ON PROPERTIES OF REPRESENTATIONS IN CERTAIN LINEAR NUMERATION SYSTEMS

MARCIA EDSON

Abstract. Given $a \geq b$, let $G_0 = 1, G_1 = a+1$, and $G_{n+2} = aG_{n+1} + bG_n$ for $n \geq 0$. For each choice of a and b, we have a linear recurrence that defines a numeration system. Every positive integer n may be written as the sum of the G_n, with alphabet $A = \{0, 1, \ldots a\}$, in one or more different ways. Let $R_{(a,b)}(n)$ be the function that counts the number of distinct representations of an integer as a sum of the G_n. We extend results of J. Berstel, P. Kocábová, Z. Masáková, and E. Pelantová, and M. Edson and L. Q. Zamboni and give two distinct methods for calculating $R_{(a,b)}(n)$. One formula involves products of 2×2 matrices and the other sums of binomial coefficients modulo 2. For the main result, we consider the limiting measure μ_β of a convergent infinite convolution of measures (Bernoulli convolutions), where β is the dominating root of the characteristic equation of the recurrence above. We study the Garsia entropy of these measures and calculate explicitly the limiting entropy associated with μ_β. This result extends those of J. Alexander and D. Zagier, and P. J. Grabner, P. Kirschenhofer, and R. F. Tichy. We then see that all these results can be generalized further to confluent numeration systems.

1. Introduction and Preliminaries

In this paper, we study the sequence-based numeration systems given by the linear recurrence

$$\begin{aligned} G_{n+2} &= aG_{n+1} + bG_n \text{ for } n \geq 0, \\ G_0 &= 1, G_1 = a+1 \text{ where } a, b \in \mathbb{N}, a \geq b. \end{aligned} \tag{1.1}$$

The most well known of these is the Fibonacci numeration system, obtained when $a = b = 1$.

Each positive integer n may be expressed as a sum of the following form,

$$n = \sum_{i=0}^{k} d_i G_i \tag{1.2}$$

where $d_i \in \{0, 1, \ldots, a\}$, for $0 \leq i \leq k$ and $d_k > 0$. We call the associated word $d_k d_{k-1} \ldots d_0$ a *representation* of n over the alphabet $A = \{0, 1, \ldots a\}$. We may obtain a unique representation for each n via the greedy algorithm. Let k be the unique integer such that $G_k \leq n < G_{k+1}$. Then $n = d_k G_k + n_k$, where $0 \leq n_k < G_k$. Generally, let $n_{i+1} = d_i G_i + n_i$,

Date: July 4, 2011.

The author is supported by the Austrian Science Foundation FWF, project S9605, part of the Austrian National Research Network “Analytic Combinatorics and Probabilistic Number Theory”.

where $G_i \leq n_{i+1} < G_{i+1}$ and $0 \leq n_i < G_i$. Iterating this process for $0 \leq i \leq k$, we may obtain a unique expression of the form (1.2). We call the representation obtained by the greedy algorithm, the greedy representation. While this particular representation is unique for all n, not all positive integers have only one representation in the G_k-based numeration system. However, the greedy representation has the property that it is the largest with respect to the lexicographic order. General information about representations in various bases is given in [8] and [17] while results about Fibonacci representations and representations in more generalized settings can be found, for example, in [2, 5, 6, 11, 14].

Define the sets $S = \{00, 01, \ldots, 0(a-b)\}$ and $T = \{aa, a(a-1), \ldots, ab\}$, and suppose that a word w is the representation of an integer n in base G_k. If w avoids elements of T, then w is the greedy representation of n. In order to see how to get one representation from another, let $s \in S$, and x be such that $1 \leq x \leq a$. Then any occurrence of a subword of the form xs in w may be replaced by the word $(x-1)t$, for some $t \in T$, to obtain an equivalent representation of n, and vice versa. If two words w and v are representations of the same positive integer n, we write $w \equiv v$. Words avoiding elements of the sets S and T have exactly one representation, and we call these words *ST-free*.

