numbers β are all Pisot numbers, and satisfy

$$1 = \frac{1}{\beta} + \frac{1}{\beta^2} + \cdots + \frac{1}{\beta^k}.$$

5.2 Theorem. *With notations as before, let $S = T_g$ and $T = T_\beta$, with $g \in \mathbf{N}_{\geq 2}$, and $\beta > 1$ a 'pseudo golden number'. Then*

$$\lim_{n\to\infty} \frac{m_g^{(\beta)}(n)}{n} = \frac{\log g}{\log \beta}, \quad \text{(a.e.)}.$$

Proof. In case $k = 2$, a digit 1 is always followed by a zero, and only the β-cylinders $B_m^{(\beta)}$ corresponding to a sequence of β-digits ending with 0 are refined. Thus from the definition of $m(n)$ it follows that the last digit of $B^{(\beta)}_{m(n)}(x)$ is always 0 (if it were 1, the choice was wrong). Note that in this case

$$\lambda(B^{(\beta)}_{m(n)}(x)) = \beta^{-m(n)}.$$

Let $\ell(n)$ be defined as before, and notice that at a 'jump-time' $n = n_k$ one has

$$\beta^{-(m(n)+2)} \leq g \cdot g^{-n} \leq g \cdot \beta^{-\ell(n)}, \tag{5.3}$$

from which

$$\ell(n) - 2 - \frac{\log g}{\log \beta} \leq m(n).$$

Since $m(n) \leq \ell(n)$ we see that (5.1) follows in case β equals the golden mean.

In case $k \geq 3$ the situation is slightly more complicated; let us consider here $k = 3$. Now any sequence of β-digits ending with two consecutive 1's must be followed by a zero, and therefore – by the definition of $m(n)$ – the last digit of $B^{(\beta)}_{m(n)}(x)$ is either 0, or is 1 which is preceded by 0. One can easily convince oneself then that

$$\lambda(B^{(\beta)}_{m(n)}(x)) = \begin{cases} \beta^{-m(n)} & \text{if } d_{m(n)}(x) = 0, \\ \beta^{-(m(n)+1)} + \beta^{-(m(n)+2)} & \text{if } d_{m(n)}(x) = 1. \end{cases}$$

Here $d_{m(n)}(x)$ is the $m(n)^{\text{th}}$ β-digit of x.

Now let $n = n_k$ be a 'jump-time'. In case $d_{m(n)}(x) = 0$ we see that $B^{(\beta)}_{m(n)}(x)$ consists of two β-cylinders, one of length $\beta^{-(m(n)+1)}$ and one of length $\beta^{-(m(n)+2)} + \beta^{-(m(n)+3)}$. One of these two cylinders is contained in $A_{n-1}(x)$ (due to the fact that n is a 'jump-time'), and (5.3) is therefore satisfied. Of course one has $m(n) \leq \ell(n)$ in this case. In case $d_{m(n)}(x) = 1$ we see that $B^{(\beta)}_{m(n)}(x)$ consists of two β-cylinders, one of length $\beta^{-(m(n)+1)}$ and one of length $\beta^{-(m(n)+2)}$. Now

$$\lambda(B^{(\beta)}_{m(n)}(x)) = \beta^{-(m(n)+1)} + \beta^{-(m(n)+2)} \leq \beta^{-m(n)},$$

and by definition of $\ell(n)$ one has $m(n) \leq \ell(n)$. Again one of these two sub-cylinders of $B^{(\beta)}_{m(n)}(x)$ is contained in $A_{n-1}(x)$, and one has

$$\beta^{-(m(n)+2)} \leq g \cdot g^{-n} \leq g \cdot \beta^{-\ell(n)}.$$

We see that (5.3) is again satisfied. For $k \geq 4$ the proof is similar; one only needs to consider more cases.

This proof does not easily generalize. However we still expect the following to hold.

5.4 Conjecture. *Let $g \in \mathbf{N}_{\geq 2}$ and $\beta > 1$; then, with notations as before:*

$$\lim_{n\to\infty} \frac{m^{(\beta)}_g(n)}{n} = \frac{\log g}{\log \beta}, \quad \text{(a.e.)}.$$

Clearly the 'ultimate' conjecture here is the following.

