# BBP-numbers and normality 

Bacheloronderzoek Wiskunde
Augustus 2008
Student: J.M. Heegen
Studentnummer: 1056484
Begeleiders: Prof.dr. J. Top en Prof.dr. H. Waalkens

## Contents

Introduction ..... 2

1. Normal numbers ..... 3
2. The Bailey-Borwein-Plouffe formula ..... 4
3. Using BBP-type formulae to compute individual digits of arithmetical constants ..... 7
4. Computing digits in Mathematica ..... 9
5. Generating correct base-b expansions for BBP-numbers ..... 12
6. Normality of BBP-numbers ..... 16
7. Rational BBP-numbers ..... 18
8. Conclusion ..... 22
Bibliography ..... 24

## Introduction

One of the many unsolved problems in mathematics is the question whether the digits of various arithmetical constants exhibit a certain randomness, that is, if one randomly picks a string of some length out of the expansion of the constant, each possible string of that length is equally likely to come up. This is the notion of a so-called normal number.
To take a first step in the direction of a possible solution for this problem, we will introduce a very special class of numbers that do not only possess a method to generate digit expansions with increasing accuracy, and in some cases, even exactly; but also, under a certain hypothesis, are provably normal. This class of numbers include some fundamental, naturally occurring constants, most notably the fascinating number $\pi$.

## 1. Normal numbers

Definition 1.1. Let $\alpha$ be a real number and $b \geq 2$ an integer. The number $\alpha$ is called normal to base $b$ if every string of $k$ digits occurs in the base-b expansion of $\alpha$ with a limiting frequency of $b^{-k}$.
If $\alpha$ is normal to every integer base $b \geq 2$, it is called absolutely normal.
For example, if $\alpha$ is normal to base 2 , the digits 0 and 1 will both appear in the binary expansion of $\alpha$ with an equal frequency of $\frac{1}{2}$; the strings $00,01,10$ and 11 will all appear with an equal frequency of $\frac{1}{4}$, and so on.

In the case of a number that is normal to base 10 , every digit $0,1,2, \ldots, 9$ will appear one-tenth of the time in its decimal expansion. A randomly picked string of five digits will have a probability of $\frac{1}{100000}$ to be " 07817 ", and so on.
For a number that is normal to base 16, in its hexadecimal expansion (with a 'digit alphabet' consisting of the integers 0 through 9 and the capital letters A through F) the probabilities for a randomly picked 6 -digit string to be " 1 CACA 0 " or "C0FFEE" are both equal to $16^{-6}$.

The expansion of a rational number is eventually periodic, i.e. from some position in the expansion, a particular string of digits is repeated infinitely often. This means that there are strings of digits that never appear, so the number cannot be normal. Hence, every normal number must be irrational.

There are uncountably many numbers that are not normal. This is easily seen by the fact that, for instance, there are uncountably many numbers without a 0 in their decimal expansion, and these are obviously not normal.

Some important properties of normal numbers are the following [Kuipers \& Niederreiter, 1974]:

- Let $b, k, m \in \mathbb{N}, b \geq 2$ and $\alpha \in \mathbb{R}$, such that $\alpha$ is normal to base $b^{k}$. Then $\alpha$ is also normal to base $b^{m}$.
- Let $q, r$ be rational numbers with $r \neq 0$. If $\alpha$ is normal to base $b$, then $r \alpha+q$ is also normal to base $b$. Furthermore, $\alpha$ is normal to base $c=b^{r}$, in case $c$ is an integer.

An important result by Émile Borel, who first introduced the concept of normality in the beginning of the 20th century, is that almost all real numbers are absolutely normal; that is, the set of all real numbers that are not absolutely normal has a Lebesgue measure of zero.

Despite the abundance of such numbers, it has turned out to be very difficult to prove the normality of a given irrational number. Even though empirical data strongly seems to suggest that constants like $\pi, e, \log 2, \sqrt{2}$, the golden mean $\frac{1}{2}(1+\sqrt{5})$ (and in fact all irrational algebraic numbers) and numerous others are absolutely normal, yet no proof nor disproof of normality in any base has ever been given for any of these numbers.

Normality has been established for some numbers that have been specifically constructed for these purposes. An example is the Champernowne constant, which is obtained by concatenating (i.e.
putting behind one another in ascending order) the positive integers:

$$
C_{10}=0.123456789101112131415 \ldots
$$

This number has been shown to be normal to base 10. Similarly, the binary Champernowne constant $C_{2}=0.11011100101110111 \ldots$ is known to be normal to base 2 , the ternary Champernowne constant $C_{3}=0.12101112202122 \ldots$ is normal to base 3 , etcetera.
Another artificially constructed number, the Copeland-Erdős constant, which is obtained by concatenating all primes: $0.2357111317192329 \ldots$, is normal to base 10 .

A fairly recent result [Bailey, 2005] is that all constants of the form

$$
\alpha_{p, q}=\sum_{k=1}^{\infty} \frac{1}{p^{q^{k}} q^{k}}
$$

in which $p$ and $q$ are relatively prime, the so-called Stoneham numbers, are normal to base $p$.

However, it has long been a desire to prove the widely conjectured normality of the aforementioned "naturally occurring" constants $(\pi, e, \log 2, \sqrt{2}, \ldots)$. An important new direction in the quest for these proofs has been initiated by the discovery of a formula known as the Bailey-BorweinPlouffe formula.

## 2. The Bailey-Borwein-Plouffe formula

The Bailey-Borwein-Plouffe formula (shortly: BBP) is a formula for computing $\pi$. It is named after Simon Plouffe, who discovered it in 1995, and David Bailey and Peter Borwein, the co-authors of the paper in which it was first published. It states that

$$
\begin{equation*}
\pi=\sum_{i=0}^{\infty} \frac{1}{16^{i}}\left(\frac{4}{8 i+1}-\frac{2}{8 i+4}-\frac{1}{8 i+5}-\frac{1}{8 i+6}\right) \tag{1}
\end{equation*}
$$

More generally, a BBP-type formula is a convergent series of the form

$$
\alpha=\sum_{k=0}^{\infty} \frac{p(k)}{b^{k} q(k)}
$$

where $b \geq 2$ is an integer and $p, q$ are polynomials with integer coefficients. The BBP-fomula for $\pi$ is of this type, because we can combine the four fractions into one. A real number $\alpha$ that can be represented in this way is called a $B B P$-number.

## Proof of the BBP formula for $\pi$.

For any $k<8$,

$$
\int_{0}^{1 / \sqrt{2}} \frac{x^{k-1}}{1-x^{8}} d x=\int_{0}^{1 / \sqrt{2}} x^{k-1} \cdot \frac{1}{1-x^{8}} d x=\int_{0}^{1 / \sqrt{2}} x^{k-1}\left(\sum_{i=0}^{\infty}\left(x^{8}\right)^{i}\right) d x
$$

(we can substitute the geometric series here, because $\left|x^{8}\right|<1$ for $0 \leq x \leq \frac{1}{\sqrt{2}}$ )

$$
\begin{gathered}
=\int_{0}^{1 / \sqrt{2}}\left(\sum_{i=0}^{\infty} x^{k-1+8 i}\right) d x=\sum_{i=0}^{\infty}\left(\int_{0}^{1 / \sqrt{2}} x^{k-1+8 i} d x\right)=\sum_{i=0}^{\infty}\left[\frac{1}{k+8 i} x^{k+8 i}\right]_{x=0}^{x=1 / \sqrt{2}} \\
=\sum_{i=0}^{\infty} \frac{1}{k+8 i}\left(\frac{1}{\sqrt{2}}\right)^{k+8 i}=\sum_{i=0}^{\infty} \frac{1}{k+8 i} \cdot \frac{1}{2^{k / 2}} \cdot \frac{1}{16^{i}}=\frac{1}{2^{k / 2}} \sum_{i=0}^{\infty} \frac{1}{16^{i}(8 i+k)} .
\end{gathered}
$$

