STATISTICAL PROPERTIES OF THE CALKIN–WILF TREE: REAL AN $p$–ADIC DISTRIBUTION

GIEDRIUS ALKAUSKAS, JÖRN STEUDING

Abstract. We examine statistical properties of the Calkin–Wilf tree and give number-theoretical applications.

1. A mean-value related to the Calkin–Wilf tree

The Calkin–Wilf tree is generated by the iteration

$$\frac{a}{b} \mapsto \frac{a}{a+b}, \frac{a+b}{b},$$

starting from the root $\frac{1}{1}$; the number $\frac{a}{a+b}$ is called the left child of $\frac{a}{b}$ and $\frac{a+b}{b}$ the right child; we also say that $\frac{a}{b}$ is the mother of its children. Recently, Calkin & Wilf [1] have shown that this tree contains any positive rational number once and only once, each of which represented as a reduced fraction. The first iterations lead to

Reading the tree line by line, the Calkin–Wilf enumeration of $\mathbb{Q}^+$ starts with

$$\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{1}{3}, \frac{3}{2}, \frac{2}{3}, \frac{3}{3}, \frac{3}{4}, \frac{4}{3}, \frac{4}{5}, \frac{5}{4}, \frac{5}{3}, \frac{3}{5}, \frac{5}{2}, \frac{2}{5}, \frac{5}{3}, \frac{3}{4}, \frac{4}{1}, \cdots$$

As recently pointed out by Reznick [10], this sequence was already investigated by Stern [12] in 1858. This sequence satisfies also the iteration

$$x_1 = 1, \quad x_{n+1} = 1/(2[x_n] + 1 - x_n),$$
where $[x]$ denotes the largest integer $\leq x$; this observation is due to Newman (cf. [8]), answering a question of D.E. Knuth, resp. Vandervelde & Zagier (cf. [11]).

The Calkin–Wilf enumeration of the positive rationals has many interesting features. For instance, it encodes the hyperbinary representations of all positive integers (see [1]). Furthermore, it can be used as model for the game Euclid first formulated by Cole & Davie [2]; see Hofmann, Schuster & Steuding [5]. In this short note we are concerned with statistical properties of the Calkin–Wilf tree.

We write the $n$th generation of the Calkin–Wilf tree as $\mathcal{CW}^{(n)} = \{x_j^{(n)}\}_j$, where the $x_j^{(n)}$ are the elements ordered according to their appearance in the $n$th line of the Calkin–Wilf tree. So $\mathbb{Q}^+ = \bigcup_{n=1}^{\infty} \mathcal{CW}^{(n)}$. Obviously, $\mathcal{CW}^{(n)}$ consists of $2^{n-1}$ elements. Denote by $\Sigma(n)$ the sum of all elements of the $n$th generation of the Calkin–Wilf tree,

$$\Sigma(n) = \sum_{j=1}^{2^{n-1}} x_j^{(n)}.$$

Our first result gives the mean-value of the elements of the $n$th generation of the Calkin–Wilf tree:

**Theorem 1.** For any $n \in \mathbb{N}$,

$$\Sigma(n) = 3 \cdot 2^{n-2} - \frac{1}{2}.$$

This result may be interpreted as follows. We observe that $x_1^{(n)} = \frac{1}{n}$ and $x_{2n-1}^{(n)} = \frac{n}{1}$ for all $n \in \mathbb{N}$, and thus $\mathcal{CW}^{(n)}$ is supported on an unbounded set as $n \to \infty$. However, the average value of the $2^{n-1}$ elements of the $n$th generation $\mathcal{CW}^{(n)}$ is approximately $\frac{3}{2}$, which is, surprisingly, a finite number. This has a simple explanation: in some sense, small values are taken in earlier generations than large values. For instance, in each generation $\mathcal{CW}^{(n)}$ takes as many values form the interval $(0, 1)$ as from $(1, \infty)$. This result was also recently proved by Reznick [10]; his proof differs slightly from our argument.\(^1\)

\(^1\)The problem of determining the average value of the Calkin–Wilf tree was posed by the second named author as a problem in the problem session of the IV International conference on analytic and probabilistic number theory in Palanga 2006; an independent solution was given by Eduard Wirsing.
Proof by induction on \( n \). The statement of the theorem is correct for \( n = 1 \) and \( n = 2 \). Now suppose that \( n \geq 3 \). In order to prove the statement for \( n \) we first observe a symmetry in the Calkin–Wilf tree with respect to its middle: for \( n \geq 2 \),

\[
x_j^{(n)} = \frac{a}{b} \quad \iff \quad x_{2^{n-1}+1-j}^{(n)} = \frac{b}{a};
\]

this is easily proved by another induction on \( n \) (and we leave its simple verification to the reader). Further, we note that \( x_j^{(n)} \leq 1 \) if and only if \( j \) is odd; here equality holds if and only if \( n = 1 \).

Now we start to evaluate \( \Sigma(n) \). For this purpose we compute

\[
y_j^{(n)} := \begin{cases} x_j^{(n)} + x_{2^{n-1}-j}^{(n)} & \text{for } j = 1, 2, \ldots, 2^{n-2} - 1, \\
x_j^{(n)} + x_{2j}^{(n)} & \text{for } j = 2^{n-2},
\end{cases}
\]

and add these values over \( j = 1, 2, \ldots, 2^{n-2} \). Clearly, \( \Sigma(n) = \sum_{j=1}^{2^{n-2}} y_j^{(n)} \).

First, assume that \( j \) is odd. Then both, \( x_j^{(n)} \) and \( x_{2^{n-1}-j}^{(n)} \) are strictly less than 1. In view of (1) the mothers of \( x_j^{(n)} \) and \( x_{2^{n-1}-j}^{(n)} \) are of the form \( \frac{a}{b} \) and \( \frac{b}{a} \), respectively. Hence,

\[
x_j^{(n)} + x_{2^{n-1}-j}^{(n)} = \frac{a}{a+b} + \frac{b}{a+b}
\]

and thus we find \( y_j^{(n)} = 1 \) in this case.