Example 1. Let $a = 5, b = 2$. Then $G_0 = 1, G_1 = 6, G_2 = 32, G_3 = 172, G_4 = 924, G_5 = 4964, G_6 = 26,668 \ldots$. The set $S = \{00, 01, 02, 03\}$ and the set $T = \{55, 54, 53, 52\}$. The integer $5481 = 4964 + 3(172) + 1$ leading to the greedy representation 103001. Using the replacement rule above, $xs \equiv (x-1)t$, we obtain the following four representations for 5481.

103001
102521
055001
054521

Consider the sequence $R_{(a,b)}(n)$ that counts the number of distinct partitions of n in the G_k base. Denote by A^* the set of all words over A, including the empty word, and set

$$\Omega_{(a,b)}(n) = \{w = w_0 w_2 \ldots w_k \in A^* : w_0 > 0 \text{ and } n = \sum_{i=0}^{k} w_i G_{k-i}\}.$$

Then $R_{(a,b)}(n) = \#\Omega(n)$. Further, consider the natural decomposition of $\Omega_{(a,b)}(n)$ given as follows. Let G be the largest term in the sequence $\{G_k\}$ less or equal to n, and let m be the largest integer such that mG remains less or equal to n. Let $\Omega^+(n)$ be the set of representations of n involving mG and $\Omega^-(n)$ the set of representations that do not. Then set $R^+_{(a,b)}(n) = \#\Omega^+(n)$ and $R^-_{(a,b)}(n) = \#\Omega^-(n)$. Clearly, $R_{(a,b)}(n) = R^+_{(a,b)}(n) + R^-_{N,(a,b)}(n)$. For simplicity, when no ambiguity exists, we simply write $R^+(n)$ and $R^-(n)$. Using the previous example, we see that $\Omega_{(5,2)}(5481) = \{103001, 102521, 055001, 054521\}$ so that $R_{(5,2)}(5481) = 4$, $\Omega^+_{(5,2)}(5481) = \{103001, 102521\}$ so that $R^+_{(5,2)}(5481) = 2$, and $\Omega^-_{(5,2)}(5481) = \{055001, 054521\}$ so that $R^-_{(5,2)}(5481) = 2$ Note that with a slight abuse of notation, we will sometimes write $R_{(a,b)}(w)$ instead of $R_{(a,b)}(n)$ for $w \in \Omega(n)$. Furthermore, we will simply write $R(n)$ instead of $R_{(a,b)}(n)$ when there is no possible ambiguity.

The function that counts the number of representations in a given base has been studied by many authors; some references include [2,4,5,6,14]. In section 2, we will give two formulas for the number of representations in these G_k-based numeration systems. These results make use of formulas previously established in [2,6].

Recurrence (1.1) is such that the dominating root $\beta_{(a,b)}$ of its characteristic equation satisfying

$$\beta_{(a,b)}^2 = a\beta_{(a,b)} + b$$

is a Pisot number. This follows directly from a result of A. Brauer in [3] as $a \geq b \geq 1$. We simply write β in place of $\beta_{(a,b)}$ unless there is a chance for ambiguity.

We consider sums of the form $\sum_{n=1}^{N} a_n \beta^{-n}$ where $a_n \in A = \{0, 1, \ldots, \lceil \beta \rceil - 1\}$. Let $A_N = \{x | x = \sum_{n=1}^{N} a_n \beta^{-n}\}$, and define a measure $\mu_N = (a+1)^{-N} \sum_{x \in A_N} r(x)\delta_x$, where $r(x)$ is the number of representations of x of length N in base β and δ_x denotes the unit point mass at x. Then these measures converge weakly to a measure μ_β. Jessen and Wintner [13] show that any convergent infinite convolution is either purely singular or absolutely continuous. In particular, we have that the measures μ_β are either purely singular or absolutely continuous.

