

Analysis of the High Water Mark Convergents of Champernowne's Constant in Various Bases

John K. Sikora
john.sikora@xtera.com

August 4, 2014

Abstract

In this paper, we show that patterns exist in the properties of the High Water Mark (HWM) convergents of Champernowne's Constant in various bases (C_b), specifically in bases 2 through 124. The convergents are formed by truncating the Continued Fraction Expansion (CFE) of C_b immediately before the CFE HWMs. These patterns have been extended from the known patterns in the CFE HWMs of Champernowne's Constant in base ten (C_{10}). We show that the patterns may be used to efficiently calculate the CFE coefficients of C_b , and to calculate and predict the lengths of the HWM coefficients, the number of correct digits of C_b as calculated by the convergent, and the convergent error. We have discovered that minor corrections to the pattern formulations in base 10 are required for some bases, and these corrections are presented and discussed. The resulting formulations may be used to make the calculations for C_b in any base.

1 Introduction

Champernowne's Constant, C_{10} , was formulated in 1933 by British mathematician and economist D. G. Champernowne as an example of a normal number [1]. It is formed by concatenating the positive integers to the right of a decimal point, i.e., 0.123456789101112..., without end. In 1937, Kurt Mahler proved that C_{10} is transcendental [2].

It is well known that the CFE of C_{10} consists mostly of coefficients with a reasonably small number of digits, interspersed with coefficients with a very large number of digits. A coefficient that has a value that is larger than any previous coefficient is called a HWM. As one progresses along the CFE, the pattern of relatively small coefficients sprinkled with very large coefficients continues between the HWMs, even though the other large coefficients are not themselves HWMs.

There are many patterns in the properties of the convergents formed by truncating the CFE of C_{10} immediately before the HWMs for HWM numbers greater than 3. These properties and their patterns have been formulated and discussed [3, Conjectures 1-7]. The purpose of this paper is to extend the formulations and the conjectures from base 10 to bases 2 through 124, with exceptions noted and the pertinent formulae provided.

In addition, another conjecture which describes the lengths of the largest coefficients between the HWMs [4, Conjecture 8] is generalized from C_{10} to C_b . We show that deviations from the expected behavior occur in some bases and these deviations are noted and briefly discussed.

2 Definitions

Definition 1: C_{10} - Champernowne's Constant in base 10.

Definition 2: C_b - Champernowne's Constant in base b . For example C_3 is 0.12101112202122100101...₃

Definition 3: C_b position, or position - For an integer, it is the position of the first digit of the integer in the consecutive sequence used to generate C_b , starting at position 0 for the 0 to the left of the radix point.

The radix point is not counted. Therefore, for C_{10} , the integer 1 is in position 1 and the integer 12 is in position 14. For C_3 , the integer 11_3 is in position 5.

Definition 4: **CFE coefficient number**, or **coefficient number** - The number of the CFE coefficient, starting at coefficient number 0. Coefficient number 0 is always 0 for C_b .

Definition 5: **HWM number**, or **HWM #** - The number of the HWM. The first HWM is 0 and it is HWM number 1.

Definition 6: **HWM length** - The number of digits of the HWM of C_b as represented in base b .

Definition 7: 10_b^X - This is equal to b^X and it is read one-zero, or “ten”, to the power of X in base b . This notation is used instead of b^X primarily because the computer programs that were used in the analysis calculate the HWM convergents with two primary inputs: base b , and the power of “ten”, X . This terminology was chosen so that there would be consistency between this paper and the programs.

Definition 8: $P10_{b,X}$, or C_b position of 10_b^X - The formula for the C_b position of the integer 10_b^X , $X \geq 0$, is:

$$C_b \text{ position of } 10_b^X = P10_{b,X} = 1 + \sum_{x=0}^X (b-1) \cdot x \cdot b^{(x-1)} \quad (2.1)$$

Thus, the C_b position of 10_b^0 is 1, the position of 10_b^1 is b , and the position of 10_b^2 is 15.

Definition 9: 10_b^X **HWM convergent**, or 10_b^X **convergent** - The convergent formed from the numerator, $N_{b,X}$, and the denominator, $D_{b,X}$, as calculated using Conjecture 7 and Conjecture 6, respectively. The convergent is named in this manner because the C_b position 10_b^X is used in the calculation of the convergent. The CFE derived from the convergent does not calculate the 10_b^X HWM coefficient as defined in Definition 10. The HWM coefficient is, however, the next term in the CFE.

Definition 10: 10_b^X **HWM coefficient**, or 10_b^X **coefficient** - The CFE coefficient immediately following the last coefficient as calculated by the 10_b^X convergent. With a few exceptions as noted later in this paper, these coefficients constitute the HWMs of the CFE of C_b .

Definition 11: **The denominator of the 10_b^X HWM convergent**, or $D_{b,X}$ - The denominator for the 10_b^X HWM convergent as constructed using Conjecture 6.

Definition 12: **The numerator of the 10_b^X HWM convergent**, or $N_{b,X}$ - The numerator for the 10_b^X HWM convergent as calculated using Conjecture 7.

Definition 13: **Number of correct digits**, or **NCD**, or $NCD(b, X)$ - The number of correct consecutive digits of C_b as calculated by a convergent, starting at, and counting the 0 to the left of the radix point. The radix point itself is not counted. Therefore, if the position of the last correct digit, p_c , is known, the number of correct digits is $p_c + 1$ since the 0 is counted as the first correct digit, but it is in position 0. When given with arguments, base b , and power of “ten” X , $NCD(b, X)$ is the NCD of the 10_b^X HWM convergent as defined in Definition 9.