Next, we consider the case that \( j \) is even. Then both, \( x_j^{(n)} \) and \( x_{2^{n-1}-j}^{(n)} \) are strictly greater than 1. If the mothers of \( x_j^{(n)} \) and \( x_{2^{n-1}-j}^{(n)} \) are of the form \( \frac{a}{b} \) and \( \frac{a'}{b'} \), respectively, then

\[
x_j^{(n)} = \frac{a+b}{b} = 1 + \frac{a}{b} \quad \text{and} \quad x_{2^{n-1}-j}^{(n)} = 1 + \frac{a'}{b'}.
\]

Hence, we find for their sum

\[
x_j^{(n)} + x_{2^{n-1}-j}^{(n)} = 2 + \frac{a}{b} + \frac{a'}{b'}
\]

and so \( y_j^{(n)} = 2 + y_k^{(n-1)} \), where \( y_k^{(n-1)} \) is either the sum of two elements \( x_k^{(n-1)} \) and \( x_{2^{n-2}-k}^{(n-1)} \) or the sum of \( x_{2^{n-3}}^{(n-1)} \) and \( x_{2^{n-2}}^{(n-1)} \).

It remains to combine both evaluations. Since both cases appear equally often, namely each \( 2^{n-3} \) times, we obtain the recurrence formula

\[
\Sigma(n) = \Sigma(n-1) + (1 + 2) \cdot 2^{n-3},
\]
being valid for $n \geq 3$. This implies the assertion of the theorem. ■

2. AN APPLICATION TO FINITE CONTINUED FRACTIONS

Theorem 1 has a nice number-theoretical interpretation. It is well-known that each positive rational number $x$ has a representation as a finite (regular) continued fraction

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \cdots + \frac{1}{a_m}}}$$

with $a_0 \in \mathbb{N} \cup \{0\}$ and $a_j \in \mathbb{N}$ for some $m \in \mathbb{N} \cup \{0\}$. In order to have a unique representation, we assume that $a_m \geq 2$ if $m \in \mathbb{N}$. We shall use the standard notation $x = [a_0, a_1, \ldots, a_m]$. Continued fractions are of special interest in the theory of diophantine approximation.

As Bird, Gibbons & Lester [3] showed, the $n$th generation of the Calkin–Wilf tree consists exactly of those rationals having a continued fraction expansion $[a_0, a_1, \ldots, a_m]$ for which the sum of the partial quotients $a_j$ is constant $n$, the continued fractions of even length in the left subtree, and the continued fractions with odd length in the right subtree. Thus Theorem 1 yields

**Corollary 2.** For any $n \in \mathbb{N}$,

$$2^{1-n} \sum_{a_0+a_1+\ldots+a_m=n} [a_0, a_1, \ldots, a_m] = \frac{3}{2} - 2^{-n}.$$

One can use the approach via continued fractions to locate any positive rational in the tree. This observation is due to Bird, Gibbons & Lester [3] (actually, their reasoning is based on Graham, Knuth & Patasnik [4] who gave such a description for the related Stern–Brocot tree). Given a reduced fraction $x$ in the Calkin–Wilf tree with continued fraction expansion

$$x = [a_0, a_1, \ldots, a_{m-2}, a_{m-1}, a_m],$$

we associate the path

- $L^{a_{m-1}} R^{a_{m-2}} \cdots L^{a_1} R^{a_0}$ if $m$ is odd, and
- $R^{a_{m-1}} L^{a_{m-2}} \cdots R^{a_1} L^{a_0}$ if $m$ is even;
Note that $a_m - 1 \geq 1$ for $m \in \mathbb{N}$. The notation $R^a$ with $a \in \mathbb{N} \cup \{0\}$ means: $a$ steps to the right, whereas $L^b$ with $b \in \mathbb{N} \cup \{0\}$ stands for $b$ steps to the left. Then, starting from the root $\frac{1}{1}$ and following this path from left to right, we end up with the element $x$. This follows easily from the iteration with which the tree was build; notice that this claim is essentially already contained in Lehmer [7] (this was also observed by Reznick [10]).

**Corollary 3.** Given any non-empty interval $(\alpha, \beta)$ in $\mathbb{R}^+$, and any finite path in the Calkin–Wilf tree, there exists a continuation of this path which contains a rational number from the interval $(\alpha, \beta)$.

**Proof.** We expand $\alpha$ and $\beta$ into continued fractions, $\alpha = [a_0, a_1, \ldots]$ and $\beta = [b_0, b_1, \ldots]$, say. Let $k$ be the least index such that $a_k \neq b_k$. According to the parity of $k$ we have $a_k < b_k$ (if $k$ is even) or $a_k > b_k$ (if $k$ is odd). Without loss of generality we may assume that $|b_k - a_k| \geq 2$ (since otherwise we may consider a subinterval of $(\alpha, \beta)$). Moreover we may suppose that the path in question is starting from the root and is given in the form $L^{c_{m-1}}R^{c_{m-1}}L^{c_{m-2}} \cdots L^{c_1}R^{c_0}$ (the other case may be treated analogously). Then we construct a rational number $x$ by assigning the finite continued fraction

$$x = [a_0, a_1, \ldots, a_{k-1}, x_k, x_{k+1}, c_0, c_1, \ldots, c_{m-2}, c_{m-1}, c_m],$$

where $x_k := \min\{a_k, b_k\} + 1$ and $x_{k+1}$ denotes the string 1 if $k$ is odd, resp. 1, 1 if $k$ is even. Since $b_j = a_j$ for $0 \leq j < k$ and

$$\min\{a_k, b_k\} < x_k < \max\{a_k, b_k\},$$

it follows that $\alpha < x < \beta$. Since the length of the continued fraction expansion has the same parity as $m$ (thanks to the definition of $x_{k+1}$), the element $x$ can be reached by the path $L^{c_{m-1}}R^{c_{m-1}}L^{c_{m-2}} \cdots L^{c_1}R^{c_0}$. This proves the corollary. ■

### 3. A random walk on the Calkin–Wilf tree

Starting with $X_1 = \frac{1}{1}$, we define a sequence of random variables by the following iteration: if $X_n = \frac{a}{b}$, then $X_{n+1} = \frac{a}{a+b}$ with probability $\frac{1}{2}$ and $X_{n+1} = \frac{a+b}{b}$ with probability $\frac{1}{2}$. The sequence $\{X_n\}$ may be
regarded as a random walk on the Calkin–Wilf tree where \( n \) is a discrete time parameter.