In [7], Erdős proved that for $\beta = \frac{1+\sqrt{5}}{2}$, μ_β is purely singular. For further results, we refer to [15, 18, 19]. Garsia in [10], in order to study the measures μ_β further, introduced the idea of the Garsia entropy which is defined as

$$H(A_n) = -\sum_{x \in A_n} p(x) \ln p(x)$$

where $p(x) = \frac{r(x)}{(a+1)^n}$ is the weight assigned to x by μ_n. Then set

$$H_\beta = \lim_{N \to \infty} \frac{H(A_N)}{N \ln \beta}.$$

Garsia proved for general β (not just β satisfying recurrence (1.1)) that if $H_\beta < 1$, then μ_β is purely singular. Additionally, he showed that $H_\beta < 1$ for any Pisot number β. Though Garsia proved significant results involving H_β and μ_β, he did not give numerical values for H_β.

Alexander and Zagier in [1] consider the case $a = b = 1$, so that $\beta = \frac{1+\sqrt{5}}{2}$. Usually the problem of computing entropies is quite difficult but through a graph-theoretical argument, Alexander and Zagier give an explicit value for H_β, where $\beta = \frac{1+\sqrt{5}}{2}$. They make use of the Fibonacci graph, which can be built from the Euclidean tree. The Euclidean tree begins with one node at level 0 labeled with the pair $(1, 1)$ and one node at level 1 labeled with the pair $(2, 1)$. Then the nodes at level n are defined inductively as follows. Given a node at level n labeled (a, b), there are two edges (left and right) to nodes at level $n + 1$ labeled $(a + b, a)$ and $(a + b, b)$, respectively. Therefore this tree corresponds to the *subtractive Euclidean algorithm*, the Euclidean algorithm without division.

For any pair of relatively prime integers (k,i), we define the *length* $e(k,i)$ of the pair (k,i) to be the number of steps in the subtractive Euclidean algorithm applied to the pair k and i. In other words, $e(i,i)=0$ and $e(i+k,i)=e(i+k,k)=e(i,k)+1$.

Grabner, Kirschenhofer, and Tichy [11] give an explicit value for H_β in the case β is the dominating characteristic root of the m-bonacci recurrence which satisfies

$$\beta^m = \beta^{m-1} + \cdots + \beta + 1,$$

extending the results given in [1]. The graph-theoretic approach taken by Alexander and Zagier becomes significantly more complicated in this case. Therefore, they abandon this approach in favor of one using generating functions and the method of Guibas and Odlyzko for counting strings with forbidden subwords [12].

A generalization of the results of Grabner, Kirschenhofer, and Tichy [11] can be found in the doctoral dissertation of M. Lamberger, see [16]. Here, the case that is treated is given by the recurrence

$$\begin{aligned} G_{n+m} &= aG_{n+m-1} + \cdots + aG_{n+1} + aG_n \text{ for } n \geq 0, \\ G_0 &= 1, G_i = (a+1)^i, \text{ for } 1 \leq i \leq m-1, \text{ where } a \in \mathbb{N}. \end{aligned}$$

Therefore, when we discuss the Garsia entropy, we assume that $a > b$. We note here that the counting is necessarily more complicated in the case where $a > b$, due to the number of forbidden subwords. In the case $a = b$, the sets S and T only contain one element each.

In the situation of the general a and b we discuss in this paper, a graph-theoretic approach would lead to a non-planar graph. Therefore, we will abandon the more complicated graph-theoretical setting in favor of arguments using combinatorics on words. This leads to the use of generating functions and the method of Guibas and Odlyzko [12]. In Section 3, we prove the main result, which is as follows.