Now define

$$
\begin{aligned}
& T_{k}:=\sum_{i=0}^{\infty} \frac{1}{16^{i}(8 i+k)}, \text { so that } T_{k}=2^{k / 2} \int_{0}^{1 / \sqrt{2}} \frac{x^{k-1}}{1-x^{8}} d x . \text { Then } \\
& \begin{aligned}
& \sum_{i=0}^{\infty} \frac{1}{16^{i}}\left(\frac{4}{8 i+1}-\frac{2}{8 i+4}-\frac{1}{8 i+5}-\frac{1}{8 i+6}\right)=4 T_{1}-2 T_{4}-T_{5}-T_{6} \\
&= 4 \cdot 2^{1 / 2} \int_{0}^{1 / \sqrt{2}} \frac{1}{1-x^{8}} d x-2 \cdot 2^{4 / 2} \int_{0}^{1 / \sqrt{2}} \frac{x^{3}}{1-x^{8}} d x \\
&-2^{5 / 2} \int_{0}^{1 / \sqrt{2}} \frac{x^{4}}{1-x^{8}} d x-2^{6 / 2} \int_{0}^{1 / \sqrt{2}} \frac{x^{5}}{1-x^{8}} d x \\
& \quad=\int_{0}^{1 / \sqrt{2}} \frac{4 \sqrt{2}-8 x^{3}-4 \sqrt{2} x^{4}-8 x^{5}}{1-x^{8}} d x .
\end{aligned}
\end{aligned}
$$

We now substitute $y:=\sqrt{2} x$, so that $x=\frac{1}{\sqrt{2}} y$ and $d x=\frac{1}{\sqrt{2}} d y$, and the above expression becomes

$$
\begin{gathered}
\int_{0}^{1} \frac{4 \sqrt{2}-8 \cdot \frac{1}{2 \sqrt{2}} y^{3}-4 \sqrt{2} \cdot \frac{1}{4} y^{4}-8 \cdot \frac{1}{4 \sqrt{2}} y^{5}}{1-\frac{1}{16} y^{8}} \cdot \frac{1}{\sqrt{2}} d y \\
=\int_{0}^{1} 16 \cdot \frac{4-2 y^{3}-y^{4}-y^{5}}{16-y^{8}} d y=16 \int_{0}^{1} \frac{y^{5}+y^{4}+2 y^{3}-4}{y^{8}-16} d y \\
=16 \int_{0}^{1} \frac{(y-1)\left(y^{4}+2 y^{3}+4 y^{2}+4 y+4\right)}{\left(y^{4}-2 y^{3}+4 y-4\right)\left(y^{4}+2 y^{3}+4 y^{2}+4 y+4\right)} d y \\
=16 \int_{0}^{1} \frac{(y-1)}{y^{4}-2 y^{3}+4 y-4} d y=16 \int_{0}^{1} \frac{y-1}{\left(y^{2}-2\right)\left(y^{2}-2 y+2\right)} d y .
\end{gathered}
$$

To evaluate this integral, we will first decompose the integrand into partial fractions. We need to find $A, B, C, D \in \mathbb{R}$ such that

$$
\frac{y-1}{\left(y^{2}-2\right)\left(y^{2}-2 y+2\right)}=\frac{A y+B}{y^{2}-2}+\frac{C y+D}{y^{2}-2 y+2} .
$$

This means that

$$
\begin{aligned}
y-1 & =(A y+B)\left(y^{2}-2 y+2\right)+(C y+D)\left(y^{2}-2\right) \\
& =(A+C) y^{3}+(-2 A+B+D) y^{2}+(2 A-2 B-2 C) y+(2 B-2 D)
\end{aligned}
$$

So the integral becomes

$$
\begin{gathered}
16 \int_{0}^{1} \frac{\frac{1}{4} y}{y^{2}-2}+\frac{-\frac{1}{4} y+\frac{1}{2}}{y^{2}-2 y+2} d y \\
=2 \int_{0}^{1} \frac{2 y}{y^{2}-2} d y-2 \int_{0}^{1} \frac{2 y-2}{y^{2}-2 y+2} d y+4 \int_{0}^{1} \frac{1}{y^{2}-2 y+2} d y \\
=2\left[\log \left|y^{2}-2\right|\right]_{y=0}^{y=1}-2\left[\log \left|y^{2}-2 y+2\right|\right]_{y=0}^{y=1}+4[\arctan (y-1)]_{y=0}^{y=1} \\
=2 \log 1-2 \log 2-2 \log 1+2 \log 2+4 \arctan 0-4 \arctan (-1) \\
=-4 \arctan (-1)=\pi .
\end{gathered}
$$

Hence the BBP-formula (1) has been proved.
The above is a formal analytical proof of the BBP-formula. However, the way it was originally discovered was by doing numerical searches on a computer with the PSLQ-algorithm, an algorithm designed to find integer relations between real numbers $x_{1}, \ldots, x_{n}$ of the form $\sum_{i=1}^{n} a_{i} x_{i}=0$, where $a_{i} \in \mathbb{Z}$ but not all equal to zero.

## Other BBP-type formulae

Various other, similar BBP-type formulae are known, some of which have been found using the same PSLQ-algorithm. We list a few examples here.

$$
\begin{align*}
\pi= & \frac{1}{2} \sum_{k=0}^{\infty} \frac{1}{16^{k}}\left(\frac{8}{8 k+2}+\frac{4}{8 k+3}+\frac{4}{8 k+4}-\frac{1}{8 k+7}\right) \\
\pi= & \frac{1}{2^{6}} \sum_{k=0}^{\infty} \frac{(-1)^{k}}{2^{10 k}}\left(-\frac{2^{5}}{4 k+1}-\frac{1}{4 k+3}+\frac{2^{8}}{10 k+1}\right. \\
& \left.\quad-\frac{2^{6}}{10 k+3}-\frac{2^{2}}{10 k+5}-\frac{2^{2}}{10 k+7}+\frac{1}{10 k+9}\right)  \tag{2}\\
\pi^{2}= & \frac{9}{8} \sum_{k=0}^{\infty} \frac{1}{64^{k}}\left(\frac{16}{(6 k+1)^{2}}-\frac{24}{(6 k+2)^{2}}-\frac{8}{(6 k+3)^{2}}-\frac{6}{(6 k+4)^{2}}+\frac{1}{(6 k+5)^{2}}\right) \\
\log 2= & \sum_{k=1}^{\infty} \frac{1}{k 2^{k}}  \tag{3}\\
\log 2= & \frac{2}{3} \sum_{k=0}^{\infty} \frac{1}{9^{k}(2 k+1)} \\
\log \left(\frac{10}{9}\right)= & \sum_{k=1}^{\infty} \frac{1}{k 10^{k}} \tag{4}
\end{align*}
$$

$$
\begin{aligned}
\log 3= & \sum_{k=0}^{\infty} \frac{1}{4^{k}(2 k+1)} \\
\arctan \left(\frac{1}{3}\right)= & \sum_{k=0}^{\infty} \frac{1}{16^{k}}\left(\frac{1}{8 k+1}-\frac{1}{8 k+2}-\frac{1 / 2}{8 k+4}-\frac{1 / 4}{8 k+5}\right) \\
& \frac{25}{2} \log \left(\frac{781}{256}\left(\frac{57-5 \sqrt{5}}{57+5 \sqrt{5}}\right)^{\sqrt{5}}\right)=\sum_{k=0}^{\infty} \frac{1}{5^{5 k}}\left(\frac{5}{5 k+2}+\frac{1}{5 k+3}\right)
\end{aligned}
$$

Many other constants for which BBP-type formulae are known can be found in the literature. These include many rather exotic numbers and formulae with very large numbers of terms.

Note that the equalities (3) and (4) have actually been well-known for centuries. Another example are the polylogarithms $\operatorname{Li}_{s}\left(\frac{1}{b}\right)$, which have BBP-type formulae by definition:

$$
\operatorname{Li}_{s}\left(\frac{1}{b}\right)=\sum_{k=1}^{\infty} \frac{1}{b^{k} k^{s}} .
$$

## 3. Using BBP-type formulae to compute individual digits of arithmetical constants

We will use the following notations:
Let $\lfloor\cdot\rfloor$ denote the floor function, that is, for $x \in \mathbb{R},\lfloor x\rfloor$ is the greatest integer less than or equal to $x$.
For $x, y \in \mathbb{R}, y \neq 0$, let the modulo operator be defined as

$$
x \bmod y:=x-\left\lfloor\frac{x}{y}\right\rfloor y .
$$

That is, the smallest nonnegative number representing $x+\mathbb{Z} \cdot y$ that lies in a certain interval of length $y$.