In general, in this paper we refer to the HWM convergents in a manner different from that in our previous paper which detailed the HWM convergent properties of C_{10} [3]. In the previous paper, a HWM convergent is referred to as “the convergent before HWM #N”. In this paper a HWM convergent is referred to as “the 10_b^X HWM convergent”, or as the “ 10_b^X convergent” as defined in Definition 9. The reason is that in this paper, bases 2 through 124 are discussed, rather than base 10 exclusively, and the HWM numbers are different for the various bases.¹

¹The HWM numbers match for bases 6 through 124, but not for bases 2 through 5. See Conjecture 1 and Conjecture 5 for details.

3 Method

A program was written in Ruby to calculate C_b in the desired base by taking the digits of each of the integers used to form C_b , dividing by the appropriate power of the base for each digit, and summing. The Euclidean Algorithm applied to real numbers was then used to calculate the CFE coefficients of C_b . The use of this method was difficult and care was taken to calculate C_b to sufficient accuracy so that the desired number of coefficients could be accurately calculated. Bases 2, 7, 8, 9, and 10 were used in the initial investigation.

It was noticed that large coefficients (HWMs) in the CFE of C_b appeared in other bases, similar to the behavior of the CFE of C_{10} . The CFE coefficients of C_b , terminated before the HWMs were then used to calculate the numerator and denominator of each convergent. Patterns were observed in the denominators of the convergents as given in Conjecture 6. The means of calculating the numerators of the convergents was determined as given in Conjecture 7.

Programs were then written in C to calculate the numerators and denominators using these conjectures and the resulting calculations led to the formulation of Conjecture 1 through Conjecture 5. These calculations were checked for each 10_b^X convergent and HWM as described in Section 5.1.

4 Conjectures

In this section we extend the C_{10} conjectures [3, Conjectures 1-7] to C_b . The order of the conjectures in the present paper matches the order of the conjectures in the previous paper for ease of comparison. However, Conjecture 1 through Conjecture 5 follow from analyzing the convergents formed using Conjecture 6 and Conjecture 7.

Conjecture 1: *The 10_b^X convergent, for $X \geq 0$, calculates the value of C_b correctly until the integer $10_b^{X+1} - 2$ in the integer sequence used to form C_b is reached. Furthermore, this integer is calculated as $10_b^{X+1} - 1$ instead of $10_b^{X+1} - 2$.*

Exception 1.1 For base 2, the convergents that calculate C_2 per Conjecture 1 start with the 10_2^X convergent, $X \geq 2$, rather than $X \geq 0$.

Exception 1.2 For base 3 and base 4, the HWM convergents that calculate C_3 and C_4 per Conjecture 1 start with the 10_3^X and 10_4^X convergent, $X \geq 1$, rather than $X \geq 0$.

The following are examples of the conjecture:

the 10_{10}^1 HWM convergent calculates C_{10} as “0.123456...95969799...” instead of “0.123456...95969798...”. Similarly, the 10_5^2 HWM convergent calculates C_5 as “0.12341011...440441442444...5” instead of “0.12341011...440441442443...5”.

Conjecture 2: *The NCD of the 10_b^X HWM convergent for base b , $X \geq 0$ is given by:*

$$NCD(b, X) = C_b \text{ position of } 10_b^{X+1} - X - 2 \tag{4.1}$$

Exception 2.1 For $b = 2$, Equation (4.1) is valid for $X \geq 2$.

Exception 2.2 For $b = 3$ and $b = 4$, Equation (4.1) is valid for $X \geq 1$.

The C_b position of 10_b^X is given in Equation (2.1). Thus, the $NCD(10,0)$ is 8, the $NCD(5,3)$ is 2340, the $NCD(16,5)$ is 99544809, and the $NCD(10,7)$ is 788888881.

Conjecture 3: *The 10_b^X HWM coefficient numbers are even. The coefficient numbers start at 0 per Definition 4.*

The 10_b^X HWM convergents calculate C_b slightly larger than C_b . Per Definition 4, this means that the 10_b^X HWM coefficient numbers will be even.

This conjecture is very useful at the end of the calculation of the CFE coefficients as the algorithm cannot distinguish the degenerate case where the last coefficient is 1 from the case where the last coefficient is not 1. In the degenerate case where the last coefficient equals 1, the algorithm calculates the last term as $Y + 1$ instead of Y , and the calculation ends on an even numbered coefficient. However, if the calculation ends on an even numbered coefficient, it is taken from Conjecture 3 that this is the degenerate case, the even numbered coefficient is Y , and the last coefficient (the one before the HWM) is 1.

Conjecture 4: *The error of the 10_b^X HWM convergent is given by:*

$$\text{error} = \begin{cases} 1.0 \times b^{-(\text{exponent}-1)} & \text{base } b \geq 5, X = 0 \\ (b-1).1_b \times b^{-\text{exponent}} & \text{base } b \geq 3, X = 1 \\ 1.01_2 \times b^{-\text{exponent}} & \text{base } b = 2, X = 2 \\ (b-1).\underbrace{00\dots00}_{\text{number of zeroes}}(b-1)(b-1)_b \times b^{-\text{exponent}} & b = 2, X \geq 3; b \geq 3, X \geq 2 \end{cases} \quad (4.2)$$

$$\text{number of zeroes} = X$$

$$\text{exponent} = \text{NCD}(b, X) + X + 2$$

For example, the 10_8^0 HWM has an error of 1.0×8^{-7} , the 10_3^1 HWM has an error of $2.1_3 \times 3^{-15_{10}}$, the 10_2^2 HWM has an error of $1.01_2 \times 2^{-18_{10}}$, the 10_{10}^2 HWM has an error of $9.0099_{10} \times 10^{-2890_{10}}$, and the 10_{15}^5 HWM has an error of $e.00000ee_{15} \times 15^{-67530135_{10}}$. Note that the mantissa is given in base b while the exponent is given in base ten, similar to the output from the GNU Multiple Precision Arithmetic (GMP) library [5] function `mpf_out_str`. This library was utilized by the programs used in the analysis.