**Theorem 4.** Let \((\alpha, \beta)\) be any non-empty interval in \(\mathbb{R}^+\). Then, with probability 1, the random walk \(\{X_n\}\) visits the interval \((\alpha, \beta)\), i.e., with probability 1, there exists \(m \in \mathbb{N}\) such that \(x_m \in (\alpha, \beta)\).

**Proof.** The interval \((\alpha, \beta)\) contains a non-empty subinterval \([A, B]\) such that for any \(\zeta \in [A, B]\) the initial partial quotients \(c_0, c_1, \ldots, c_m\) are identical: \(\zeta = [c_0, c_1, \ldots, c_m, \ldots]\). Hence, with the interval \([A, B]\) we may associate a path pattern \(L^{c_m-1}R^{c_m-1}L^{c_m-2}\ldots L^{c_1}R^{c_0}\) in the Calkin-Wilf tree such that any path in the tree starting from the root and ending with \(L^{c_m-1}R^{c_m-1}L^{c_m-2}\ldots L^{c_1}R^{c_0}\) points to an element in \([A, B]\).

Since the probability is \(\frac{1}{2}\) for both \(\frac{a}{b} \mapsto \frac{a}{a+b}\) and \(\frac{a}{b} \mapsto \frac{a+b}{b}\), each pattern of fixed length \(m\) appears with the same probability and so we may restrict on the path pattern \(R^k\).

In the case \(k = 1\) we find in each generation exactly one which ends with \(R\) but does not contain any \(R\) before (actually, this is \(L^{n-1}R\) in generation \(n\)). Adding up all probabilities for these paths, we get

\[
\sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j = 1,
\]

and so, with probability 1, the random walk \(X_n\) will go to the right child for some \(n\). Now assume that the statement is true for \(k\). We shall show that then it is also true for \(k+1\). For each path of the form \(XR^k\), where \(X\) is any combination of powers of \(L\) and \(R\), there are two paths \(XR^kL\) and \(XR^{k+1}\), so by induction the probability that the random walk eventually follows the path \(R^{k+1}\) is at least \(\frac{1}{2}\). However, for each path \(XR^k\) one also has to consider the subtrees starting from \(XR^kL^d\) for \(d = 1, 2, \ldots\), each of which containing paths which end \(R^{k+1}\). By self-similarity, the probability that the random walk eventually follows the path \(R^{k+1}\) is

\[
\sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j \sum_{d=1}^{\infty} \left(\frac{1}{2}\right)^d = 1.
\]

This proves the theorem. \(\blacksquare\)
4. Statistical properties of the Calkin–Wilf tree

In view of Corollary 2 it is interesting to have a better understanding of the statistics of the Calkin–Wilf tree. The following theorem gives the limit distribution function in explicit form.

**Theorem 5.** Let \( F_n(x) \) denote the distribution function of the \( n \)-th generation, i.e.,

\[
F_n(x) = 2^{1-n} \# \{ j : x_j^{(n)} \leq x \}.
\]

Then uniformly \( F_n(x) \to F(x) \), where

\[
F([a_0, a_1, a_2, a_3, ...]) = 1 - 2^{-a_0} + 2^{-(a_0+a_1)} - 2^{-(a_0+a_1+a_2)} + ...
\]

(for rational numbers \( x = [a_0, a_1, ...] \) this series terminates at the last non-zero partial quotient of the continued fraction). Thus, \( F(0) = 0 \), \( F(\infty) = 1 \), and \( F(x) \) is a monotonically increasing function. Moreover, \( F(x) \) is continuous and singular, i.e., \( F'(x) = 0 \) almost everywhere.

**Proof.** Let \( x \geq 1 \). One half of the fractions in the \( n+1 \)-st generation do not exceed 1, and hence also do not exceed \( x \). Further,

\[
\frac{a + b}{b} \leq x \iff \frac{a}{b} \leq x - 1.
\]

Hence,

\[
2F_{n+1}(x) = F_n(x-1) + 1, \quad n \geq 1.
\]

Now assume \( 0 < x < 1 \). Then

\[
\frac{a}{a+b} \leq x \iff \frac{a}{b} \leq \frac{x}{1-x}.
\]

Therefore,

\[
2F_{n+1}(x) = F_n\left(\frac{x}{1-x}\right).
\]

The distribution function \( F \), defined in the formulation of the theorem, satisfies the functional equation

\[
2F(x) = \begin{cases} 
F(x-1) + 1 & \text{if } x \geq 1, \\
F\left(\frac{x}{1-x}\right) & \text{if } 0 < x < 1.
\end{cases}
\]

For instance, the second identity is equivalent to \( 2F(\frac{t}{t+1}) = F(t) \) for all positive \( t \). If \( t = [b_0, b_1, ...] \), then \( \frac{t}{t+1} = [0, b_0, b_1, ...] \) for \( t \geq 1 \), and \( \frac{t}{t+1} = [0, b_1+1, b_2, ...] \) for \( t < 1 \), and the statement follows immediately.
Now define $\delta_n(x) = F(x) - F_n(x)$. In order to prove the first assertion of the theorem, the uniform convergence $F_n \to F$, it is sufficient to show that
\[
\sup_{x \geq 0} |\delta_n(x)| \leq 2^{-n}.
\] (2)
It is easy to see that the assertion is true for $n = 1$. Now suppose the estimate is true for $n$. In view of the functional equation for both $F_n(x)$ and $F(x)$, we have
\[
2\delta_{n+1}(x) = \delta_n\left(\frac{x}{1-x}\right)
\] for $0 < x < 1$, which gives $\sup_{0 \leq x < 1} |\delta_{n+1}(x)| \leq 2^{-n-1}$. Moreover, we have
\[
2\delta_{n+1}(x) = \delta_n(x - 1)
\] for $x \geq 1$, which yields the same bound for $\delta_n(x)$ in the range $x \geq 1$. This proves (2).