Theorem 1. *Let*

$$\kappa_n = \sum_{\substack{0<i<k \\ \gcd(k,i)=1 \\ e(k,i)=n}} k \ln k \;\; \text{and} \;\; \widetilde{\alpha}_n(x) = \sum_{\substack{0<i<n \\ \gcd(n,i)=1}} x^{2e(n,i)}.$$

Furthermore, let

$$\mathcal{T}(x) = \ln(a+1) - \widehat{M}(x) \sum_{N=1}^{\infty} \kappa_N x^{2N},$$

where

$$\widehat{M}(x) = \frac{(a-b+1)(1-x)\gamma(x)(1-3x^2)^2}{(a+1)(1+x)^3(1-(3+2a-2b)x^2)^2},$$

and

$$\gamma(x) = a + 2ax - (2+3a+2a^2-2b-2ab)x^2 + (2+4a+2a^2-6b-6ab+4b)x^3.$$

Then

$$H_{\beta(a,b)} = \frac{1}{\ln \beta_{(a,b)}} \mathcal{T}\left(\frac{1}{a+1}\right).$$

2. Counting Representations

Suppose W is a greedy representation of a positive integer n. What follows is a factorization of W, whereby we eliminate subwords of W that may not be replaced by equivalent representations. We shall call it the *principal factorization* of W, as in [6]. We may write

$$W = V_1U_1V_2U_2 \dots V_JU_JZ$$

where

- $V_1, V_2, \dots, V_J, Z$ are ST-free
- If V_i ends in 0, then U_i begins in a letter greater than $a - b + 1$
- If V_i ends in a, then U_i begins in a letter less than b
- Each U_i is of the form
$$U_i = r0x_k0x_{k-1} \dots 0x_00y$$
with $1 \leq r \leq a, 0 \leq y \leq a - b$, and $x_i \in \{0, 1, \dots, a - b + 1\}$.

Observe that the V_i and Z do not contribute to the number of ways to rewrite W using the replacement rule. Since the V_i and Z are all ST-free, there are no replacements to be made within these factors. Furthermore, with the restrictions placed on the ending of the V_i, we are guaranteed that no V_i "moves" into the U_i beside it. More precisely, if V_i ends in 0, then we may write $V_i = vx0$ where $x > 0$, and concatenating V_i and U_i, we obtain $vx0r0x_k0x_{k-1} \dots 0x_00y$ where $r > a - b + 1$. But we may not employ the replacement rule for the subword $x0r$, since $r > a - b + 1$. A similar argument holds when V_i ends in a. This leads us to the following result.

Lemma 1. *The number of representations of* W *is the product of the number of representations of the* U_i.

$$R_{(a,b)}(W) = \prod_{i=1}^{J} R_{(a,b)}(U_i).$$

Example 2. Let $a = 5$ and $b = 2$. We have that $W = 434100245111030011212121212$ is the greedy representation of some positive integer in the numeration system generated by the pair $(5, 2)$. Note that $W = (434)(100)(24511)(10300)(112121212) = V_1U_1V_2U_2Z$. We have that $R(U_1) = 2$ since $100 \equiv 052$, and $R(U_2) = 4$ from Example 1. Therefore, $R(W) = R(U_1)R(U_2) = 8$.

The lemma that follows is essentially Lemma 2, in [6].

Lemma 2. *Let w be a greedy representation of an integer n, in base G_k, having $a-b+1$ as its first letter. Let $1 \leq r_1, r_2 \leq a$. Then*

$$\begin{pmatrix} R^-(r_1 0^\ell w) \\ R^+(r_1 0^\ell w) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R^-(r_2 0^{\ell-2} w) \\ R^+(r_2 0^{\ell-2} w) \end{pmatrix} \text{ for } \ell \geq 3, \tag{2.1}$$

$$\begin{pmatrix} R^-(r_1 00 w) \\ R^+(r_1 00 w) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} R^-(w) \\ R^+(w) \end{pmatrix}, \tag{2.2}$$

$$\begin{pmatrix} R_-(r_1 0 w) \\ R_+(r_1 0 w) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} R^-(w) \\ R^+(w) \end{pmatrix}. \tag{2.3}$$

Proof. Consider equation (2.1). Since $R^-(r_1 0^\ell w)$ counts the number of representations of the form $(r_1-1)ab0^{\ell-2}w$, $R^-(r_1 0^\ell w) = R(b0^{\ell-2}w) = R(r_2 0^{\ell-2} w) = R^-(r_2 0^{\ell-2} w) + R^+(r_2 0^{\ell-2} w)$. And, since $R^+(r_1 0^\ell w)$ and $R^+(r_2 0^{\ell-2} w)$ count the number of representations fixing r_1 and r_2, respectively, $R^+(r_1 0^\ell w) = R(w) = R^+(r_2 0^{\ell-2} w)$.