The fractional part of a number $x \in \mathbb{R}$ is defined as $\operatorname{frac}(x):=x-\lfloor x\rfloor=x \bmod 1$.
Now consider the sums

$$
T_{j}=\sum_{k=0}^{\infty} \frac{1}{16^{k}(8 k+j)}
$$

as used in the derivation of the BBP-formula. Note that the hexadecimal digits of these sums (i.e. the digits in the base-16 expansion), starting at position $d+1$ after the hexadecimal point, can be obtained from the fractional part of $16^{d} T_{j}$. Now

$$
\operatorname{frac}\left(16^{d} T_{j}\right)=\sum_{k=0}^{\infty} \frac{16^{d-k}}{8 k+j} \bmod 1=\sum_{k=0}^{d} \frac{16^{d-k}}{8 k+j} \bmod 1+\sum_{k=d+1}^{\infty} \frac{16^{d-k}}{8 k+j} \bmod 1
$$

$$
\begin{aligned}
& =\sum_{k=0}^{d}\left(\frac{16^{d-k}}{8 k+j}-\left\lfloor\frac{16^{d-k}}{8 k+j}\right\rfloor\right) \bmod 1+\sum_{k=d+1}^{\infty} \frac{16^{d-k}}{8 k+j} \bmod 1 \\
& =\sum_{k=0}^{d} \frac{16^{d-k}-\left\lfloor\frac{16^{d-k}}{8 k+j}\right\rfloor(8 k+j)}{8 k+j} \bmod 1+\sum_{k=d+1}^{\infty} \frac{16^{d-k}}{8 k+j} \bmod 1 \\
& =\sum_{k=0}^{d} \frac{16^{d-k} \bmod (8 k+j)}{8 k+j} \bmod 1+\sum_{k=d+1}^{\infty} \frac{16^{d-k}}{8 k+j} \bmod 1
\end{aligned}
$$

The second summation can be evaluated using floating-point arithmetic. It has only negative exponents of 16 and is rapidly converging. Only a few terms are needed, just until it is ensured that the remaining terms add up to less than the "machine epsilon" of the floating-point arithmetic used. The first summation can be evaluated efficiently using a scheme known as the Binary Exponentiation Algorithm, reduced modulo an integer.

## The Binary Exponentiation Algorithm

To compute $x^{n}$, where $n$ is a nonnegative integer, first note that $n$ can be represented uniquely as $n=\sum_{i=0}^{N} b_{i} 2^{i}$, where $N=\left\lfloor\log _{2}(n)\right\rfloor$ and each $b_{i} \in\{0,1\}$. Then

$$
x^{n}=x^{\sum_{i=0}^{N} b_{i} 2^{i}}=\prod_{i=0}^{N} x^{b_{i} 2^{i}}
$$

in which there are at most $N$ multiplications of powers of $x$. (For $b_{i}=0$, no multiplication is necessary.) Now for all $i$,

$$
x^{2^{i}}=\left(x^{2^{i-1}}\right)^{2}=\left(\left(x^{2^{i-2}}\right)^{2}\right)^{2}=\ldots=\left(\cdots\left(\left(x^{2}\right)^{2}\right)^{2} \cdots\right)^{2}
$$

in which there are $i-1$ pairs of parentheses, corresponding to $i-1$ multiplications. So in total, the number of multiplications needed will not be greater than $N+N$, so it will never exceed $2 \log _{2}(n)$, which is of an order far less than if one would naively multiply $x$ with itself $n-1$ times.

The corresponding Binary Exponentiation Algorithm (shortly BEA) can be written as

$$
\operatorname{Pow}(x, n)\left\{\begin{aligned}
& 1, \\
& \operatorname{Pow}\left(x^{2}, \frac{n}{2}\right), \quad n \in \mathbb{N} \text { even } \\
& x \cdot \operatorname{Pow}\left(x^{2}, \frac{n-1}{2}\right), \\
& n \in \mathbb{N} \text { odd }
\end{aligned}\right.
$$

In case of the BEA modulo an integer $m$, each multiplication result is reduced modulo $m$. A simple program that would compute $r=b^{n} \bmod c$ in this manner can be represented as follows:

$$
\begin{aligned}
& t:=\max \left\{2^{i} \mid 2^{i} \leq n\right\} \\
& r:=1
\end{aligned}
$$

A if $n \geq t$ then $r:=b r \bmod c ; n:=n-t$; endif;
$t:=t / 2$;
if $t \geq 1$ then $r:=r^{2} \bmod c$; go to $\mathbf{A}$; endif;
This computation can be entirely performed with integers no greater than $c^{2}$.

The above method can be applied to compute each of the four sums $T_{j}$. Combining these using the BBP-formula, this allows us to correctly compute the $(d+1)$-th hexadecimal digit of $\pi$ without the need to know the first $d$ digits.
We can now also compute individual binary digits of $\pi$ : for every $d$, the hexadecimal digit at position $d$ corresponds to the four binary digits at positions $4 d-3$ till $4 d$.
It is clear that these methods can be easily generalized to extract individual digits in base $b$ for any constant $\alpha$ that has a BBP-type formula

$$
\alpha=\sum_{k=0}^{\infty} \frac{p(k)}{b^{c k} q(k)}
$$

where $p, q$ are polynomials with integer coefficients and $b, c \in \mathbb{N}, b \geq 2$. For example, formula (3) can be used to obtain individual binary digits of $\log 2$ and (4) to obtain individual decimal digits of $\log \left(\frac{10}{9}\right)$. Formula (2), known as Bellard's formula, is used in the currently fastest known and most widely used method for computing individual binary digits of $\pi$.

## 4. Computing digits in Mathematica

Using the methods described in the previous section, we can efficiently compute digits of various arithmetical constants with a software package such as Mathematica. We will start here with the simple BBP-type formula (4).

Recall that

$$
\log \left(\frac{10}{9}\right)=\sum_{k=1}^{\infty} \frac{1}{10^{k} k}
$$

so that the decimal digits of $\log \left(\frac{10}{9}\right)$ beginning at position $d+1$ can be obtained from the fractional part
$\operatorname{frac}\left(10^{d} \sum_{k=1}^{\infty} \frac{1}{10^{k} k}\right)=\left(10^{d} \sum_{k=1}^{\infty} \frac{1}{10^{k} k}\right) \bmod 1=(\underbrace{\sum_{k=1}^{d}\left(\frac{10^{d-k} \bmod k}{k}\right) \bmod 1}_{=: S_{d+1}}+\sum_{k=d+1}^{\infty} \frac{10^{d-k}}{k}) \bmod 1$
It is easily seen that if we multiply the above expression by 10 and then take the largest integer below that, the result will be equal to the $d+1$-th decimal digit of $\log \left(\frac{10}{9}\right)$.
Now we look at the first part of the sum, $S_{d+1}$, ignoring the tail sum starting at $d+1$. We will consider the sequence $s_{d}:=\left(\left\lfloor 10 S_{d}\right\rfloor\right)$. This sequence will be used in the computation of the decimal digits from position 2 onwards. We can define this in Mathematica in the following way:

```
In[1]:= S[d_] := FractionalPart[ Sum [ Mod [10^(d - k), k]/k, {k, 1, d}]]
In[2]:= s[d_] := IntegerPart[10 S[d - 1]]
```

We want to see if these computations, in which the tail sums are not considered yet, already produce accurate digits. To do this we generate two tables of 1000 digits, one for the digits obtained by the above method and the other for the actual decimal digits of $\log \left(\frac{10}{9}\right)$, respectively:

```
In[3]:= stable := Table[s[d], {d, 2, 1001}]
In[4]:= srealdigits := RealDigits[Log[10/9], 10, 1001][[1]]
In[5]:= stableoftruedigits := Table[srealdigits[[d]], {d, 2, 1001}]
```

We compare these tables and count the number of matching digits:

```
In[6]:= sdifftable := stableoftruedigits - stable
In[7]:= Count[sdifftable, 0]
Out[7]= 991
```

We see that there are only 9 differences between the digits produced by $s_{d}$ and the correct digits of $\log \left(\frac{10}{9}\right)$.