Clearly, $F$, as a distribution function, is monotonic; obviously, it is also continuous. It remains to prove that $F(x)$ is singular. Given an irrational number $\alpha = [a_0, a_1, a_2, ...]$, we consider the sequence
\[
\alpha_n = [a_0, a_1, ..., a_{n-1}, a_n + 1, a_{n+1}, ...];
\] obviously, $\alpha_n$ is the real number which is defined by the continued fraction expansion of $\alpha$, where the $n$th partial quotient $a_n$ is replaced by $a_n + 1$. Depending on the parity of $n$, $\alpha_n$ is less than or greater than $\alpha$. Thus, any real number $y$, which is sufficiently close to $\alpha$, is contained between two terms of the sequence, $\alpha_L$ and $\alpha_{L+2}$ say. Then
\[
\left|\frac{F(y) - F(\alpha)}{y - \alpha}\right| \leq \left|\frac{F(\alpha_L) - F(\alpha)}{\alpha_{L+2} - \alpha}\right|.
\]
From the explicit form of $F$ we deduce
\[
|F(\alpha_L) - F(\alpha)| \leq \frac{1}{2}2^{-(a_0 + a_1 + ... + a_L)}.
\]
On the other hand,
\[
|\alpha_{L+2} - \alpha| \geq ( [a_1, a_2, ..., a_{L+2} + 1, ...] - [a_1, a_2, ..., a_{L+2}, ...]) (a_0 + 1)^{-2}
\] \[
\geq \left( (a_0 + 1)(a_1 + 1)...(a_{L+2} + 1) \right)^{-2}
\]
by induction. Thus,
\[ \left| \frac{F(y) - F(\alpha)}{y - \alpha} \right| \leq 2^{1-(a_0+a_1+\ldots+a_L)} \prod_{i=1}^{L+2} (a_i + 1)^2. \]

The theorem of Khinchin ([9], p. 86, implies that \( \prod_{i=1}^{n} (a_i + 1)^{1/n} \) tends to a fixed constant limit almost everywhere. On the other hand, the same reasoning shows that \( \frac{1}{n} \sum_{i=1}^{n} a_n \) tends to infinity for almost all \( x \). Thus, almost everywhere the limit
\[ \lim_{y \to \alpha} \left( F(y) - F(\alpha) \right) \frac{1}{(y - \alpha)^{-1}} \]
exists and is equal 0. This finishes the proof of the theorem. □

By the same argument as for the singular behaviour of \( F \) we can show that \( F'(\frac{\sqrt{5}+1}{2}) = \infty \). Actually, the terms of \( CW^{(n)} \) are densely concentrated around numbers with \( F'(x) = \infty \) and scarcely around those where \( F'(x) = 0 \). The value of \( F(x) \) is rational iff \( x \) is either rational or quadratic irrationality, e.g.
\[ F(1) = \frac{1}{2}, \quad F(\sqrt{2}) = \frac{3}{5}, \quad F((\sqrt{5} + 1)/2) = \frac{2}{3}. \]

This follows immediately from Lagrange’s theorem which characterizes the quadratic irrationals by their eventually periodic continued fraction expansion. For Euler’s number \( e = [2, 1, 2n, 1] \) we find that \( F(e) \) can be expressed in terms of special values of Jacobi theta functions.

5. Characteristics of the distribution function

In view of Corollary [2], the mean of the distribution function \( F \) is \( \frac{3}{2} \). Since \( F \) has a tail of exponential decay, more precisely \( 1 - F(x) = O(2^{-x}) \), it follows that all moments exist. For \( k \in \mathbb{N}_0 \), the \( k \)th moment is defined by
\[ M_k = \int_{0}^{\infty} x^k \, dF(x). \]

In order to give an asymptotic formula for \( M_k \) let
\[ m_k = \int_{0}^{\infty} \left( \frac{x}{x+1} \right)^k \, dF(x) \]
We will see that the generating function of $m_k$ has some interesting properties. Let $\omega(x)$ be a continuous function of at most power growth: $\omega(x) \ll x^T$ as $x \to \infty$. By the functional equation for $F$ we find $F(x+n) = 1 - 2^{-n} + 2^{-n}F(x)$, $x \geq 0$. Hence

$$
\int_0^\infty \omega(x) \, dF(x) = \sum_{n=0}^\infty \int_0^1 \omega(x+n) \, dF(x+n)
$$

$$
= \int_0^1 \sum_{n=0}^\infty \frac{\omega(x+n)}{2^n} \, dF(x);
$$

these integrals exist in view of our assumptions and the fact that $F(x)$ has a tail of exponential decay. Let $x = \frac{t}{t+1}$ for $t \geq 0$. Since $F\left(\frac{L}{t+1}\right) = \frac{1}{2}F(t)$, this change of variables gives

$$
\int_0^\infty \omega(x) \, dF(x) = \sum_{n=0}^\infty \int_0^\infty \frac{\omega\left(\frac{t}{t+1} + n\right)}{2^{n+1}} \, dF(t)
$$

(All changes of order of summation and integration are justified by the condition we put on $\omega(x)$). Now let $\omega(x) = x^L$ for some $L \in \mathbb{N}_0$ and define

$$b_s = \sum_{n=0}^\infty \frac{n^s}{2^{n+1}}.
$$

Then

$$
\int_0^\infty x^L \, dF(x) = \int_0^\infty \sum_{i=0}^L \binom{L}{i} \frac{x^i}{x+1} b_{L-i} \, dF(x),
$$

whence the relation

$$M_L = \sum_{i=0}^L m_i \binom{L}{i} b_{L-i}
$$

(3)

for $L \in \mathbb{N}_0$. The generating function of the sequence of the $b_s$ is given by

$$b(t) = \sum_{L=0}^\infty \frac{b_L t^L}{L!} = \sum_{L=0}^\infty \sum_{n=0}^\infty \frac{n^L t^L}{2^{n+1} L!} = \sum_{n=0}^\infty \frac{e^{nt}}{2^{n+1}} = \frac{1}{2 - e^t}.$$
Denote by \(M(t)\) and \(m(t)\) the corresponding generating functions of the coefficients \(M_k\) and \(m_k\), respectively. Then we can rewrite (3) as
\[
M(t) = \sum_{L=0}^{\infty} \frac{M_L}{L!} t^L = \frac{1}{2 - e^t} \sum_{L=0}^{\infty} \frac{m_L}{L!} t^L = \frac{1}{2 - e^t} m(t).
\]
The function \(m(t)\) is entire, and \(M(t)\) has a positive radius of convergence. This already allows us to find approximate values of the moments \(M_L\).