The value of *exponent* is also equal to the C_b position of 10_b^{X+1} . Note that for the base $b = 2, X = 2$ case we could have explicitly provided the value of *exponent*, but we wanted to show that the value of *exponent* was consistent for all of the given cases as formulated.

Conjecture 5: *Let:*

$$(b == 2) = \begin{cases} 0 & b \neq 2 \\ 1 & b = 2 \end{cases}$$

$$(b == 4) = \begin{cases} 0 & b \neq 4 \\ 1 & b = 4 \end{cases}$$

The length, as represented in base b , of the HWM coefficient that occurs after the last coefficient of the CFE calculated by the 10_b^X HWM convergent for $X \geq 0$ is given by:

$$\text{length} = \text{NCD}(b, X) - 2 \cdot \text{NCD}(b, X - 1) - 3 \cdot (X - (b == 2)) - 2 + (b == 4) \quad (4.3)$$

Exception 5.1 For base 2, Equation (4.3) is valid for $X \geq 2$, rather than $X \geq 0$. In addition, the 10_2^2 coefficient is coefficient number 6 and it is of length 3, per Conjecture 5. It is not a HWM but the convergent does calculate C_2 per Conjecture 1. The coefficient number of the next HWM, which is the 10_2^3 HWM coefficient, is 14.² All subsequent HWMs are 10_2^X HWM coefficients.

Exception 5.2 For base 3 and base 4, Equation (4.3) is valid for $X \geq 1$, rather than $X \geq 0$. For base 3, the 10_3^1 HWM coefficient is coefficient number 6, and for base 4, the 10_4^1 HWM coefficient is coefficient number 12. Both are HWMs, and subsequent HWMs are 10_b^X HWM coefficients.

²The 10_2^2 coefficient has a value of 5, one less than the value of coefficient numbers 2 and 5, which are the largest terms along with coefficient number 10 until the 10_2^3 HWM coefficient at coefficient number 14.

Exception 5.3 For base 5, the 10_5^0 coefficient is coefficient number 4 and it is of length 1, per Conjecture 5.³ It is not a HWM but the convergent does calculate C_5 per Conjecture 1. The coefficient number of the 10_5^1 HWM coefficient is 10. The HWMs at coefficient number 7 and coefficient number 9 are not 10_5^X HWM coefficients.

For example, the 10_2^3 HWM as represented in binary has a length of 9, the 10_{10}^2 HWM as represented in decimal has a length of 2504, and the 10_{16}^5 HWM as represented in hexadecimal has a length of 89198852.

For the 10_b^0 coefficient, $b \geq 5$, the coefficient length is $b - 4$. For bases greater than 5, the 10_b^0 coefficient is a HWM, and it will be HWM #4 at coefficient number 4, per Definition 4 and Definition 5. All subsequent HWMs are those generated by a 10_b^X convergent, $X \geq 1$.⁴

Furthermore, it appears that for $b \geq 5$, the first 4 CFE terms are $[0; b - 2, b - 1, 1] = [0; b - 2, b]$. This gives a 10_b^0 convergent of

$$10_b^0 \text{ convergent} = \frac{b}{(b-1)^2}, \quad b \geq 5 \quad (4.4)$$

As stated above, the 10_b^0 convergent is a HWM for $b \geq 6$. For the numerators and denominators for other convergents, see Equation (4.6) and Equation (4.5), respectively.

It should be noted that for the lowest value of X allowed in any base, the $NCD(b, X - 1)$ term is not the true NCD as the 10_b^{X-1} convergent is not a valid HWM convergent. However, the calculated value as used in Equation (4.3) gives the correct HWM length. It should also be reiterated that the HWM following the 10_b^X HWM convergent is not calculated by the convergent. It is calculated by the 10_b^{X+1} HWM convergent, or by other means [6].

It is unknown why the HWM lengths in base 4, as formulated, are longer by 1 digit than the HWM lengths in other bases (other than base 2). It is also unknown why the HWM lengths in base 2, as formulated, are longer by 3 digits than the HWM lengths in other bases (other than base 4). However, in the base 2 case, we feel that since there are no 10_b^X HWM convergents for $X < 2$, this may increase the lengths of the HWM coefficients.

Conjecture 6: The denominator, $D_{b,0}$, used to calculate the 10_b^0 HWM convergent for base, $b \geq 5$, $X = 0$, is given in Equation (4.4). The denominator used to calculate the 10_b^X HWM convergent for base, b , $X \geq 1$, is given by:

$$D_{b,X} = \begin{cases} \left(\underbrace{(b-1) \cdot (b-1) \dots (b-1)}_{\text{number of } (b-1) \text{ digits}} \cdot \underbrace{(b-2) \overbrace{0 \dots 0}^{\text{number of zeroes}} 1}_{\text{number of zeroes}} \right) \times b^{\text{exponent}} & b \text{ is odd} \\ \left(\underbrace{\left(\frac{b}{2} - 1\right) \cdot (b-1) \dots (b-1)}_{\text{number of } (b-1) \text{ digits}} \cdot \underbrace{0 \dots 0}_{\text{number of zeroes}} \left(\frac{b}{2}\right) \right) \times b^{\text{exponent}} & b \text{ is even} \end{cases} \quad (4.5)$$

$$\begin{aligned} & \text{number of } (b-1) \text{ digits} = X \\ & \text{number of zeroes} = X \\ & \text{number of } (b-1) \text{ digits} = X \\ & \text{number of zeroes} = X + 1 \\ & \text{exponent} = NCD(b, X - 1) + 2X + 1 \end{aligned}$$

Exception 6.1 For base 2, Equation (4.5) is valid for $X \geq 2$, rather than $X \geq 1$.⁵

³The coefficient is 1.