In the same way, we can look at the polylogarithm $\operatorname{Li}_{2}\left(\frac{1}{10}\right)$, which is defined as

$$
\operatorname{Li}_{2}\left(\frac{1}{10}\right)=\sum_{k=1}^{\infty} \frac{1}{10^{k} k^{2}}
$$

Notice that this formula is very similar to the one for $\log \left(\frac{10}{9}\right)$, the only difference being $k^{2}$ in the denominator instead of $k$. Its decimal digits beginning at position $d+1$ can be obtained from the fractional part

$$
\left(10^{d} \sum_{k=1}^{\infty} \frac{1}{10^{k} k^{2}}\right) \bmod 1=(\underbrace{\sum_{k=1}^{d}\left(\frac{10^{d-k} \bmod k^{2}}{k^{2}}\right) \bmod 1}_{=: T_{d+1}}+\sum_{k=d+1}^{\infty} \frac{10^{d-k}}{k^{2}}) \bmod 1
$$

Multiplying this expression by 10 and taking the largest integer below the result, we get the $d+1$-th decimal digit of $\operatorname{Li}_{2}\left(\frac{1}{10}\right)$.
We ignore the tail sum again and only look at the first part of the sum, $T_{d+1}$. We now use the sequence $t_{d}:=\left(\left\lfloor 10 T_{d}\right\rfloor\right)$ in the computation of the decimal digits from position 2 onwards. In Mathematica:

```
In[8]:= T[d_] := FractionalPart[ Sum [ Mod [10^(d - k), k^2]/k^2, {k, 1, d}]]
In[9]:= t[d_] := IntegerPart[10 T[d - 1]]
```

Again we want to check if this already produces accurate digits. We generate two tables again, one for the digits computed this way and the other for the actual decimal digits of $\operatorname{Li}_{2}\left(\frac{1}{10}\right)$ :

```
In[10]:= ttable := Table[t[d], {d, 2, 1001}]
In[11]:= trealdigits := RealDigits[PolyLog[2, 1/10], 10, 1001][[1]]
In[12]:= ttableoftruedigits := Table[trealdigits[[d]], {d, 2, 1001}]
```

We compare the number of matching digits in these tables again:

```
In[13]:= tdifftable := ttableoftruedigits - ttable
In[14]:= Count[tdifftable, 0]
Out[14]= 999
```

Actually, all 1000 digits match here; the one discrepancy is due to rounding. (The actual 1001-st and 1002-nd digits are 2 and 9 , respectively. This causes the function RealDigits to return a 3 as the 1001-st and last digit in this computation.)

In a similar way, we can use the BBP-formula

$$
\pi=\sum_{k=0}^{\infty} \frac{1}{16^{k}}\left(\frac{4}{8 k+1}-\frac{2}{8 k+4}-\frac{1}{8 k+5}-\frac{1}{8 k+6}\right)
$$

to perform a computation of hexadecimal digits of $\pi$. For computational convenience, we gather all four terms under the same denominator and then shift the summation index by 1 . We get

$$
\pi=\sum_{k=1}^{\infty} \frac{1}{16^{k-1}} \frac{p(k)}{q(k)} \quad \text { where } \quad \begin{aligned}
p(k) & =120 k^{2}-89 k+16 \\
q(k) & =512 k^{4}-1024 k^{3}+712 k^{2}-206 k+21
\end{aligned}
$$

The hexadecimal digits starting at the $d+1$-th position can be obtained from

$$
\left(16^{d+1} \sum_{k=1}^{\infty} \frac{p(k)}{16^{k} q(k)}\right) \bmod 1=(\underbrace{\sum_{k=1}^{d}\left(\frac{16^{d+1-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1}_{=: U_{d+1}}+\sum_{k=d+1}^{\infty} \frac{16^{d+1-k} p(k)}{q(k)}) \bmod 1
$$

Again, we only look at the first part of this sum, $U_{d+1}$, and define the sequence $u_{d}:=\left(\left\lfloor 16 U_{d}\right\rfloor\right)$.

$$
\begin{aligned}
& \operatorname{In}[15]:=p\left[k_{-}\right]:=120 k^{\wedge} 2-89 k+16 \\
& \operatorname{In}[16]:=q\left[k_{-}\right]:=512 k^{\wedge} 4-1024 k^{\wedge} 3+712 k^{\wedge} 2-206 k+21
\end{aligned}
$$

```
In[17]:= U[d_] := FractionalPart[ Sum[ Mod[ 16^(d - k + 1)p[k], q[k]]/q[k],
    {k, 1, d}]]
In[18]:= u[d_] := IntegerPart[16 U[d - 1]]
```

To get an indication of the accuracy of the digits produced this way, we generate two tables of 1000 digits again, one for the computed digits and one for the actual hexadecimal digits of $\pi$.

```
In[19]:= utable := Table[u[d], {d, 3, 1002}]
In[20]:= urealdigits := RealDigits[Pi, 16, 1002, -2][[1]]
In[21]:= utableoftruedigits := Table[urealdigits[[d]], {d, 2, 1001}]
```

Comparing these tables, we get

```
In[22]:= udifftable := utableoftruedigits - utable
In[23]:= Count[udifftable, 0]
Out[23]= 1000
```

Again we observe no differences between our computed digits and the true hexadecimal digits of $\pi$. In fact, in various researches, millions of hexadecimal digits have been computed in this manner, none of these showing any differences.

## 5. Generating correct base- $b$ expansions for BBP-numbers

Lemma 5.1. Let $\left(x_{n}\right)$ be the sequence defined by

$$
\begin{equation*}
x_{0}=0 \quad \text { and } \quad x_{n}=\left(b x_{n-1}+\frac{p(n)}{q(n)}\right) \bmod 1, \quad n=1,2, \ldots \tag{5}
\end{equation*}
$$

Then for all $n=1,2, \ldots$ the following holds:

$$
x_{n}=\left(\sum_{k=1}^{n} \frac{b^{n-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1 .
$$

Proof. We will prove this equality by induction. For $n=1$ we have

$$
\begin{gathered}
x_{1}=\left(b x_{0}+\frac{p(1)}{q(1)}\right) \bmod 1=\frac{p(1)}{q(1)} \bmod 1=\frac{b^{1-1} p(1) \bmod q(1)}{q(1)} \bmod 1 \\
=\left(\sum_{k=1}^{1} \frac{b^{1-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1
\end{gathered}
$$

so the equality holds for $n=1$.
Now suppose that for some $n \geq 1$,

$$
x_{n}=\left(b x_{n-1}+\frac{p(n)}{q(n)}\right) \bmod 1=\sum_{k=1}^{n}\left(\frac{b^{n-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1 .
$$

Then

$$
\begin{gathered}
\left(b x_{n}+\frac{p(n+1)}{q(n+1)}\right) \bmod 1=\left(b\left(b x_{n-1}+\frac{p(n)}{q(n)}\right) \bmod 1+\frac{p(n+1)}{q(n+1)}\right) \bmod 1 \\
=\left(b\left(\sum_{k=1}^{n}\left(\frac{b^{n-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1\right) \bmod 1+\frac{p(n+1) \bmod q(n+1)}{q(n+1)} \bmod 1\right) \bmod 1 \\
=\left(b\left(\sum_{k=1}^{n}\left(\frac{b^{n-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1\right) \bmod 1+b \cdot \frac{b^{n-(n+1)} p(n+1) \bmod q(n+1)}{q(n+1)} \bmod 1\right) \bmod 1 \\
=\left(b \sum_{k=1}^{n+1}\left(\frac{b^{n-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1\right) \bmod 1=\sum_{k=1}^{n+1}\left(\frac{b^{n+1-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1
\end{gathered}
$$

which means that the equality also holds for $n+1$.

For a real number $\alpha \in[0,1)$, we can define the norm $\|\alpha\|=\min (\alpha, 1-\alpha)$. In this way, $\|\alpha-\beta\|$ measures the shortest distance between $\alpha$ and $\beta$ on a circle with circumference 1 .

Lemma 5.2. Let $\alpha$ be a real number with BBP-type formula

$$
\alpha=\sum_{k=1}^{\infty} \frac{p(k)}{b^{k} q(k)}
$$

where $p, q$ are integer polynomials with $0 \leq \operatorname{deg}(p)<\operatorname{deg}(q)$ such that $q$ has no zeros for the positive integers. Define the sequence $\left(x_{n}\right)$ as in (5) and let $\alpha_{n}$ denote the part of the base-b expansion of $\alpha$ starting at position $n+1$ after the base-b "decimal" point. Then

$$
\left\|x_{n}-\alpha_{n}\right\| \longrightarrow 0 \text { for } n \longrightarrow \infty
$$

Proof. For all $n \in \mathbb{N}$ we have that

$$
\begin{aligned}
\alpha_{n} & =\left(b^{n} \alpha\right) \bmod 1=\left(\sum_{k=1}^{\infty} \frac{b^{n-k} p(k)}{q(k)}\right) \bmod 1 \\
& =\left(\left(\sum_{k=1}^{n} \frac{b^{n-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1+\sum_{k=n+1}^{\infty} \frac{b^{n-k} p(k)}{q(k)}\right) \bmod 1 \\
& =\left(x_{n}+\sum_{k=n+1}^{\infty} \frac{b^{n-k} p(k)}{q(k)}\right) \bmod 1 .
\end{aligned}
$$

Furthermore, since $0 \leq \operatorname{deg}(p)<\operatorname{deg}(q)$, we have $\left|\frac{p(k)}{q(k)}\right| \searrow 0$, that is, for all $\varepsilon>0$, there is an $N=N(\varepsilon)$ such that $\left|\frac{p(k)}{q(k)}\right|<\varepsilon$ for all $k \geq N$. Hence, if $n$ is large enough,

$$
\left\|x_{n}-\alpha_{n}\right\| \leq\left|\sum_{k=n+1}^{\infty} \frac{b^{n-k} p(k)}{q(k)}\right|<\varepsilon \sum_{k=n+1}^{\infty} b^{n-k}=\varepsilon \sum_{j=1}^{\infty} \frac{1}{b^{j}}=\varepsilon \frac{1}{b-1} \leq \varepsilon .
$$

This means that $\left\|x_{n}-\alpha_{n}\right\| \longrightarrow 0$ for $n \longrightarrow \infty$.