**Theorem 6.** For \(L \in \mathbb{N}_0\),
\[
M_L = \frac{m(\log 2)}{2 \log 2} \left( \frac{1}{\log 2} \right)^L L! + O_\varepsilon \left( ((4\pi^2 + (\log 2)^{1/2} - \varepsilon)^{-L} \right) L!
\]

**Proof.** By Cauchy’s formula, for any sufficiently small \(r\),
\[
M_L = \frac{L!}{2\pi i} \int_{|z|=r} \frac{M(z)}{z^{L+1}} \, dz.
\]
Changing the path of integration, we get by the calculus of residues
\[
M_L = -\text{Res}_{z=\log 2} \left( \frac{m(z)}{(2 - e^z)z^{L+1}} \right) - \frac{L!}{2\pi i} \int_{|z|=R} \frac{m(z)}{2 - e^z z^{L+1}} \, dz,
\]
where \(R\) satisfies \(\log 2 < R < |\log 2 + 2\pi i|\) (which means that there is exactly one simple pole of the integrand located in the interior of the circle \(|z| = R\)). It is easily seen that the residue coincides with the main term in the formula of the lemma; the error term follows from estimating the integral. ■

We obtain the inverse to the linear equations (3):
\[
m_L = M_L - \sum_{s=0}^{L-1} M_s \binom{L}{s}
\]
for \(L \in \mathbb{N}_0\). Since \(b(t)(2 - e^t) = 1\), the coefficients \(b_L\) can be calculated recursively
\[
b_L = \sum_{s=0}^{L-1} \binom{L}{s} b_s.
\]
Thus, \(b_0 = 1, b_1 = 1, b_2 = 3, b_3 = 13, b_4 = 75, b_5 = 541\).
We proceed with a property of the function $m(t)$ which reflects the symmetry of the distribution function: $F(y) + F(1/y) = 1$. Unfortunately, this property is still insufficient for determining the coefficients $m_L$. As a matter of fact,

$$m_L = \int_0^\infty \left(\frac{x}{x+1}\right)^L dF(x) = -\int_0^\infty \left(\frac{1/x}{1/x+1}\right)^L dF(1/x)$$

$$= \int_0^\infty \left(\frac{1}{x+1}\right)^L dF(x).$$

Since

$$\left(\frac{x}{x+1}\right)^L = \left(\frac{x+1-1}{x+1}\right)^L = \sum_{s=0}^L \binom{L}{s} (-1)^s \left(\frac{1}{x+1}\right)^{L-s},$$

this gives

$$m_L = \sum_{s=0}^L \binom{L}{s} (-1)^s m_s$$

for $L \geq 0$. For example, $m_1 = m_0 - m_1$, which gives $m_1 = \frac{1}{2}$ (since $m_0 = 1$), and thus $M_1 = \frac{3}{2}$ (see Theorem 1). For the other coefficients we only get linear relations. Thus, $2m_3 = -\frac{1}{2} + 3m_2$. In terms of $m(t)$ the recursion formula above yields the identity

$$m(t) = m(-t)e^t.$$

We conclude this chapter with the result, which uniquely determines the function $m(t)$ (along with the condition $m(0) = 1$).

**Theorem 7.** The function $m(s)$ satisfies the integral equation

$$m(-s) = (2e^s - 1) \int_0^\infty m'(-t)J_0(2\sqrt{st}) dt, \quad s \in \mathbb{R}_+,$$

where $J_0(*)$ stands for the Bessel function:

$$J_0(z) = \frac{1}{\pi} \int_0^\pi \cos(z \sin x) \, dx.$$
The proof of this theorem and the solution of this integral equation, and thus the explicit description of the moments will be given in a subsequent paper.

6. \(p\)-adic distribution

In the previous sections, we were interested in the distribution of the \(n\)th generation of the tree \(\mathcal{CW}\) in the field of real numbers. Since the set of non-equivalent valuations of \(\mathbb{Q}\) contains a valuation, associated with any prime number \(p\), it is natural to consider the distribution of the set of each generation in the field of \(p\)-adic numbers \(\mathbb{Q}_p\). In this case we have an ultrametric inequality, which implies that two circles are either co-centric, or do not intersect. We define

\[
F_n(z, \nu) = 2^{-n+1} \# \left\{ \frac{a}{b} \in \mathcal{CW}^{(n)} : \text{ord}_p \left( \frac{a}{b} - z \right) \geq \nu \right\}, \quad z \in \mathbb{Q}_p, \quad \nu \in \mathbb{Z}.
\]

(When \(p\) is fixed, the subscript \(p\) in \(F\) is omitted). Note that in order to calculate \(F_n(z, \nu)\) we can confine to the case \(\text{ord}_p(z) < \nu\); otherwise \(\text{ord}_p(z) \geq \nu \iff \text{ord}_p(\frac{a}{b}) \geq \nu\). We shall calculate the limit distribution \(\mu_p(z, \nu) = \lim_{n \to \infty} F_n(z, \nu)\), and also some characteristics of it, e.g. the zeta function

\[
Z_p(s) = \int_{u \in \mathbb{Q}_p} |u|^s d\mu_p, \quad s \in \mathbb{C}, \quad z \in \mathbb{Q}_p,
\]

where \(|*|\) stands for the \(p\)-adic valuation.