⁴There is a close call for base 6 as coefficient number 8 is not a 10_6^X coefficient and it is only one less than coefficient number 4, the 10_6^0 coefficient.

⁵Although the 10_2^1 HWM convergent is not valid, $NCD(2, 1) = 3$ which gives a valid exponent of 8 for the 10_2^2 HWM convergent denominator, $0.110001_2 \times 2^{8 \cdot 10}$.

For example, $D_{2,4}$, the denominator for the 10^4_2 HWM convergent is $0.1111000001_2 \times 2^{54_{10}}$, $D_{5,3}$ is $4.4430001_5 \times 5^{348_{10}}$, $D_{10,8}$ is $4.999999990000000005_{10} \times 10^{788888898_{10}}$, and $D_{15,5}$ is $e.eeee000001_{15} \times 15^{3742640_{10}}$.

If $X = 1$ for an odd base, then the $(b - 1)$ term is to the left of the radix point and the $(b - 2)$ term is immediately to the right of the radix point. For example, the denominator for the 10^1_7 HWM convergent is $6.501_7 \times 7^{8_{10}}$.

Conjecture 7: *The numerator, $N_{b,0}$, used to calculate the 10^0_b HWM convergent for base, $b \geq 5$, $X = 0$, is given in Equation (4.4). For other b and X , $N_{b,X}$ is calculated as described below.*

Given:

1. The denominator, $D_{b,X}$, of a 10^X_b convergent as calculated from Conjecture 6
2. The C_b position of the integer 10^X_b , $P10_{b,X}$, as given by Equation (2.1)
3. $C_b[0.1 \dots P10_{b,X}]$ denotes C_b truncated to position $P10_{b,X}$ ⁶
4. $Ceil(f)$ denotes the integer immediately above non-integer f .

Then the numerator, $N_{b,X}$, $X \geq 1$, of the convergent is calculated as:

$$N_{b,X} = \begin{cases} Ceil(D_{b,X} \cdot C_b[0.1 \dots P10_{b,X}]) + 1 & b \text{ is odd} \\ Ceil(D_{b,X} \cdot C_b[0.1 \dots P10_{b,X}]) & b \text{ is even} \end{cases} \quad (4.6)$$

Exception 7.1 For base 2, Equation (4.6) is valid for $X \geq 2$, rather than $X \geq 1$.

For example, the calculation for $N_{2,3}$, the numerator for the 10^3_2 HWM convergent is: $Ceil(0.11100001_2 \times 2^{21_{10}} \cdot 0.110111001011101111_2) = 110000100000000100001_2$.

The calculation for $N_{3,2}$ is:

$$Ceil(2.21001_3 \times 3^{17_{10}} \cdot 0.121011122021221_3) + 1 = 112222220011222111_3.$$

The calculation for $N_{10,1}$ is:

$$Ceil(4.9005_{10} \times 10^{11_{10}} \cdot 0.1234567891_{10}) = 60499999499_{10}.$$

And the calculation for $N_{15,1}$ is:

$$Ceil(e.d01_{15} \times 15^{16_{10}} \cdot 0.123456789abcd1_{15}) + 1 = 120eeeeeeeeeeeed1_{15}.$$

5 Results, Verification, and Data Location

5.1 Results and Verification

The C_b CFE coefficients may be calculated very efficiently from the 10^X_b HWM convergents. For $X \geq 1$, the denominator of the convergent is given by Conjecture 6 and the numerator of the convergent may be calculated by multiplying the denominator by a sufficient number of digits of C_b . The required number of digits is $P10_{b,X} + 1$ as given in Conjecture 7, resulting in a very efficient calculation for the C_b CFE coefficients.⁷ The Euclidean Algorithm is then used to calculate the CFE coefficients from the numerator and the denominator. Conjecture 3 is used to determine if the last coefficient is 1. For $b \geq 5$, $X = 0$, Equation (4.4) is used to calculate the numerator and the denominator.

The properties of the convergents as given in Conjecture 1 through Conjecture 4 may be observed by dividing the numerator of the convergent by the denominator of the convergent and comparing the result with C_b for each base and power. The property of the convergents as given in Conjecture 5 may be checked by examination of the C_b CFE coefficients.

In the current study, the calculations and comparisons were performed for bases 2 to 124. Table 1 gives the base, b , and the powers of “ten”, X , in base b that were computed. The calculated results for *each base and each power* were checked against the predictions made in Conjecture 1 through Conjecture 5.

⁶The last digit is a 1 as it is the first digit of 10^X_b in the integer sequence used to form C_b .

⁷The required number of digits is $P10_{b,X} + 1$ because the zero to the left of the radix point has been included in the count.

The maximum power for each base was usually dictated by memory size. The amount of memory required to calculate C_b , the C_b CFE coefficients, and the error for the 10_b^{X-1} HWM convergent is approximately the same as the amount of memory required to calculate the C_b CFE coefficients for the 10_b^X HWM convergent, without calculating C_b and the error. There are separate columns for each type of calculation in Table 1. For more details, see Section 5.2 and Section 7.2.