Lemma 5.2 shows us that the sequence $\left(x_{n}\right)$ is a good approximation of the sequence $\left(\alpha_{n}\right)$ of shifted digit expansions of $\alpha$.
In the previous section we saw that for some BBP-numbers

$$
\alpha=\sum_{k=1}^{\infty} \frac{1}{b^{k}} \frac{p(k)}{q(k)}
$$

the sequence $\left(y_{n}\right):=\left(\left\lfloor b x_{n}\right\rfloor\right)$, where $x_{n}$ is the fractional part of a finite summation,

$$
x_{n}=\left(\sum_{k=1}^{n} \frac{b^{n-k} p(k) \bmod q(k)}{q(k)}\right) \bmod 1,
$$

produces digits of the base- $b$ expansion of $\alpha$ quite accurately, regardless of the tail summation; for some of these BBP-numbers, the produced digits even appear to exactly match the true digits.
Note that the sequence $\left(x_{n}\right)$ is precisely the one used in Lemmas 5.1 and 5.2. Hence, another way of stating this is that the sequence $\left(y_{n}\right)=\left(\left\lfloor b x_{n}\right\rfloor\right)$, where $\left(x_{n}\right)$ is defined as

$$
x_{0}=0 \quad \text { and } \quad x_{n}=\left(b x_{n-1}+\frac{p(n)}{q(n)}\right) \bmod 1, \quad n=1,2, \ldots
$$

very accurately (and in some cases exactly) appears to produce digits of the base-b expansion of $\alpha$, regardless of the tail sequence

$$
t_{n}=\sum_{k=n+1}^{\infty} \frac{b^{n-k} p(k)}{q(k)} .
$$

For instance, in the BBP-type formula (3) for $\alpha=\log 2$, we have $p(k) \equiv 1$ and $q(k)=k$ and this corresponds to the sequence $\left(x_{n}\right)$ where

$$
x_{0}=0 \quad \text { and } \quad x_{n}=\left(2 x_{n-1}+\frac{1}{n}\right) \bmod 1 .
$$

The digit sequence $\left(y_{n}\right)=\left(\left\lfloor 2 x_{n}\right\rfloor\right)$ generates the binary digits of $\log 2$ quite accurately. In the same way, in the formula (4) for $\alpha=\log \left(\frac{10}{9}\right)$, we also have $p(k) \equiv 1$ and $q(k)=k$, corresponding to the sequence $\left(x_{n}\right)$ with

$$
x_{0}=0 \quad \text { and } \quad x_{n}=\left(10 x_{n-1}+\frac{1}{n}\right) \bmod 1 .
$$

Here the digit sequence $\left(y_{n}\right)=\left(\left\lfloor 10 x_{n}\right\rfloor\right)$ quite accurately produces the decimal digits of $\log \left(\frac{10}{9}\right)$.
For $\alpha=\operatorname{Li}_{2}\left(\frac{1}{10}\right)$, we have $p(k) \equiv 1$ and $q(k)=k^{2}$, with its corresponding sequence $\left(x_{n}\right)$, where

$$
x_{0}=0 \quad \text { and } \quad x_{n}=\left(10 x_{n-1}+\frac{1}{n^{2}}\right) \bmod 1,
$$

the digit sequence $\left(y_{n}\right)=\left(\left\lfloor 10 x_{n}\right\rfloor\right)$ appears to correctly produce the decimal digits of $\operatorname{Li}_{2}\left(\frac{1}{10}\right)$.
In the same way, the BBP-formula gives us $p(k)=120 k^{2}-89 k+16$ and $q(k)=512 k^{4}-$ $1024 k^{3}+712 k^{2}-206 k+21$ and the sequence

$$
x_{0}=0 \quad \text { and } \quad x_{n}=\left(16 x_{n-1}+\frac{p(k)}{q(k)}\right) \bmod 1
$$

Here, the digit sequence $\left(y_{n}\right)=\left(\left\lfloor 16 x_{n}\right\rfloor\right)$ appears to correctly produce hexadecimal digits of $\pi$.
We quickly observe that for $\log 2$ and $\log \left(\frac{10}{9}\right)$, we have $\operatorname{deg}(q)=\operatorname{deg}(p)+1$ and for $\operatorname{Li}_{2}\left(\frac{1}{10}\right)$ and $\pi$, we have $\operatorname{deg}(q)=\operatorname{deg}(p)+2$. In fact, computations for various other BBP-numbers all seem to suggest that whenever $\operatorname{deg}(q) \geq \operatorname{deg}(p)+2$, the correct digits are generated, whereas errors are expected to occur if this is not the case. This leads to the following conjecture:

Conjecture 5.3. If a real number $\alpha$ can be written as a BBP-type formula

$$
\alpha=\sum_{k=1}^{\infty} \frac{1}{b^{k}} \frac{p(k)}{q(k)},
$$

with $\operatorname{deg}(q) \geq \operatorname{deg}(p)+2$, the sequence $\left(\left\lfloor b x_{n}\right\rfloor\right)$, where $\left(x_{n}\right)$ is defined as

$$
x_{0}=0 \quad \text { and } \quad x_{n}=\left(b x_{n-1}+\frac{p(n)}{q(n)}\right) \bmod 1, \quad n=1,2, \ldots
$$

generates the correct base-b expansion of $\alpha$.

Note that when $\operatorname{deg}(q) \geq \operatorname{deg}(p)+2$, the expression $\frac{p(k)}{q(k)}$ is summable. If the conjecture is true, the tail sequence

$$
t_{n}=\sum_{k=n+1}^{\infty} \frac{b^{n-k} p(k)}{q(k)}
$$

which can be seen as the error term for the prediction of the digit at place $n+1$, never changes a digit. For $\operatorname{Li}_{2}\left(\frac{1}{10}\right)$, the tail sequence is given by

$$
t_{n}=\sum_{k=n+1}^{\infty} \frac{10^{n-k}}{k^{2}} .
$$

This is approximately equal to the first summand $(k=n+1)$. For the sum of all tail sequences (i.e. the sum of all prediction errors) we have

$$
\sum_{n=1}^{\infty} t_{n} \approx 0.0644934 \ldots
$$

Multiplied by 10, this number can be regarded as the expected total number of errors in the generated digit sequence; it is a number smaller than 1 . Similarly, for $\pi$, we have the tail sequence

$$
t_{n}=\sum_{k=n+1}^{\infty} \frac{16^{n-k}\left(120 k^{2}-89 k+16\right)}{512 k^{4}-1024 k^{3}+712 k^{2}-206 k+21}
$$

and

$$
\sum_{n=1}^{\infty} t_{n} \approx 0.0157946 \ldots
$$

Multiplied by 16 , this number gives us an indication of the expected total number of errors in the generated digit sequence for $\pi$; again it is smaller than 1 , meaning no errors are expected to occur. Obviously, the corresponding tail sequences for $\log 2$ and $\log \left(\frac{10}{9}\right)$, where $\frac{p(k)}{q(k)}=\frac{1}{k}$, are not summable. This indicates that errors are expected to occur infinitely often. However, it is difficult to be precise here; no proof of Conjecture 5.3 is known, nor of a concrete relation between the sums of the tail sequences and the number of errors that occur in our digit computations.

## 6. Normality of BBP-numbers

We now return to the concept of normality. If Conjecture 5.3 is true, we have a way to correctly represent individual digits of a certain class of BBP-numbers by means of the sequence $\left(\left\lfloor b x_{n}\right\rfloor\right)$. Even under the weaker (but proven) statement Lemma 5.2, we know that for all BBP-numbers, the sequence at least eventually agrees quite well with the true base- $b$ expansions. This leads to the hope that statements can be made about the distribution of the digits of BBP-numbers, or even provide a way to establish normality for these numbers.