To illustrate how the method works, we will calculate the value of \(F_n\) in two special cases. Let \(p = 2\), and let \(E(n)\) be the number of rational numbers in the \(n\)th generation with one of \(a\) or \(b\) being even, and let \(O(n)\) be the corresponding number fractions with both \(a\) and \(b\) odd. Then \(E(n) + O(n) = 2^{n-1}\). Since \(\frac{a}{b}\) in the \(n\)th generation generates \(\frac{a}{a+b}\) and \(\frac{2a+k}{b}\) in the \((n+1)\)st generation, each fraction \(\frac{a}{b}\) with one of the \(a, b\) even will generate one fraction with both numerator and denominator odd. If both \(a, b\) are odd, then their two offsprings will not be of this kind. Therefore, \(O(n+1) = E(n)\). Similarly, \(E(n+1) = E(n) + 2O(n)\). This gives the recurrence \(E(n+1) = E(n) + 2E(n-1), n \geq 2\), and
this implies
\[ E(n) = \frac{2n + 2(-1)^n}{3}, \quad O(n) = \frac{2^{n-1} + 2(-1)^{n-1}}{3}, \quad \mu_2(0, 0) = \frac{2}{3}. \]

(For the last equality note that \( \frac{a}{b} \) and \( \frac{b}{a} \) simultaneously belong to \( CW(n) \), and so the number of fractions with \( \text{ord}_2(*) > 0 \) is \( E(n)/2 \).

We will generalize this example to odd prime \( p \geq 3 \). Let \( L_i(n) \) be the part of the fractions in the \( n \)th generations such that \( \frac{a}{b} \equiv i \mod p \) for \( 0 \leq i \leq p - 1 \) or \( i = \infty \) (that is, \( b \equiv 0 \mod p \)). Thus,

\[ \sum_{i \in \mathbb{F}_p \cup \{\infty\}} L_i(n) = 1; \]

in other words, \( L_i(n) = F_n(i, 1) \). For our later investigations we need a result from the theory of finite Markov chains.

**Lemma 1.** Let \( A \) be a matrix of a finite Markov chain with \( s \) stages. That is, \( a_{i,j} \geq 0 \), and \( \sum_{j=1}^{s} a_{i,j} = 1 \) for all \( i \). Suppose that \( A \) is irreducible (for all pairs \( (i, j) \), and some \( m \), the entry \( a_{i,j}^{(m)} \) of the matrix \( A^m \) is strictly positive), acyclic and recurrent (this is satisfied, if all entries of \( A^m \) are strictly positive for some \( m \)). Then the eigenvalue 1 is simple, if \( \lambda \) is another eigenvalue, then \( |\lambda| < 1 \), and \( A^m \), as \( m \to \infty \), tends to the matrix \( B \), with entries \( b_{i,j} = \pi_j \), where \( (\pi_1, ..., \pi_s) \) is a unique left eigenvector with eigenvalue 1, such that \( \sum_{j=1}^{s} \pi_j = 1 \).

A proof of this lemma can be found in [6], Section 3.1., Theorem 1.3.

**Theorem 8.** \( \mu_p(z, 1) = \frac{1}{p+1} \) for \( z \in \mathbb{Z}_p \).

**Proof.** Similarly as in the above example, a fraction \( \frac{a}{b} \) from the \( n \)th generation generates \( \frac{a}{a+b} \) and \( \frac{a+b}{b} \) in the \( (n+1) \)st generation, and it is routine to check that

\[ L_i(n+1) = \frac{1}{2} L_{i-1}(n) + \frac{1}{2} L_{i-1}(n) \quad \text{for} \quad i \in \mathbb{F}_p \cup \{\infty\}, \quad (4) \]

(Here we make a natural convention for \( \frac{i}{1-i} \) and \( i-1 \), if \( i = 1 \) or \( \infty \)). In this equation, it can happen that \( i - 1 \equiv \frac{i}{1-i} \mod p \); thus, \( (2i - 1)^2 \equiv -3 \mod p \). The recurrence for this particular \( i \) is to be understood in the obvious way, \( L_i(n+1) = L_{i-1}(n) \). Therefore, if we
denote the vector-column $(L_\infty(n), L_0(n), ..., L_{p-1}(n))^T$ by $v_n$, and if $A$ is a matrix of the system (4), then $v_{n+1} = Av_n$, and hence

$$v_n = A^{n-1}v_1,$$

where $v_1 = (0, 0, 1, 0, ..., 0)^T$. In any particular case, this allows us to find the values of $L_i$ explicitly. For example, if $p = 7$, the characteristic polynomial is

$$f(x) = \frac{1}{16}(x - 1)(2x - 1)(2x^2 + 1)(4x^4 + 2x^3 + 2x + 1).$$

The list of roots is

$$\alpha_1 = 1, \quad \alpha = \frac{1}{2}, \quad \alpha_3,4 = \pm \frac{i}{\sqrt{2}}, \quad \alpha_{5,6,7,8} = \frac{-1 - \sqrt{17}}{8} \pm \frac{\sqrt{1 + \sqrt{17}}}{2\sqrt{2}},$$

(with respect to the two values for the root $\sqrt{17}$), the matrix is diagonalisable, and the Jordan normal form gives the expression

$$L_i(n) = \sum_{s=1}^{8} C_{i,s} a_s^n.$$

Note that the elements in each row of the $(p + 1) \times (p + 1)$ matrix $A$ are non-negative and sum up to 1, and thus, we have a matrix of a finite Markov chain. We need to check that it is acyclic. Let $\tau(i) = i - 1$, and $\sigma(i) = \frac{i}{1-i}$ for $i \in \mathbb{F}_p \cup \{\infty\}$. The entry $a_{i,j}^{(m)}$ of $A^m$ is

$$a_{i,j}^{(m)} = \sum_{i_1, \ldots, i_{m-1}} a_{i,i_1} \cdot a_{i_1,i_2} \cdot \ldots \cdot a_{i_{m-1},j}.$$