Sequences A244330 [8], A244331 [9] for base 2, and A244332 [10], and A244333 [11], for base 3 have been added to the OEIS which correspond to sequences A038705 [12] and A143534 [13] in base 10, respectively. The first member of each pair of sequences gives the CFE coefficient number of each HWM coefficient and the second gives the length of each HWM coefficient as represented in the specified base.⁸

In addition, sequences A244758 [14] for base 2, and A244759 [15] for base 3 have been added to the OEIS which correspond to sequence A143532 [16] in base 10. These sequences give the number of digits as represented in the specified base for all of the calculated CFE coefficients.

We are considering adding similar sequences for base 4 through base 9. For these bases, there are existing OEIS sequences for the digits of C_b , but not for the CFE coefficients of C_b . Therefore, sequences for the CFE coefficients of C_b would be added as well as sequences for the HWM coefficients, the coefficient number of the HWMs, the length of all of the calculated coefficients, and the length of the HWM coefficients.

It should be noted that the maximum numbers of CFE coefficients as calculated and presented in Table 1 are not necessarily the maximum number of coefficients that have been calculated in a particular base. For example, at present for C_{10} , 82328 coefficients have been calculated [17].

Table 1: Calculated Bases, Powers, and Number of CFE Coefficients

base b	Min X	Max X w/ error	Max X w/o error	Number of CFE Coefficients (for max X w/o error)
2	2	25	24	98093504
3	1	15	16	2982556
4	1	12	13	629420
5	0	10	11	195554
6	0	9	10	105806
7	0	8	9	53596
8–10	0	7	8	base 8: 26362 base 10: 34062
11–14	0	6	7	base 11: 14424 base 14: 16386
15–16	0	5	6	base 15: 6080 base 16: 6258
17–27	0	4	5	base 17: 2382 base 27: 3184
28–62	0	3	4	base 28: 1104 base 62: 1830
63–90	0	2	3	base 63: 540 base 90: 574
91–124	0	1	2	base 91: 136 base 124: 128

A tabular view of the properties for a specific base, 9, appears in Table 2. To the best of our knowledge, the values in the shaded cells are predictions, and they have not been confirmed. A similar table, for base 10, is given in [3, Table 1].

5.2 Data Location and Digit Symbols

The computed data may be downloaded [19] and found under the “data” folder. The data is available under the Open Data Commons Public Domain Dedication and License (ODC PDDL) [20].

⁸The coefficient numbers start at 1 per the precedent set in A038705 [12].

Table 2: C_9 HWM Convergents Summary

Power of $10_9, X$	HWM Coefficient Number (next CFE term)	Integer that Fails ; Fails As, in Convergent Calculation	$NCD(9, X)$	Convergent Error	Convergent Denominator	Length of HWM (next CFE term)
0	4	$7_9; 8_9$	7	$1.0 \times 9^{-8_{10}}$	64_{10}	5
1	16	$87_9; 88_9$	150	$8.1_9 \times 9^{-153_{10}}$	$8.701_9 \times 9^{10_{10}}$	131
2	52	$887_9; 888_9$	2093	$\frac{8.0088_9}{\times 9^{-2097_{10}}}$	$\frac{8.87001_9}{\times 9^{155_{10}}}$	1785
3	152	$8887_9; 8888_9$	25420	$\frac{8.00088_9}{\times 9^{-25425_{10}}}$	$\frac{8.8870001_9}{\times 9^{2100_{10}}}$	21223
4	492	$88887_9; 88888_9$	287859	$\frac{8.000088_9}{\times 9^{-287865_{10}}}$	$\frac{8.888700001_9}{\times 9^{25429_{10}}}$	237005
5	1598	$888887_9; 888888_9$	3122210	$\frac{8.0000088_9}{\times 9^{-3122217_{10}}}$	$\frac{8.88887000001_9}{\times 9^{287870_{10}}}$	2546475
6	4512	$8888887_9; 8888888_9$	32882905	$\frac{8.00000088_9}{\times 9^{-32882913_{10}}}$	$\frac{8.8888870000001_9}{\times 9^{3122223_{10}}}$	26638465
7	12164	$88888887_9; 88888888_9$	338992920	$\frac{8.000000088_9}{\times 9^{-338992929_{10}}}$	$\frac{8.888888700000001_9}{\times 9^{32882920_{10}}}$	273227087
8	30410	$888888887_9; 888888888_9$	3438356831	$\frac{8.0000000088_9}{\times 9^{-3438356841_{10}}}$	$\frac{8.88888887000000001_9}{\times 9^{338992937_{10}}}$	2760370965
	Conjecture 3 (even numbered)	Conjecture 1	Conjecture 1, Conjecture 2	Conjecture 1, Conjecture 4	Conjecture 6, Equation (4.4)	Conjecture 5

The downloaded data must be unzipped and extracted. The data includes the numerator of the convergent and the CFE coefficients as calculated from the convergent for each base, b , and power of “ten”, X , as given in Table 1 under the “Max X w/o error” column. For the bases and powers under the “Max X w/ error” column, the data also includes C_b as calculated by the convergent as well as the convergent error. However, for $b = 2$, $X = 25$, the CFE coefficients are not given due to the lengthy calculation time.⁹

The base, b , and power, X , appear in each file name. The following prefixes identify the types of files:

- `cn_calc_` – C_b as calculated by the convergent
- `cn_cfe_coeffs_` – the CFE coefficients calculated from the convergent
- `cn_error_` – C_b as calculated by the convergent minus C_b (approximately 30 digits)
- `cn_numer_` – the numerator of the convergent

It should be noted that there are various suffixes, after the base and power, that indicate the specific program that was used to generate the data. These suffixes include, for example, whether or not the error was calculated.