Similar arguments hold for the remaining equations. □

Using the identities

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{d-1} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} d & d \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{d} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} d+1 & d \\ 1 & 1 \end{pmatrix}$$

we obtain that for any word of the form $U = r0^\ell w$, with w beginning in $a-b+1$,

$$\begin{pmatrix} R^-(U) \\ R^+(U) \end{pmatrix} = \begin{pmatrix} \lceil \frac{\ell}{2} \rceil & \lfloor \frac{\ell}{2} \rfloor \\ 1 & 1 \end{pmatrix} \begin{pmatrix} R^-(w) \\ R^+(w) \end{pmatrix} \text{ for } \ell \geq 1.$$

This gives the following result originally proven by Berstel in [2] for the case of Fibonacci.

Proposition 1. *Let $U = r0^{d_1}x_1 0^{d_2} x_2 \dots x_k 0^{d_k-1} y$, where $x_i \in \{0, a-b+1\}$ and $0 \leq y \leq a-b$. Then*

$$R(U) = \begin{pmatrix} 1 & 1 \end{pmatrix} \left(\prod_{i=1}^{k} \begin{pmatrix} \lceil \frac{d_j}{2} \rceil & \lfloor \frac{d_j}{2} \rfloor \\ 1 & 1 \end{pmatrix} \right) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We consider now the more general case where for $U = r0^{d_1}x_1 0^{d_2} x_2 \dots x_k 0^{d_k-1} y$, some $x_i \in \{1, \dots, a-b\}$. Denote such x_i as $y_1, y_2, \dots, y_j$ with $1 \leq j \leq m$, and rewrite

$$U = y_{j+1} t_j y_j t_{j-1} y_{j-1} \dots y_1 t_0 y_0$$

where $y_{i+1} t_i$ is as in Proposition (1). We now "inflate" U with a second copy of the y_i in order to apply the formula. Let $\widetilde{U} = (rt_j y_j)(y_j t_{j-1} y_{j-1}) \dots (y_1 t_0 y_0) = L_j L_{j-1} \dots L_0$. The following lemma shows that the number of representations of U is equal to the number of representations of the inflated copy of U since the L_i are independent. Since we may apply Berstel's formula to each factor L_i, we may use it to calculate $R(U)$ where $U = r0x_k0x_{k-1}\dots 0x_00y$ with $1 \le r \le a, 0 \le y \le a-b$, and $x_i \in \{0,1,\dots,a-b+1\}$, as in the principal factorization. Note that when $a = b$, Proposition (1) yields the number of representations $R(U)$, and the inflation rule is not defined in this case.

Lemma 3. *Let $\widetilde{U} = (rt_jy_j)(y_jt_{j-1}y_{j-1})\dots(y_1t_0y_0) = L_jL_{j-1}\dots L_0$. Then $R(\widetilde{U}) = R(U)$.*

Proof. We begin with the observation that since each $y_k \le (a-b)$ for $0 \le k \le j$, each y_k can be used to write a new representation of U_i by applying the exchange rule with y_k in the rightmost position. In other words, since y_k is preceded by 0, and $0y_k \in S$, we can obtain a new representation in which $0y_k$ is exchanged with $a(y_k+b)$. Similarly, since $y_k \ge 1$, each y_k can be used to write a new representation of u_i by applying the exchange rule with y_k in the leftmost position. Furthermore, these two ways of creating a new representation of U_i involving y_k, from the left or to the right, are independent of one another. So we may insert the extra copy of y_k without affecting the frequency. □

Berstel's approach gives us one method to calculate the number of representations of an integer in base G_k. We now discuss another approach considered in [6] for the m-bonacci base.