Therefore, we need to check that for some fixed $m$, the composition of $m$ $\sigma$’s or $\tau$’s leads from any $i$ to any $j$. One checks directly that for any positive $k$, and $i, j \in \mathbb{F}_p$,

$$\tau^{p-1-j} \circ \sigma \circ \tau^k \circ \sigma \circ \tau^{i-1}(i) = j,$$

$$\tau^{p-1-j} \circ \sigma \circ \tau^k(\infty) = j,$$

$$\tau^k \circ \sigma \circ \tau^{i-1}(i) = \infty;$$

(for $i = 0$, we write $\tau^{-1}$ for $\tau^{p-1}$). For each pair $(i, j)$, choose $k$ in order the amount of compositions used to be equal (say, to $m$). Then obviously all entries of $A^m$ are positive, and this matrix satisfies the conditions of lemma. Since all columns also sum up to 1, $(\pi_1, ..., \pi_{p+1})$, 

| ST ATISTICAL PROPERTIES OF THE CALKIN–WILF TREE 15 |

[173x690]
\[ \pi_j = \frac{1}{p+1}, \quad 1 \leq j \leq p + 1, \] is the needed eigenvector. This proves the theorem. ■

**Theorem 9.** Let \( \nu \in \mathbb{Z} \) and \( z \in \mathbb{Q}_p \), and \( \text{ord}_p(z) < \nu \) (or \( z = 0 \)). Then, if \( z \) is \( p \)-adic integer,

\[ \mu(z, \nu) = \frac{1}{p^\nu + p^{\nu-1}}. \]

If \( z \) is not integer, \( \text{ord}_p(z) = -\lambda < 0 \),

\[ \mu(z, \nu) = \frac{1}{p^{\nu+2\lambda} + p^{\nu+2\lambda-1}}. \]

For \( z = 0, -\nu \leq 0 \), we have

\[ \mu(0, -\nu) = 1 - \frac{1}{p^{\nu+1} + p^{\nu}}. \]

This theorem allows the computation of the associated zeta-function:

**Corollary 10.** For \( s \) in the strip \(-1 < \Re s < 1\),

\[ Z_p(s) = \int_{u \in \mathbb{Q}_p} |u|^s d\mu_p = \frac{(p-1)^2}{(p-p^{-s})(p-p^s)}, \]

and \( Z_p(s) = Z_p(-s) \).

The proof is straightforward. It should be noted that this expression encodes all the values of \( \mu(0, \nu) \) for \( \nu \in \mathbb{Z} \).

**Proof of Theorem 9.** For shortness, when \( p \) is fixed, denote \( \text{ord}_p(*) \) by \( v(*) \). As before, we want a recurrence relation among the numbers \( F_n(i, \kappa), \quad i \in \mathbb{Q}_+ \). For each integral \( \kappa \), we can confine to the case \( i < p^\kappa \). If \( i = 0 \), we only consider \( \kappa > 0 \) and call these pairs \((i, \kappa)\) "admissible". We also include \( G_n(0, -\kappa) \) for \( \kappa \geq 1 \), where these values are defined in the same manner as \( F_n \), only inverting the inequality, considering \( \frac{a}{b} \in CW^{(n)} \), such that \( v(\frac{a}{b}) \leq -\kappa \); the ratio of fractions in the \( n \)th generation outside this circle. As before, a fraction \( \frac{a}{b} \) in the \( n \)th generation generates the fractions \( \frac{a}{a+b} \) and \( \frac{a+b}{b} \) in the \((n+1)\)st generation. Let \( \tau(i, \kappa) = ((i-1) \mod p^\kappa, \kappa) \). Then for all admissible pairs \((i, \kappa), \quad i \neq 0 \), the pair \( \tau(i, \kappa) \) is also admissible, and

\[ v(\frac{a+b}{b} - i) = \kappa \iff v(\frac{a}{b} - (i-1)) = \kappa. \]
Thus, we have an infinite matrix $A$ Markov chain. If $a \neq 1$, $u \in \mathbb{Z}_p$, and $(i, \kappa)$ is admissible, then

$$\frac{a}{b} - \frac{i}{1 - i} = \frac{p^\kappa u}{(1 - i)(1 - p^\kappa u)}.$$ 

Since $v(\frac{a}{1 - \frac{1}{i}}) = v(i) - v(1 - i)$, this is 0 unless $i$ is an integer, equals to $v(i)$ if the latter is $> 0$ and equals to $-v(1 - i)$ if $v(1 - i) > 0$. Further, this difference has valuation $\geq \kappa_0 = \kappa$, if $i \in \mathbb{Z}, i \neq 1 \mod p$, valuation $\geq \kappa_0 = \kappa - 2v(1 - i)$, if $i \in \mathbb{Z}, i \equiv 1 \mod p$, and valuation $\geq \kappa_0 = \kappa - 2v(i)$ if $i$ is not integer. In all three cases, easy to check, that, if we define $i_0 = \frac{i}{1 + i}$, $u \in \mathbb{Z}_p$. Then

$$\frac{a}{a + b} - \frac{i_0}{1 + i_0} = \frac{p^\kappa}{(1 + i_0)(1 + p^\kappa u)}.$$ 

If $i = \frac{i_0}{1 + i_0}$ is a $p$–adic integer, $i \neq 1 \mod p$, this has a valuation $\geq \kappa = \kappa_0$; if $i$ is a $p$–adic integer, $i \equiv 1(p)$, this has valuation

$$\geq \kappa = \kappa_0 - 2v(i_0) = \kappa_0 + 2v(1 - i);$$

if $i$ is not a $p$–adic integer, this has valuation

$$\geq \kappa = \kappa_0 - 2v(1 + i_0) = \kappa_0 + 2v(i).$$

Thus,

$$v\left(\frac{a}{a + b} - i\right) \geq \kappa \iff v\left(\frac{a}{b} - i_0\right) \geq \kappa_0.$$ 