5.5 Conjecture. *For any two number theoretic fibred maps S and T*

$$\lim_{n\to\infty} \frac{m_S^T(x,n)}{n} = \frac{h(S)}{h(T)}, \quad \text{(a.e.)},$$

with $m_S^T(x,n)$ defined as before.

6. Some experimental evidence

In this Section we report on numerical experiments. Our experiments were carried out using the computer algebra system Magma, see [3]. The general set-up of the experiments was as follows. We choose n random digits for the expansion of a number x in $[0, 1)$ with respect to a number theoretic fibred map S, and compute $m_S^T(x, n)$, the number of digits of x with respect to the number theoretic fibred map T determined completely by the first n digits of x with respect to S. This is done by comparing the T-expansions of both endpoints of the S-cylinder $A_n(x)$.

Experiment 1. The first experiment was designed to test the set-up. It aimed to verify Lochs' result: we let $S = T_{10}$, the decimal expansion map, and $T = T_{\mathrm{RCF}}$ for the regular continued fraction map. For $N = 1000$ random real numbers of $n = 1000$ decimal digits, we computed the number of partial quotients determined by the endpoints of the decimal cylinder $A_n(x)$.

We ran this experiment twice; the averages (over 1000 runs) were 970.534 and 969.178 (with standard deviations around 24). Compare these averages to the value $970.270\cdots$ predicted by Theorem (1.1).

Experiment 2. In this experiment we reversed the roles of decimal and continued fraction expansions. For $N = 1000$ random real numbers (between 0 and 1) we determined 1000 partial quotients for the regular continued fraction expansions, and we computed the number of decimal digits determined unambiguously. Note that this case is not covered by any of the Theorems (see the remark after Corollary 3.3); however, based on Conjecture (5.5) one expects a proportion equal to the inverse of Lochs' constant from (1.1) of accurate decimal digits, which comes down to 1030.641 decimal digits for 1000 partial quotients given. Carrying out the experiment twice gave 1030.495 and 1028.569 digits on average (with standard deviations around 25).

The random continued fraction expansions were generated from random real numbers (between 0 and 1), that were in turn chosen by picking enough decimal digits at random. To verify that indeed the partial quotients were random, we kept note of their distribution. The table below compares the observed frequency of digits (columns 3 and 4) with the theoretical distribution (in column 2) which says that digits k occurs with probability $\log(1 + \frac{1}{k(k+2)})/\log 2$.

digit	expected	observed (1)	observed (2)
1	0.4150	0.4155	0.4144
2	0.1699	0.1703	0.1698
3	0.0931	0.0930	0.0931
4	0.0589	0.0588	0.0588
5	0.0406	0.0404	0.0410
6	0.0297	0.0296	0.0299
7	0.0227	0.0227	0.0229
8	0.0179	0.0180	0.0180
9	0.0145	0.0144	0.0147
10	0.0120	0.0118	0.0121

Experiment 3. In this experiment we compared the relative speed of approximation of g-adic and h-adic expansions.

$h \rightarrow$	2			7			10		
$g \downarrow$	pred	observed		pred	observed		pred	observed	
2	–	–	–	**0.356**	0.355	0.355	**0.301**	0.300	0.300
7	**2.807**	2.805	2.805	–	–	–	**0.845**	0.844	0.844
10	**3.322**	3.320	3.320	**1.183**	1.182	1.182	–	–	–

In fact we compared binary, 7-adic and decimal expansions of $N = 1000$ random real numbers with each of the other two expansions, again by computing to how many h-digits both endpoints of the g-adic cylinder $A_{1000}(x)$ agreed. The two tables list the values found in each of two rounds of this experiment and compares the result with the value found by Theorem (4.9). Note that by (4.4) always $m(n) \leq \ell(n)$ and thus the observed ratio will always approximate $\log g/\log h$ from below. By reversing the roles of g and h, it is thus possible in this case to approximate the entropy quotient by sandwiching!