Lemma 4. *Let $U = r0x_k0x_{k-1}\dots 0x_00y$ where $r = a-b+1, x_i \in \{0, a-b+1\}$ for $0 \le i \le k$, and $0 \le y \le a-b$. Suppose that $1 \le z \le a$. Then*

$$\begin{aligned}
R^+(z0r0x_k0x_{k-1}\dots 0x_00y) &= R(U) = R^+(U) + R^-(U)\\
R^-(z0r0x_k0x_{k-1}\dots 0x_00y) &= R^-(U)\\
R^+(z000x_k0x_{k-1}\dots 0x_00y) &= R^+(U)\\
R^-(z000x_k0x_{k-1}\dots 0x_00y) &= R(U) = R^+(U) + R^-(U)
\end{aligned}$$

Proof. Note that $w \in \Omega^+(z0U)$ if and only if $w = z0w'$ for some $w' \in \Omega(U)$. Therefore, $R^+(z0U) = R(U)$. Next, we can see that $w \in \Omega^-(z0U)$ if and only if $w = (z-1)aw'$ for some $w' \in \Omega^-(U)$. It follows that $R^-(z0U) = R^-(U)$. A similar argument holds for the remaining identities. □

We may use Lemma (4) to compute the number of representations of an integer n whose representation is of the form $U = r0x_k0x_{k-1}\dots 0x_00y$ where $1 \le r \le a, x_i \in \{0, a-b+1\}$ for $0 \le i \le k$, and $0 \le y \le a-b$. We construct a tower of $k+2$ levels $L_0, L_1, \cdots, L_{k+1}$, where each level L_i consists of an ordered pair (a,b) of positive integers. We begin by setting $x_i' = 0$ if $x_i = 0$ and $x_i' = 1$ if $x_i = a-b+1$, and then fixing the positive integer $s = 1\cdot 2^{k+1} + x_k'\cdot 2^k + \dots + x_1'\cdot 2 + x_0'$. We start with level 0 by setting $L_0 = (1,1)$. Then L_{i+1} is obtained from L_i according to the value of x_i. Suppose that $L_i = (a,b)$. If $x_i = 0$, then $L_{i+1} = (a, a+b)$ and if $x_i = a-b+1$, then $L_{i+1} = (a+b, b)$. It follows from the Lemma (4) that $L_{k+1} = (R^+(U), R^-(U))$. Hence $R(U)$ is the sum of the entries of level L_{k+1}.

We note that, in the following proposition, each binomial coefficient is taken modulo 2 so that the formula for $R(U)$ simply is a sum of 0's and 1's. Because $R(U)$ is the sum of the entries of level L_{k+1}, the proof of Proposition 2 is essentially identical to that of Corollary 1 in [6], with only minor changes necessary. Therefore, the proposition will be stated without proof.

Proposition 2. *Let* $U = r0x_k0x_{k-1}\ldots0x_00y$ *where* $1 \leq r \leq a, x_i \in \{0, a-b+1\}$ *for* $0 \leq i \leq k$, *and* $0 \leq y \leq a-b$. *Further let* $s = 1\cdot 2^{k+1} + x'_k \cdot 2^k + \cdots x'_1 \cdot 2 + x'_0$. *Then*

$$R(U) = \sum_{j=0}^{s} \left[\binom{2s-j}{j} \pmod 2 \right].$$

Lemma 5. *Suppose* $U = r0x_k0x_{k-1}\ldots0x_00y$ *where* $1 \leq r \leq a, x_i \in \{0, 1, \ldots a-b+1\}$ *for* $0 \leq i \leq k$, *and* $0 \leq y \leq a-b$, *and let* $\widetilde{U} = L_\ell L_{\ell-1}\ldots L_0$ *be the inflated form of* U, *so that each* L_i *is of the form* $L_i = r_i 0x_{k,i}0x_{k-1,i}\ldots0x_{0,i}0y_i$. *If* $s_i = 1\cdot 2^{k+1} + x'_{k,i}\cdot 2^k + \cdots + x'_{1,i}\cdot 2 + x'_{0,i}$, *then*

$$R(U) = \prod_{i=0}^{\ell} R(L_i) = \prod_{i=0}^{\ell}\sum_{j=0}^{s_i} \left[\binom{2s_i-j}{j} \pmod 2 \right].$$