Definition 6.1. A sequence $\left(x_{n}\right)$ in $[0,1)$ is called equidistributed if for all $0 \leq c<d<1$,

$$
\lim _{N \rightarrow \infty} \frac{\#\left\{x_{i} \in[c, d) \mid i<N\right\}}{N}=d-c
$$

Definition 6.2. A sequence $\left(x_{n}\right)$ in $[0,1)$ has a finite attractor $W=\left(w_{0}, w_{1}, \ldots w_{P-1}\right)$ if for all $\varepsilon>0$ there exists a $K=K(\varepsilon)$ such that $\left\|x_{K+k}-w_{t(k)}\right\|<\varepsilon$ for all $k \geq 0$, where $t$ is some function $t: \mathbb{N} \rightarrow\{0, \ldots P-1\}$.

This is used in the following important hypothesis formulated by [Bailey \& Crandall, 2001]:

Hypothesis 6.3. Let $p, q$ be polynomials with integer coefficients, such that $0 \leq \operatorname{deg}(p)<\operatorname{deg}(q)$, with $q$ having no zeros on $\mathbb{N}$, and let $r_{n}=\frac{p(n)}{q(n)}$. Then the sequence $\left(x_{n}\right)$, defined by

$$
x_{0}=0 \quad \text { and } \quad x_{n}=\left(b x_{n-1}+r_{n}\right) \bmod 1, \quad n=1,2, \ldots
$$

is either equidistributed in $[0,1)$ or has a finite attractor.

We now state some important results that will be used to show the consequence of Hypothesis
6.3 for the normality of BBP-numbers.

Theorem 6.4. If $\left(x_{n}\right)$ is an equidistributed sequence and $\left(y_{n}\right)$ a sequence for which $\lim _{n \rightarrow \infty}\left(y_{n} \bmod 1\right)=L$, then the sequence $\left(\left(x_{n}+y_{n}\right) \bmod 1\right)$ is also equidistributed.

Theorem 6.5. If $\left(y_{n}\right)$ is a sequence in $[0,1)$ that has a limit point $\lim _{n \rightarrow \infty} y_{n}=L$, then a sequence $\left(x_{n}\right)$ in $[0,1)$ has a finite attractor if and only if the sequence $\left(\left(x_{n}+y_{n}\right) \bmod 1\right)$ has a finite attractor.

Theorem 6.6. A number $\alpha \in \mathbb{R}$ is normal to base-b if and only if the sequence $\left(\left(b^{n} \alpha\right) \bmod 1\right)_{n \in \mathbb{N}}$ is equidistributed.

Theorem 6.7. A number $\alpha \in \mathbb{R}$ is rational if and only if the sequence $\left(\left(b^{n} \alpha\right) \bmod 1\right)_{n \in \mathbb{N}}$ has a finite attractor.

It would go too far to prove these results here; full proofs have been given in [Kuipers \& Niederreiter, 1974] and [Bailey \& Crandall, 2001], respectively.

Theorem 6.8. Let $\alpha$ be a $B B P$-number of the form

$$
\begin{equation*}
\alpha=\sum_{k=1}^{\infty} \frac{p(k)}{b^{k} q(k)} \tag{6}
\end{equation*}
$$

where $p, q$ are polynomials with integer coefficients, $0 \leq \operatorname{deg}(p)<\operatorname{deg}(q), q$ having no zeros for positive integers, and $b \geq 2$ an integer. Then $\alpha$ is rational if and only if the sequence defined by

$$
\begin{equation*}
x_{0}=0 \quad \text { and } \quad x_{n}=\left(b x_{n-1}+\frac{p(n)}{q(n)}\right) \bmod 1, \quad n=1,2, \ldots \tag{7}
\end{equation*}
$$

has a finite attractor.
Proof. From Lemma 5.2, we know that the sequence $\left(x_{n}-\alpha_{n}\right)$, where $\alpha_{n}=\left(b^{n} \alpha\right) \bmod 1$, has a limit (which is 0 ). Because of Theorem 6.5, this means that $\left(x_{n}\right)$ has a finite attractor if and only if $\left(\alpha_{n}\right)$ has a finite attractor, and from Theorem 6.7, this is the case if and only if $\alpha$ is rational.

Theorem 6.9. Let $\alpha$ and $\left(x_{n}\right)$ be defined as in Theorem 6.8. If $\left(x_{n}\right)$ is equidistributed, then $\alpha$ is normal to base $b$.

Proof. Suppose that $\left(x_{n}\right)$ is equidistributed. Let $\left(\alpha_{n}\right)$ be as in the proof of the previous theorem. We know that $\left(x_{n}-\alpha_{n}\right)$ converges to 0 , so from Theorem 6.4, ( $\alpha_{n}$ ) must also be equidistributed. Because of Theorem 6.6, $\alpha$ must be normal to base $b$.

The previous two theorems lead to an interesting result about the normality of BBP-numbers, assuming Hypothesis 6.3 is true.

Theorem 6.10. Under Hypothesis 6.3, a BBP-number is either rational or normal to base $b$.
Proof. Every BBP-number $\alpha$ of the form (6) has a corresponding sequence $\left(x_{n}\right)$ of the form
(7). Under the hypothesis, this sequence either has a finite attractor or is equidistributed. If it has a finite attractor, then by Theorem $6.8 \alpha$ is rational. If it is equidistributed, then by Theorem $6.9 \alpha$ is normal to base $b$.

Obviously, this means that if Hypothesis 6.3 is valid, every irrational number that has a BBPtype formula as in (6) is normal to base $b$. For example, this would imply that $\log 2$ is normal to base $2, \log \left(\frac{10}{9}\right)$ and $\operatorname{Li}_{2}\left(\frac{1}{10}\right)$ are normal to base 10 and $\pi$ is normal to base 16 (and hence also to base 2).

## 7. Rational BBP-numbers

Of course, the question arises whether rational examples of BBP-numbers actually exist. The answer is yes; several BBP-type formulae are known that evaluate to the rational number 0 . We list some examples here, found by David Bailey using the PSLQ-algorithm. [Bailey, 2004]

$$
\begin{aligned}
& 0=\sum_{k=0}^{\infty} \frac{1}{16^{k}}\left(\frac{-8}{8 k+1}+\frac{8}{8 k+2}+\frac{4}{8 k+3}+\frac{8}{8 k+4}+\frac{2}{8 k+5}+\frac{2}{8 k+6}-\frac{1}{8 k+7}\right) \\
& 0=\sum_{k=0}^{\infty} \frac{1}{64^{k}}\left(\frac{16}{6 k+1}-\frac{24}{6 k+2}-\frac{8}{6 k+3}-\frac{6}{6 k+4}+\frac{1}{6 k+5}\right) \\
& 0=\sum_{k=0}^{\infty} \frac{1}{729^{k}}\left(\frac{81}{12 k+2}-\frac{162}{12 k+3}+\frac{27}{12 k+5}+\frac{36}{12 k+6}\right. \\
&\left.\quad+\frac{9}{12 k+8}+\frac{6}{12 k+9}+\frac{4}{12 k+10}-\frac{1}{12 k+11}\right)
\end{aligned}
$$

These BBP zero relations are relevant for finding new BBP-type formulae with algorithms like PSLQ. Of course, these formulae can be rewritten in the familiar form (6) by gathering all terms under the same denominator (and shifting the index of summation). For instance, the second formula above yields

$$
0=\sum_{k=1}^{\infty} \frac{1}{64^{k}} \frac{6804 k^{4}-14256 k^{3}+9891 k^{2}-2664 k+224}{\left(-1944 k^{5}+4860 k^{4}-4590 k^{3}+2025 k^{2}-411 k+30\right)}
$$

By Theorem 6.8 this means that the sequence $\left(x_{n}\right)$ given by $x_{0}=0$ and

$$
x_{n+1}=\left(64 x_{n}+\frac{6804 n^{4}-14256 n^{3}+9891 n^{2}-2664 n+224}{-1944 n^{5}+4860 n^{4}-4590 n^{3}+2025 n^{2}-411 n+30}\right) \bmod 1
$$

has a finite attractor. Since every digit in the base-64 expansion is 0 , this finite attractor is of course equal to zero.