Experiment 4. Comparison of decimal and r-adic expansions, for *irrational* $r > 1$. Again, for $N = 1000$ random reals with $n = 1000$ decimal digits given we determined the number of r-adic digits determined unambiguously. We ran this experiment for 2 separate values of r, namely $r = \phi = (\sqrt{5}+1)/2$ and $r = \psi = (\sqrt[3]{2}+2)/2$. Although the numerical values of these irrationals are close together (around 1.630 and 1.618), the first is the golden number, to which Theorem (5.2) applies, whereas for the second value of r we only have a Conjectural value for the entropy quotient. Theorem (5.2) with $h = \phi$ predicts that on average 4785 r-digits would be determined; we found an experimental average of 4782. For ψ the conjectural value $\log 10/\log\psi \approx 4.713$, while we found on average 4710 digits per 1000 decimals. Also the standard deviations were close: around 2.1 and 2.4.

Experiment 5. Comparison of decimal and (alternating) Lüroth series. In this experiment we generated 1000 real numbers with 1000 random decimal digits, and computed the number of Lüroth digits determined by them. In the 2 runs we found on average around 1123.7 and 1125.0 digits were determined (with standard deviations of around 32). The same experiment was repeated with the alternating Lüroth expansion, and in that case we found that 1124.6 and 1125.3 digits were determined (standard deviation around 32). Note that in the latter case Corollary (3.3) tells us to expect a ratio $\log 10/2.04627 \approx 1.12526$.

Experiment 6. By choosing random real numbers of 1000 decimal digits, we attempted to estimate the unknown entropy of the Bolyai map (cf. (2.3)(iv)). We found that on average 1000 digits determined 2178.3 Bolyai digits in the first run of $N = 250$ random reals, and 2178.0 in the second; the standard deviations were around 13. These values would correspond to an entropy of around 1.0570 or 1.0572.

Experiment 7. Choosing N random Bolyai expansions of length 1000, we determined the number of decimal digits in one experiment, and the number of regular continued fraction partial quotient in another experiment. In the first run of $N = 1000$, we found that on average 458 decimal digits were determined, in the second run 457.8 on average (with standard deviations around 4). In runs with $N = 1000$ of the second experiment we found that 444.36 and 444.37 partial quotients were determined on average, with standard deviations of around 17.

These values would correspond to an entropy of around 1.0546, 1.0541 according to Conjecture 5.5, and 1.0545 or 1.0546 according to Theorem (3.1).

In this experiment we also kept a record of the distribution of the random Bolyai digits (over the possible values $0, 1, 2$). The random Bolyai expansions were generated by choosing random real numbers and finding their Bolyai expansion to 1000 digits. We found a fraction of 0.463 of digits equal to 0, a fraction 0.305 of digits equal to 1 and a fraction of 0.232 of digits equal to 2. (Compare this with $\sqrt{2} - 1 \approx 0.414$, $\sqrt{3} - \sqrt{2} \approx 0.318$ and $2 - \sqrt{3} \approx 0.268$; see (2.3)(iv) and (2.4).)

6.1 Conjecture. *The entropy of the Bolyai map of* (2.3)(iv) *is approximately* 1.0545.

[1] Barrionuevo, J., R.M. Burton, K. Dajani and C. Kraaikamp *Ergodic properties of generalized Lúroth series*, Acta Arithm. **74** (1996) , 311–327
[2] Billingsley, P. Ergodic Theory and Information, New York-London-Sydney: John Wiley and Sons, Inc. (1965)
[3] W. Bosma, J. Cannon, C. Playoust *The Magma Algebra System I: The User Language*, J. Symbolic Computation **24** (1997) , 235–265
[4] Kalpazidou, S., A. Knopfmacher and J. Knopfmacher *Metric Properties of Alternating Lüroth Series*, Portugaliae Math. **48** (1991) , 319–325
[5] Lochs, G. *Vergleich der Genauigkeit von Dezimalbruch und Kettenbruch*, Abh. Math. Sem. Hamburg **27** (1964) , 142–144
[6] Lochs, G. *Die ersten 968 Kettenbruchnenner von π*, Monatsh. Math. **67** (1963), 311–316
[7] Nakada, H. *Metrical theory for a class of continued fraction transformations and their natural extensions*, Tokyo J. Math. **4** (1981), 399–426
[8] Rényi, A. *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hung. **8** (1957), 472–493
[9] Schweiger, F. Ergodic Theory of Fibred Systems and Metric Number Theory, Oxford University Press, Oxford (1995)