Proof. The proof follows directly from Proposition 2 and Lemma 1. □

Given a positive integer n, with principal factorization $W = V_1U_1V_2U_2\ldots V_jU_jZ$, we may calculate $R(n)$ as follows. Let $\widetilde{U_i} = L_{i_j}L_{i_{j-1}}\ldots L_{i_0}$ be the inflated version of U_i. It follows from Lemma 1 and Lemma 5 that

$$R(W) = \prod_{\ell=1}^{j} R(U_\ell) = \prod_{\ell=1}^{j}\prod_{M=0}^{i_\ell} R(L_{i_M}) = \prod_{\ell=1}^{j}\prod_{M=0}^{i_\ell}\sum_{N=0}^{s_{i_M}} \left[\binom{2s_{i_M}}{N} \pmod 2 \right].$$

As with Proposition (1), if $a = b$, Lemma (5) yields $R(U)$ and the inflated version of U is undefined.

3. The Garsia Entropy

Denote by A^* the set of all words over A, including the empty word and define an equivalence class of words on A^* as follows. We say two finite words v and w are equivalent if they are of the same length and represent the same number. We write $v \sim w$. Note that we allow leading zeros here. For a given word w, we define the *frequency* of the class represented by w as the size of the equivalence class of w. For example, for the word $w = 103001$, the frequency of the class of length six represented by w is 4. However, the frequency of the class represented by 55001 is 2, though the words 55001 and 103001 are representations of the same integer $n = 5481$. We set $\varphi(w)$ to be the frequency of the equivalence class represented by w. Then, we have that $\varphi(103001) = 4$ and $\varphi(55001) = 2$.

In Section 2, we consider the function that counts the size of the equivalence class of a word w obtained by the equivalence relation $\equiv_w$. Now if w is the greedy representation of n, we have that $R(n)$ is the size of the equivalence class of w obtained via $\equiv_w$. Note that $\varphi(w)$ is the size of the equivalence class of w obtained via $\sim_w$. Since leading zeros are allowed in our discussion of the Garsia entropy, we have that $R(n)$ and $\varphi(w)$ are not equal unless w has length greater or equal to the length of the greedy representation of w. So for each positive integer n, there is a positive integer m so that if w is a representation of n with $|w| \geq m$, then $\varphi(w) = R(n)$.

We now make use of generating functions to obtain results for the Garsia entropy. Let $F_N(k)$ denote the number of classes of words in A^* of length N having frequency k. Then $\sum_{k=1}^{\infty} kF_N(k) = (a+1)^N$, and

$$H(A_N) = -\sum_{k=1}^{\infty} kF_N(k)(a+1)^{-N} \ln\left(\frac{k}{(a+1)^N}\right) = N\ln(a+1) - \sum_{k=1}^{\infty} kF_N(k)(a+1)^{-N}\ln(k).$$

Set

$$f_k(x) = \sum_{N=0}^{\infty} F_N(k)x^N \text{ and } \Phi(x,s) = \sum_{k=1}^{\infty} k^s f_k(x).$$

Then we have,

$$\Phi(x,1) = \sum_{k=1}^{\infty} kf_k(x) = \sum_{N=0}^{\infty} (a+1)^N x^N = \frac{1}{1-(a+1)x}, \tag{3.1}$$

and

$$\left.\frac{\partial \Phi(x,s)}{\partial s}\right|_{s=1} = \sum_{\substack{k\geq 1\\ N\geq 1}} kF_N(k)\ln(k)x^N.$$