However, in an attempt to further investigate Hypothesis 6.3, we would like to find some nonzero rational BBP-numbers. We observe that BBP-numbers can be considered as special cases of hypergeometric functions:

Definition 7.1. The standard hypergeometric function ${ }_{2} F_{1}$ is given by

$$
{ }_{2} F_{1}(a, b ; c ; z)=\sum_{n=0}^{\infty} \frac{(a)_{n}(b)_{n}}{(c)_{n}} \frac{z^{n}}{n!}
$$

where $(a)_{n}$ denotes the Pochhammer symbol, $(a)_{n}=a(a+1) \cdots(a+n-1)$.
For instance,

$$
\begin{aligned}
\frac{1}{2}{ }_{2} F_{1}\left(1,1 ; 2 ; \frac{1}{2}\right) & =\frac{1}{2} \sum_{n=0}^{\infty} \frac{(1)_{n}(1)_{n}}{(2)_{n} n!}\left(\frac{1}{2}\right)^{n}=\frac{1}{2} \sum_{n=0}^{\infty} \frac{(1 \cdot 2 \cdot 3 \cdots n)(1 \cdot 2 \cdot 3 \cdots n)}{(2 \cdot 3 \cdots n \cdot(n+1))(1 \cdot 2 \cdot 3 \cdots n)}\left(\frac{1}{2}\right)^{n} \\
& =\frac{1}{2} \sum_{n=0}^{\infty} \frac{1}{n+1} \frac{1}{2^{n}}=\sum_{m=1}^{\infty} \frac{1}{m 2^{m}}=\log 2 .
\end{aligned}
$$

Another relevant example is, for $j=1, \ldots 8$,

$$
\begin{aligned}
\frac{1}{j}{ }_{2} F_{1}\left(\frac{j}{8}, 1 ; \frac{j}{8}+1 ; \frac{1}{16}\right) & =\frac{1}{j} \sum_{n=0}^{\infty} \frac{\left(\frac{j}{8}\right)_{n}(1)_{n}}{\left(\frac{j}{8}+1\right)_{n} n!}\left(\frac{1}{16}\right)^{n} \\
& =\frac{1}{j} \sum_{n=0}^{\infty} \frac{\frac{j}{8}\left(\frac{j}{8}+1\right)\left(\frac{j}{8}+2\right) \cdots\left(\frac{j}{8}+n-1\right)(1 \cdot 2 \cdots n)}{\left(\frac{j}{8}+1\right)\left(\frac{j}{8}+2\right) \cdots\left(\frac{j}{8}+n-1\right)\left(\frac{j}{8}+n\right)(1 \cdot 2 \cdots n)}\left(\frac{1}{16}\right)^{n} \\
& =\frac{1}{j} \sum_{n=0}^{\infty} \frac{\frac{j}{8}}{\frac{j}{8}+n} \frac{1}{16^{n}}=\sum_{n=0}^{\infty} \frac{1}{8 n+j} \frac{1}{16^{n}}=: T_{j}
\end{aligned}
$$

in which we recognize the series used in the original BBP formula $\pi=4 T_{1}-2 T_{4}-T_{5}-T_{6}$.
This makes us hopeful that possible nonzero rational examples of hypergeometric function evaluations will lead to nonzero rational examples of BBP-numbers. Now for some particular points, the hypergeometric function is known to assume rational values. For example,

$$
{ }_{2} F_{1}\left(\frac{1}{3}, \frac{2}{3} ; \frac{5}{6} ; \frac{27}{32}\right)=\frac{8}{5} \quad \text { and } \quad{ }_{2} F_{1}\left(\frac{1}{4}, \frac{1}{2} ; \frac{3}{4} ; \frac{80}{81}\right)=\frac{9}{5}
$$

corresponding to

$$
\sum_{n=0}^{\infty} \frac{\left(\frac{1}{3}\right)_{n}\left(\frac{2}{3}\right)_{n}}{\left(\frac{5}{6}\right)_{n} n!}\left(\frac{27}{32}\right)^{n}=\frac{8}{5} \quad \text { and } \quad \sum_{n=0}^{\infty} \frac{\left(\frac{1}{4}\right)_{n}\left(\frac{1}{2}\right)_{n}}{\left(\frac{3}{4}\right)_{n} n!}\left(\frac{80}{81}\right)^{n}=\frac{9}{5}
$$

and even

$$
{ }_{2} F_{1}\left(\frac{1}{3}, \frac{2}{3} ; \frac{3}{2} ; \frac{27}{4} x^{2}\left(1-x^{2}\right)^{2}\right)=\frac{1}{1-x^{2}} \quad \text { for } 0 \leq x \leq \frac{1}{3} \sqrt{3} .
$$

However, these examples do not correspond to BBP-type formulae; not enough terms in the numerator and denominator of the summands cancel.

An interesting theorem along these lines, dealing with the (ir)rationality of a class of numbers containing the BBP-numbers, is the following [Lagarias, 2001]:

Theorem 7.2. Let $p, q \in \mathbb{Q}[X]$ such that $q$ has no zeros for the nonnegative integers and $p$ and $q$ have no common factors. Let

$$
f(z)=\sum_{n=0}^{\infty} \frac{p(n)}{q(n)} z^{n} .
$$

If $q(x)$ splits into distinct linear factors over $\mathbb{Q}$, then for each $r \in \mathbb{Q}$ that is in the open disk of convergence of $f(z)$ around $z=0$, the value $f(r)$ is either rational or transcendental.

The proof of this theorem relies on results about irrationality and transcendence for linear combinations of logarithms. We will illustrate this for the case of $\pi$.

Take a look at the function

$$
f_{1}(x)=\sum_{n=0}^{\infty} \frac{x^{n}}{8 n+1} .
$$

Setting $x=y^{8}$, we get

$$
f_{1}(x)=\sum_{n=0}^{\infty} \frac{y^{8 n}}{8 n+1}=\frac{1}{y} \sum_{n=0}^{\infty} \frac{y^{8 n+1}}{8 n+1}=\frac{1}{y} g_{1}(y) \quad \text { where } \quad g_{1}(y):=\sum_{n=0}^{\infty} \frac{y^{8 n+1}}{8 n+1} .
$$

Then, assuming $|y|<1$,

$$
\begin{equation*}
g_{1}^{\prime}(y)=\sum_{n=0}^{\infty} y^{8 n}=\frac{1}{1-y^{8}} \quad \text { and } \quad g_{1}(0)=0 \tag{8}
\end{equation*}
$$

We know that $1-y^{8}=0$ if and only if $y \in\left\{\zeta_{8}^{k} \mid k=1, \ldots, 8\right\}$, where the $\zeta_{8}^{k}=e^{2 \pi i k / 8}$ are the 8th complex roots of unity:

$$
1-y^{8}=\prod_{k=1}^{8}\left(1-\zeta_{8}^{k} y\right)
$$

Now we want to find constants $a_{k}$ such that

$$
\frac{1}{1-y^{8}}=\sum_{k=1}^{8} \frac{a_{k}}{1-\zeta_{8}^{k} y}
$$

from which it follows with (8) that

$$
g_{1}(y)=\sum_{k=1}^{8}-a_{k} \zeta_{8}^{-k} \log \left(1-\zeta_{8}^{k} y\right)
$$

We know that for each $k=1, \ldots, 8$

$$
\frac{a_{k}}{1-\zeta_{8}^{k} y}=\frac{a_{k} \prod_{m=1, m \neq k}^{8}\left(1-\zeta_{8}^{m} y\right)}{1-y^{8}}
$$

Summing over all $k$, we get

$$
\frac{1}{1-y^{8}}=\sum_{k=1}^{8} \frac{a_{k}}{1-\zeta_{8}^{k} y}=\frac{1}{1-y^{8}} \sum_{k=1}^{8}\left(a_{k} \prod_{m=1, m \neq k}^{8}\left(1-\zeta_{8}^{m} y\right)\right)
$$

and from this we are able to calculate $a_{k}, k=1, \ldots, 8$. This rather lengthy calculation, which is most easily done in Mathematica, is equivalent to solving the system of linear equations

$$
\begin{aligned}
& \mathbf{Z} \mathbf{a}=(1,0,0,0,0,0,0,0)^{T}, \quad \text { in which } \mathbf{c}=\left(a_{1}, \ldots, a_{8}\right)^{T} \text { and } \\
& \mathbf{Z}=\left(\begin{array}{ccccccc}
\zeta_{8}^{8} & \zeta_{8}^{8} & \zeta_{8}^{8} & \zeta_{8}^{8} & \zeta_{8}^{8} & \zeta_{8}^{8} & \zeta_{8}^{8} \\
\zeta_{8}^{8} \\
\zeta_{8}^{1} & \zeta_{8}^{2} & \zeta_{8}^{3} & \zeta_{8}^{4} & \zeta_{8}^{5} & \zeta_{8}^{6} & \zeta_{8}^{7} \\
\zeta_{8}^{8} \\
\zeta_{8}^{2} & \zeta_{8}^{4} & \zeta_{8}^{6} & \zeta_{8}^{8} & \zeta_{8}^{2} & \zeta_{8}^{4} & \zeta_{8}^{6} \\
\zeta_{8}^{8} \\
\zeta_{8}^{8} & \zeta_{8}^{6} & \zeta_{8}^{1} & \zeta_{8}^{4} & \zeta_{8}^{7} & \zeta_{8}^{2} & \zeta_{8}^{5} \\
\zeta_{8}^{8} \\
\zeta_{8}^{4} & \zeta_{8}^{8} & \zeta_{8}^{4} & \zeta_{8}^{8} & \zeta_{8}^{4} & \zeta_{8}^{8} & \zeta_{8}^{4} \\
\zeta_{8}^{8} & \zeta_{8}^{2} & \zeta_{8}^{7} & \zeta_{8}^{4} & \zeta_{8}^{1} & \zeta_{8}^{6} & \zeta_{8}^{3} \\
\zeta_{8}^{8} \\
\zeta_{8}^{6} & \zeta_{8}^{4} & \zeta_{8}^{2} & \zeta_{8}^{8} & \zeta_{8}^{6} & \zeta_{8}^{4} & \zeta_{8}^{2} \\
\zeta_{8}^{8} \\
\zeta_{8}^{7} & \zeta_{8}^{6} & \zeta_{8}^{5} & \zeta_{8}^{4} & \zeta_{8}^{3} & \zeta_{8}^{2} & \zeta_{8}^{1} \\
\zeta_{8}^{8}
\end{array}\right)
\end{aligned}
$$