The symbols used for the digits in various bases are summarized in Table 3. Of course, for any base, b , the symbol range is valid only for digits 0 through $b - 1$. Referring to the table, it may be seen, for example, that when examining a data file for a base 36 calculation, the symbol for the digit 35_{10} is “z”. However, when examining a data file for a base 37 calculation, the symbol for the digit 35_{10} is “Z” and the symbol for the digit 36_{10} is “a”. When examining a data file for a base 64 calculation, the symbol for the digit 63_{10} is “|1”, with the symbol representing a single character.

Table 3: Digit Symbols for the Various Bases

Base b	Digit Symbols					
	0–9	10–35	36–61	62–71	72–97	98–123
2–10	0–9					
11–36	0–9	a–z				
37–62	0–9	A–Z	a–z			
63–72	0–9	A–Z	a–z	0–9		
73–98	0–9	A–Z	a–z	0–9	a–z	
99–124	0–9	A–Z	a–z	0–9	A–Z	a–z

Discussions of the results of the processing of the data and compilations of the data may be found in Section 7.4.

6 2nd Generation HWM Length Extension to Bases 2 through 62

It is well known that in addition to the C_{10} HWM CFE coefficients, there are other coefficients that have a large number of digits that are not themselves HWMs. For example, coefficient 101 (starting from 0) is 140 digits long, but it is not a HWM since coefficient number 40 is 2504 digits long and coefficient number 18 is 166 digits long. Coefficient 101 is an example of what we have designated a 2nd Generation HWM [4].

It is known that for C_{10} , the lengths of the 2nd Generation HWMs appear to be related to the lengths of the HWMs [4, Conjecture 8]. The differences in the lengths of a HWM and its corresponding 2nd Generation HWM show a regular pattern, increasing by 10 for each HWM number increase. We have found that, in general, the same holds true for bases 2 through 62: the difference in the lengths of a HWM and its corresponding 2nd Generation HWM increases by 10 for each HWM number increase.^{10,11} However, there were irregularities in some bases as can be seen by the shaded cells in Table 4, which is a compilation of the data. Note, for example, the entries for bases 4, 5, 9, 22, and 36.

It should also be noted that starting at base 24, the first 2nd Generation HWM starts after HWM #5 rather than after HWM #6. Further analysis may show that this is true for base 23 and possibly base 22,

⁹The calculation is in progress.

¹⁰The length is the number of digits as represented in the base of C_b .

¹¹No analysis was performed for bases 63 through 124.

but this would most likely involve looking at the error of the potential 2nd Generation HWM convergents [4, Conjecture 9].

Table 4: Bases, Difference Between HWM Lengths and 2nd Generation HWM Lengths

base	HWM Length - 2 nd Generation HWM Length													
2	80	90	100	110	120	130	140	150	160	170	180	190	200	210
3	48	58	68	78	88	98	108	118	128	138	148			
4	36	47	57	67	77	87	97	107	117					
5	25	37	47	57	67	77	87	97						
6	25	35	45	55	65	75	85							
7	27	37	47	57	67	77								
8	25	35	42	55	65	75								
9	24	37	47	57	67									
10	26	36	46	56	66									
11	27	37	47	57	67									
12	26	36	46	56										
13	27	37	47	57										
14	26	36	46	56										
15	27	37	44	57										
16	26	36	46	56										
17	27	37	47											
18	24	36	46											
19	27	37	47											
20	26	36	46											
21	27	37	47											
22	26	36	43											
23	27	37	47											
24	16	26	36	46										
25	17	27	37	47										
26	14	24	36	46										
27	17	27	37	47										
28	16	26	36											
29	15	27	37											
30	16	26	36											
31	17	27	37											
32	16	26	34											
33	17	27	37											
34	16	26	36											
35	17	25	37											
36	15	26	36											
37	17	27	34											
38	16	26	36											
39	17	27	37											
40	16	26	36											
41	17	27	37											
42	16	26	36											
43	14	25	37											
44	16	26	36											
45	17	27	37											

continued on next page

Table 4 – continued from previous page

base	HWM Length - 2 nd Generation HWM Length												
46	16	26	36										
47	17	27	37										
48	16	26	36										
49	17	27	37										
50	15	26	36										
51	17	27	37										
52	16	24	36										
53	17	27	37										
54	16	26	36										
55	17	27	37										
56	16	26	36										
57	16	27	37										
58	16	26	36										
59	17	27	37										
60	15	24	36										
61	17	27	37										
62	16	26	36										

7 Notes on Computing

7.1 Ruby Programs

As mentioned in Section 3, a program was written in Ruby to look for the patterns in the CFE coefficients of C_b in various bases. In particular, the HWM coefficients were noted and another Ruby program was written to calculate the numerators and the denominators, as represented in the particular base, of the convergents terminated immediately before the HWM coefficients. The patterns in the denominators and the methods of calculating the numerators as given in Conjecture 6 and Conjecture 7, respectively, were determined. The main CFE calculating programs were then written in C due to speed considerations. Other programs were written in Ruby and C to compile the statistics from the C program outputs and to formulate and check the conjectures.

7.2 C Programs

The main C_b CFE coefficient calculation programs were written in C. Both the MPIR 32 bit [7] and GMP 64 bit multiple precision libraries were used. In general, the GMP library was used for the highest power of “ten”, X , for each base due to the speed advantage of GMP on Linux versus MPIR on Windows and also due to the larger data handling capability of the 64 bit GMP library versus the 32 bit MPIR library.