Let $i = 1$. If $\frac{a}{a + b} = 1 + p^\kappa u$, then $\kappa > 0$, $u \in \mathbb{Z}_p$, and we obtain $\frac{a}{b} = -1 - \frac{1}{p^\kappa u}$, $v\left(\frac{a}{b}\right) \leq -\kappa$. Converse is also true. Finally, for $\kappa \geq 1$,

$$v\left(\frac{a + b}{b}\right) \leq -\kappa \iff v\left(\frac{a}{b}\right) \leq -\kappa,$$

and

$$v\left(\frac{a}{a + b}\right) \leq -\kappa \iff v\left(\frac{a}{b} + 1\right) \geq \kappa.$$ 

Therefore, we have the recurrence relations:

$$
F_{n+1}(i, \kappa) = \frac{1}{2}F_n(\tau(i, \kappa)) + \frac{1}{2}F_n(\sigma(i, \kappa)), \text{ if } (i, \kappa) \text{ is admissible},
F_{n+1}(1, \kappa) = \frac{1}{2}F_n(0, \kappa) + \frac{1}{2}G_n(0, -\kappa), \kappa \geq 1,
G_{n+1}(0, -\kappa) = \frac{1}{2}G_n(0, -\kappa) + \frac{1}{2}F_n(-1, \kappa), \kappa \geq 1.
$$

(5)

Thus, we have an infinite matrix $A$, which is a change matrix for the Markov chain. If $v_n$ is an infinite vector-column of $F_n$'s and $G_n$'s, then $v_{n+1} = Av_n$, and, as before, $v_n = A^{n-1}v_1$. It is direct to check that
each column also contains exactly two nonzero entries $\frac{1}{2}$, or one entry, equal to 1. In terms of Markov chains, we need to determine the classes of orbits. Then in proper rearranging, the matrix $A$ looks like
\[
\begin{pmatrix}
P_1 & 0 & \cdots & 0 & \cdots \\
0 & P_2 & \cdots & 0 & \cdots \\
\vdots & \ddots & \ddots & \vdots & \ddots \\
0 & 0 & \cdots & P_s & 0 \\
\vdots & \vdots & \cdots & 0 & \ddots
\end{pmatrix},
\]
where $P_s$ are finite Markov matrices. Thus, we claim that the length of each orbit is finite, every orbit has a representative $G_s(0, -\kappa)$, $\kappa \geq 1$, the length of it is $p^\kappa + p^{\kappa-1}$, and the matrix is recurrent (that is, every two positions communicate). In fact, from the system above and form the expression of the maps $\tau(i, \kappa)$ and $\sigma(i, \kappa)$, the direct check shows that the complete list of the orbit of $G_s(0, -\kappa)$ consists of (and each pair of states are communicating):
\[
G_s(0, -\kappa), \\
F_s(i, \kappa) \quad (i = 0, 1, 2, \ldots, p^\kappa - 1), \\
F_s(p^{-\lambda}u, \kappa - 2\lambda) \quad (\lambda = 1, 2, \ldots, \kappa - 1, u \in \mathbb{N}, u \not\equiv 0 \mod p, u \leq p^{\kappa-\lambda}).
\]
In total, we have
\[
1 + p^\kappa + \sum_{\lambda=1}^{\kappa-1}(p^{\kappa-\lambda} - p^{\kappa-\lambda-1}) = p^\kappa + p^{\kappa-1}
\]
members in the orbit. Thus, each $P_\kappa$ in the matrix above is a finite dimensional $\ell_\kappa \times \ell_\kappa$ matrix, where $\ell_\kappa = p^\kappa + p^{\kappa-1}$. For $\kappa = 1$, the matrix $P_1$ is exactly the matrix of the system (4). As noted above, the vector column $(1, 1, ..., 1)^T$ is the left eigen-vector. As in the previous theorem, it is straightforward to check that this matrix is irreducible and acyclic (that is, the entries of $P_\kappa^n$ are strictly positive for sufficiently large $n$). In fact, since by our observation, each two members in the orbit communicate, and since we have a move $G_s(0, -\kappa) \rightarrow G_s(0, -\kappa)$, the proof of the last statement is immediate: there exists $n$ such that any position is reachable from another in exactly $n$ moves, and this can be achieved at the expense of the move just described. Therefore,
all entries of $P_n$ are strictly positive. Thus, the claim of the theorem follows from the lemma above.

Acknowledgements. The second author thanks Jürgen Sander and Jan-Hendrik de Wiljes for their interest and valuable remarks.

References

[1] N. Calkin, H. Wilf, Recounting the rationals. Amer. Math. Monthly 107 (2000), 360–363
[2] A.J. Cole, A.J.T. Davie, A game based on the Euclidean algorithm and a winning strategy for it. Math. Gaz. 53 (1969), 354-357
[3] R. Bird, J. Gibbons, D. Lester, Functional pearl: enumerating the rationals. J. Functional Programming 16 (2006), 281-291
[4] R. Graham, D. Knuth, O. Patashnik, Concrete Mathematics: a foundation for computer science. Addison-Wesley 1990
[5] S. Hofmann, G. Schuster, J. Steuding, Euclid, Calkin & Wilf — Playing with rationals —, Elem. Math. (to appear)
[6] S. Karlin, A first course in stochastic processes. Academic Press, 1968
[7] D.H. Lehmer, On Stern’s diatomic series, Amer. Math. Monthly 36 (1929), 59-67
[8] M. Newman, Recounting the rationals, Continued. Credited in Amer. Math. Monthly 110 (2003), 642–643
[9] A. Ya. Khinchin, Continued fractions. The University of Chicago Press, 1964
[10] B. Reznick, Regularity properties of the Stern enumeration of the rationals. Preprint available at [http://arxiv.org/pdf/math.NT/0610601.pdf?front]
[11] N. Sloane, S. Plouffe, The On-line Encyclopedia of integer sequences, [http://www.research.att.com/~njas/sequences/]
[12] M.A. Stern, Über eine zahlentheoretische Funktion. J. reine angew. Math. 55 (1858), 193-220

Giedrius Alkauskas, Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, 03225 Vilnius, Lithuania
giedrius.alkauskas@maths.nottingham.ac.uk

Jörn Steuding, Department of Mathematics, Würzburg University, Am Hubland, 97 218 Würzburg, Germany
steuding@mathematik.uni-wuerzburg.de