Its solution is $\mathbf{a}=\frac{1}{8}(1,1,1,1,1,1,1,1)^{T}$. This means that

$$
g_{1}(y)=-\frac{1}{8} \sum_{k=1}^{8} \zeta_{8}^{-k} \log \left(1-\zeta_{8}^{k} y\right)
$$

In the same way, we look at the function

$$
f_{4}(x)=\sum_{n=0}^{\infty} \frac{x^{n}}{8 n+4} .
$$

Again setting $x=y^{8}$, we get

$$
f_{4}(x)=\sum_{n=0}^{\infty} \frac{y^{8 n}}{8 n+4}=\frac{1}{y^{4}} \sum_{n=0}^{\infty} \frac{y^{8 n+4}}{8 n+4}=\frac{1}{y^{4}} g_{4}(y) \quad \text { where } \quad g_{4}(y):=\sum_{n=0}^{\infty} \frac{y^{8 n+4}}{8 n+4} .
$$

Assuming $|y|<1$, we have

$$
g_{4}^{\prime}(y)=\sum_{n=0}^{\infty} y^{8 n+3}=\frac{y^{3}}{1-y^{8}} \quad \text { and } \quad g_{4}(0)=0
$$

Solving a similar system of linear equations, we get

$$
g_{4}(y)=-\frac{1}{8} \sum_{k=1}^{8}(-1)^{k} \log \left(1-\zeta_{8}^{k} y\right) .
$$

Likewise (omitting details for brevity),

$$
f_{5}(x)=\sum_{n=0}^{\infty} \frac{x^{n}}{8 n+5}=\frac{1}{y^{5}} g_{5}(y) \quad \text { where } \quad g_{5}(y):=\sum_{n=0}^{\infty} \frac{y^{8 n+5}}{8 n+5}
$$

yields

$$
g_{5}(y)=-\frac{1}{8} \sum_{k=1}^{8}(-1)^{k} \zeta_{8}^{-k} \log \left(1-\zeta_{8}^{k} y\right)
$$

and

$$
f_{6}(x)=\sum_{n=0}^{\infty} \frac{x^{n}}{8 n+6}=\frac{1}{y^{6}} g_{6}(y) \quad \text { where } \quad g_{6}(y):=\sum_{n=0}^{\infty} \frac{y^{8 n+6}}{8 n+6}
$$

yields

$$
g_{6}(y)=-\frac{1}{8} \sum_{k=1}^{8} i^{k} \log \left(1-\zeta_{8}^{k} y\right)
$$

Now note that the familiar BBP-formula for $\pi$ can be written as

$$
\pi=4 f_{1}\left(\frac{1}{16}\right)-2 f_{4}\left(\frac{1}{16}\right)-f_{5}\left(\frac{1}{16}\right)-f_{6}\left(\frac{1}{16}\right)
$$

Setting $y=1 / \sqrt{2}$, so that $y^{8}=x$ and $|y|<1$, this becomes
$\pi=4 \frac{1}{y} g_{1}(y)-2 \frac{1}{y^{4}} g_{4}(y)-\frac{1}{y^{5}} g_{5}(y)-\frac{1}{y^{6}} g_{6}(y)=4 \sqrt{2} g_{1}\left(\frac{1}{\sqrt{2}}\right)-8 g_{4}\left(\frac{1}{\sqrt{2}}\right)-4 \sqrt{2} g_{5}\left(\frac{1}{\sqrt{2}}\right)-8 g_{6}\left(\frac{1}{\sqrt{2}}\right)$
Using the obtained expressions for $g_{1}(y), g_{4}(y), g_{5}(y)$ and $g_{6}(y)$, a straightforward (but very lengthy, hence omitted) computation tells us that this is equal to $2 i \log (-i)$. This is irrational and transcendental due to the aforementioned results on linear combinations of logarithms. (Of course, the transcendence of $\pi$ is already well-known.)

Even though this result is quite interesting, we still know no nontrivial rational cases of Theorem 7.2. Also, no cases of hypergeometric function evaluatons that correspond to nonzero rational BBP-numbers are currently known. Due to this, all BBP-numbers we know are either irrational or zero. If the existence of nonzero rational BBP-numbers can be proven or disproven, one could perhaps make stronger statements about properties of certain BBP-numbers, for example about their digit distributions. Maybe one could also gain new insights into the validity of Hypothesis 6.3. However, efforts in this direction do not seem to produce the desired results. They are not entirely fruitless however; along the road, some interesting side results have been encountered.

## 8. Conclusion

It is widely suspected that mathematical constants like $\pi, e, \log 2, \sqrt{2}$ and numerous others are (absolutely) normal. An interesting approach towards possibly proving normality for a certain class of constants has been initiated quite recently by the discovery of the so-called BBP-type formulae. Even though under some conditions normality to a base $b$ (and then also to bases that are a power of $b$ ) can be established for these constants, the validity of these conditions still remains equally unproven.

Despite numerous efforts, no proof of normality for any of these constants has ever been given, not even to one particular base. In fact, every attempt to gain more insight into the distribution of digits seems to give rise to an whole new class of problems, and the discovery of some interesting properties along the way, that may or may not seem directly useful for solving the question of normality.

Even if one would be able to make statements about the normality of BBP-numbers, there are a lot of other constants, for instance $\sqrt{2}$ and $e$, for which a BBP-type formula is not even known. Perhaps some day, more hidden knowledge about normality can be uncovered, but for now, it remains one of the most intriguing unsolved problems in mathematics.

## Bibliography

[Bailey, 2004] D.H. Bailey, "A Compendium of BBP-Type Formulas for Mathematical Constants", 2004. (At the author's webpage, http://crd.lbl.gov/~dhbailey/dhbpapers/)
[Bailey, 2005] D.H. Bailey, "A "Hot-Spot" Proof of Normality for the Alpha Constants", 2005. (At the author's webpage, http://crd.lbl.gov/~dhbailey/dhbpapers/)
[Bailey, Borwein, Borwein \& Plouffe, 1997] D.H. Bailey, J.M. Borwein, P.B. Borwein and S. Plouffe, "The Quest for Pi", Mathematical Intelligencer, vol. 19, no. 1, 1997.
[Bailey, Borwein \& Plouffe, 1997] D.H. Bailey, P.B. Borwein and S. Plouffe, "On the Rapid Computation of Various Polylogarithmic Constants", Mathematics of Computation, vol. 66, no. 218.
[Bailey \& Crandall, 2001] D.H. Bailey and R.E. Crandall, "On the Random Character of Fundamental Constant Expansions", Experimental Mathematics vol. 10, no. 2, 2001.
[Kuipers \& Niederreiter, 1974] L. Kuipers and H. Niederreiter, Uniform distribution of sequences, Wiley-Interscience, New York, 1974.
[Lagarias, 2001] J.C. Lagarias, "On the Normality of Arithmetical Constants", Experimental Mathematics vol. 10, no. 3, 2001.

## Addendum

The derivation on pages $22-24$ is NOT intended to be a transcendence proof for $\pi$, nor should it be read as such. We merely restated the fact that $\pi=2 i \log (-i)$ using the BBP-formula. This is transcendental, using a result by Baker ${ }^{1}$ which tells us that finite sums of linear forms in logarithms with algebraic coefficients, evaluated at algebraic points, are transcendental if and only if the sum of the logarithmic terms is nonzero.

[^0]
[^0]:    ${ }^{1}$ A. Baker, Transcendental number theory, Cambridge University Press, London, 1975