There are three main types of programs. The first two main types of programs each have two subtypes since the limit on the input and output base range of the MPIR and GMP libraries is 2 to 62, while the study was conducted to base 124. The last main type finds use only for low bases and large values of X , so the extension to higher bases is unnecessary. The following list shows the main types and subtypes of programs:

- Programs that calculate the convergent, the CFE coefficients of the convergent, C_b as calculated by the convergent, and the convergent error
 - Programs for bases 2 through 62
 - Programs for bases 63 through 124
- Programs that calculate the convergent and the CFE coefficients of the convergent

- Programs for bases 2 through 62
- Programs for bases 63 through 124
- A program that calculates the convergent, C_b as calculated by the convergent, and the convergent error

Table 5 gives the calculation times for the various bases and powers of “ten”, X . See Section 7.5 for more details on the calculations. It may be observed that the base 63 through base 124 calculation times increase rapidly with X as the base conversion processing is not as efficient as the MPIR and GMP input and output functions.

Currently, for bases from approximately 4 to 62, the limitation in these calculations for the maximum power, X , for a given base, b , is usually memory. However, for the lower bases, particularly base 2 and base 3, and higher powers, the number of coefficients that are calculated becomes very high. In these cases, the calculation time of the CFE coefficients becomes a major factor. As more memory becomes available, the CFE coefficient calculation time may become more of a factor for bases higher than base 3 at the higher powers.

The program that calculates the convergent, C_b as calculated by the convergent, and the convergent error, but not the CFE coefficients was written to further check the validity of the applicable conjectures for the highest power of “ten” in the lower bases without the very time consuming calculation of the coefficients. This program may also be used for higher bases but the calculation time savings is not nearly as beneficial.

For bases 63 to 124, the limitation is calculation time rather than memory due to the inefficient base conversion. However, as the base increases, the difference in the number of digits between powers of “ten” increases, and this factors into the increase in calculation time as well.

7.3 Program Location and Preferred Programs

All of the programs that were used in the study, including ones that are not mentioned, are available under the GNU LGPL [18]. To download the programs, see reference [19], under the “programs” folder. As there are many programs, some of which were used to check results, the preferred program for each main type used in the HWM convergent calculations is indicated in Table 6. The preferred programs for MPIR are Microsoft Visual Studio C++ projects, while the GMP programs consist of a C program and a C header file. For instructions on how to run the programs, see the `readme_mpir.txt` and `readme_gmp.txt` files that may be found in the appropriate folders.

7.4 Other Programs and Data of Interest

As already mentioned, several programs were written to process the data and compare the output of the main programs with the conjectures. One processed data set that we feel may be of general interest is a list of the CFE coefficient lengths for all coefficient numbers, not just the HWMs. A separate file has been created for each base, b , and the maximum power of “ten”, X , in which the coefficients were calculated. In some of the lower base cases, where the files may be very large (over 1 GB), compilations for lower values of X are included.

Once the `vcpp.tar.gz` file has been unzipped and extracted, these files may be found in the following folders:

- `get_hwm_lengths_generic_base`
- `get_hwm_lengths_generic_base_auto`
- `get_hwm_lengths_generic_base_v_hi_base_auto`

The name of the file includes b and X . An example name is `cn_cfe_coefs_base_2_pow_10_24_no_err_all_coefs_lengths.txt`.

There are other files in each of these directories for each b and X . These files have been processed from the files containing all of the coefficient lengths. The files with `_powers` at the end of the name give the HWM lengths for each b and X .

Table 5: Calculation Times

calculation time in seconds				
base b	X	CFE Calculation with Error	CFE Calculation without Error	With Error, No Coefficients
2	25			219
	24		2041980	105
3	15	21551		
	16		176523	
4	12	6048		
	13		59031	
5	10	1206		
	11		13169	
6	9	869		
	10		8908	
7	8	397		
	9		3161	
8-10	7	base 8: 56 base 10: 928		
	8		base 8: 569 base 10: 5524	
11-14	6	base 11: 130 base 14: 903		
	7		base 11: 476 base 14: 3330	
15-16	5	base 15: 60 base 16: 36		
	6		base 15: 139 base 16: 152	
17-27	4	base 17: 5 base 27: 88		
	5		base 17: 9 base 27: 172	
28-62	3	base 28: 1 base 62: 82		
	4		base 28: 3 base 62: 144	
63-90	2	base 63: 5271 base 90: 51821		
	3		base 63: 14275 base 90: 53217	
91-124	1	base 91: 1 base 124: 3		
	2		base 91: 1 base 124: 3	

Table 6: Preferred Programs

Type of Program	Applicable Bases	Preferred Program
w/ error	2-62	MPIR: <code>cn_generic_base_w_err_no_itoa_auto</code> GMP: <code>cn_generic_base_gmp_w_err_auto</code>
	63-124	MPIR: <code>cn_generic_base_w_err_very_hi_base_auto</code> GMP: <code>cn_generic_base_gmp_w_err_very_hi_base_auto</code>
w/o error	2-62	MPIR: <code>cn_generic_base_no_err_no_itoa_auto</code> GMP: <code>cn_generic_base_gmp_no_err</code>
	63-124	MPIR: <code>cn_generic_base_no_err_very_hi_base_auto</code> GMP: <code>cn_generic_base_gmp_no_err_very_hi_base_auto</code>
w/ error, no coeffs	2-62	GMP: <code>cn_generic_base_gmp_w_err_auto_no_coeffs_calc</code>

7.5 Hardware

The results given in Table 5 were obtained on a desktop computer with a 3.40 GHz i7-2600 CPU, running a single thread, with 8 GB of 1333 MHz DDR3 SDRAM. The programs were run with GMP on Linux, and GMP was built via the command line using `--build=x86_64-intel-linux-gnu` to optimize GMP to 64 bit operation on an Intel x86 processor. Additional tuning per GMP instructions was attempted, but no performance increase was observed. Building GMP in this manner resulted in code that ran approximately twice as fast as the generic C code.

The MPIR programs were run in 32 bit mode on Windows and the build mode was for generic C code. The GMP programs ran about 8 times faster than the MPIR programs.

8 Suggestions for Future Analysis

8.1 Extending the powers of “ten”

Naturally, extending the maximum power of “ten”, X , for each base and checking the calculations against the conjectures is desirable. Generally, the amount of memory required increases by a factor of approximately $1.25 \cdot base$ for each increment in power. When extending the calculations to a higher power of X , we feel that it is very important to make the calculations with the program that calculates C_b as calculated by the convergent as well as with the program that calculates the CFE coefficients without calculating C_b , with no more than one power difference between the two. We believe that ensuring that the calculation of C_b conforms to Conjecture 1 for power $X - 1$ increases the credibility of the calculations of the CFE coefficients for power X .

As explained in Section 7.2, for the lower bases, computing time becomes problematic. In general, for bases 62 and lower, each increment in power increases the time by a factor of approximately $2.5 \cdot base$, with a minimum of approximately a factor of 8 for base 2 and high powers of X .

8.2 Analysis of the ratio of adjacent HWM coefficient numbers

We feel that it would be interesting to compare the ratio of the coefficient numbers of successive HWMs in the various bases. A base dependent formula may exist for the limit of this ratio as the power of “ten”, X increases without bound.

8.3 More efficient algorithm for calculating the CFE coefficients

Implementation of a faster, more efficient algorithm to calculate the CFE coefficients would be of great benefit for the lower bases and higher powers of “ten”.

8.4 Further analysis of the 2nd Generation HWM coefficients

Further analysis of the 2nd Generation HWMs is suggested to see if the difference in HWM and 2nd Generation HWM lengths stabilizes at 10 as the powers of “ten” go higher. It would also be of interest to see if the 2nd

Generation HWM convergent error, and denominators in other bases have similar patterns to those in base 10 [4].

8.5 More efficient base conversion for bases greater than base 62

The current programs for bases greater than base 62 have inefficient methods for calculating C_b and for converting between bases. This results in greatly increased calculation times for higher powers of “ten”. Therefore, it is desired that more efficient algorithms be employed.

9 Conclusions

We have determined that patterns exist in the CFE HWM convergents for Champernowne’s Constant in bases 2 to 124, similar to the patterns that exist in the CFE HWM convergents for Champernowne’s Constant in base 10. These patterns have been presented in the form of formulae in conjectures. In some bases, slight changes in the formulae are required. There are also a few exceptions to the conjectures, depending upon the base and the power of “ten”. These differences and exceptions have been presented and discussed.

A cursory analysis of the 2nd Generation HWM lengths in base 2 to base 62 has been presented and the results and exceptions to the expected results have been discussed. It was suggested that further study of the 2nd Generation HWM convergents would be worthwhile.

The programs used in the analysis, and the data calculated by the programs, were discussed. The location of the programs and data was given. The programs are available under the GNU LGPL. The data is available under the ODC PDDL.

References

- [1] D.G. Champernowne, 1933: *The construction of decimals normal in the scale of ten*, J. London Math. Soc. 8, p. 254-260.
- [2] K. Mahler, 1937: *Arithmetische eigenschaften einer klasse von dezimalbrchen*, Proc. Konin. Neder. Akad. Wet. Ser. A. 40, p. 421-428.
- [3] John K. Sikora, 2012: *On the high water mark convergents of Champernowne's Constant in base ten*, <http://arxiv.org/abs/1210.1263> [math.NT], p. 4-10.
- [4] Ibid., p. 12-18.
- [5] T. Granlund, 2012: *GNU multiple precision arithmetic library, version 5.0.5*, <http://gmplib.org/>
- [6] E. Weisstein, 2013: *Champernowne Constant*, <http://mathworld.wolfram.com/notebooks/Constants/ChampernowneConstant.nb>
- [7] William Hart, 2010: *MPIR multiple precision integers and rationals library, version 2.5.1*, <http://www.mpir.org>
- [8] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A244330>.
- [9] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A244331>.
- [10] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A244332>.
- [11] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A244333>.
- [12] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A038705>.
- [13] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A143534>.
- [14] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A244758>.
- [15] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A244759>.
- [16] OEIS Foundation Inc, 2011: *The On-Line Encyclopedia of Integer Sequences*, <http://oeis.org/A143532>.
- [17] Wolfram Mathworld, 2013: *Champernowne Constant Continued Fraction*, <http://mathworld.wolfram.com/ChampernowneConstantContinuedFraction.html>.
- [18] Free Software Foundation, 2007: *GNU Lesser General Public License v3*, <https://www.gnu.org/licenses/lgpl.html>.
- [19] John K. Sikora, 2014: *Champernowne's Constant File Repository*, <https://drive.google.com/folderview?id=0B2oQudZQvHfSZ1hZN2VIYmU5eTQ&usp=sharing>.
- [20] Open Knowledge Foundation, 2008: *Open Data Commons Public Domain Dedication and License, v1.0*, <http://opendatacommons.org/licenses/pddl/1.0/>