Therefore the generating function for the quantities $H(A_N)$ is given as

$$H(x) = \sum_{N=0}^{\infty} H(A_N)x^N = \frac{x\ln(a+1)}{(1-x)^2} - \left.\frac{\partial \Phi(x/(a+1),s)}{\partial s}\right|_{s=1}. \tag{3.2}$$

The generating function $G(x)$ of all ST-free words (including the empty word) is straightforward to obtain using the method of Guibas and Odlyzko (see [12]). It is given by

$$G(x) = \frac{x+1}{1-ax+(a-2b+1)x^2}.$$

Furthermore, the classes of frequency 1 can be generated by appending an ST-free string to any word in $\{0\}^* \cup \{a\}^*$, so that we obtain the generating function

$$f_1(x) = \frac{1}{1-x}G(x) + \frac{x}{1-x}G(x) = \frac{1+x}{1-x}G(x).$$

We call the class of a word *relational*, as in [11], if it has a representative ending in xs, for $1 \le x \le a$ and $s \in S$. So all relational word classes have frequency greater than 1. Consider all relational classes of frequency 2. We call these classes *relational prefix classes*. A relational prefix class has a representative of the form $vzx0y$ where vz is ST-free (so possibly empty), $1 \le x \le a$, and $0 \le y \le a-b$, but we must exclude those classes $vzx0y$ with $zx \in S, zx \in T$, and $zx = 0(a-b+1)$.

We denote by G_d the generating function for all ST-free strings ending in d for $d \in \{0, a\}$. Again using the method of Guibas and Odlyzko, we have that

$$G_d(x) = \frac{xG(x)}{x+1}.$$

Let $P(x)$ be the generating function of all relational prefix classes. To compute $P(x)$, we begin with all classes having a representative of the form $vzx0y$ where vz is ST-free and take away those that we excluded in the preceding paragraph. To account for the prefixes from the set $\{0\}^* \cup \{a\}^*$, we must multiply by a factor of $\frac{1+x}{1-x}$. Note that we add back the words $0xoy$ and $axoy$ which are members of relational prefix classes. Therefore, we have that

$$\begin{aligned} P(x) &= \frac{1+x}{1-x}\left[G(x)x^3a(a-b+1) - x^3(a-b+1)^2(G_0(x)+G_a(x)) + \frac{2x^4(a-b+1)^2}{x+1}\right] \\ &= \frac{x^3(a-b+1)\gamma(x)}{1-x^2}G(x). \end{aligned}$$

Denote by $r_k(x)$ the generating function of all strings in relational classes of frequency k. Since we may append an ST-free string to the end of a relational class representative without affecting frequency, we have that

$$f_k(x) = G(x)r_k(x), k \ge 2.$$

We now look further at the generating function $r_k(x)$.

Suppose that w is the greedy representative (lexicographically largest) of a relational word class such that $\varphi(w) = k$. We consider a factorization of w that will enable us to write an expression for the frequency of the class of w related to the subtractive Euclidean algorithm. This factorization is essentially the same as the principal factorization of w, with the exception of v_1. It is written differently to facilitate the use of generating functions and for reference purposes, we call it the *secondary factorization* of w. Factor w as

$$w = v_1e_1u_1v_2e_2u_2 \dots v_Je_Ju_J,$$

where

- each v_i is ST-free $(1 < i \le J)$,
- each e_i is of the form $r0x$ where $r \in \{1, \dots, a\}, x \in \{0, \dots, a-b+1\}$,
- v_ir is ST-free and does not end in $0(a-b+1)$,
- v_1 is of the form zg, where $z \in \{0\}^* \cup \{a\}^*$ and g is an ST-free word,
- and e_iu_i is of the form $r0x_m0x_{m-1}\dots0x_00y$ where $r \in \{1, \dots, a\}$, $x_\ell \in \{0, \dots, a-b+1\}$ for $0 \le \ell \le m$, and $0 \le y \le (a-b)$.

Considering this factorization, a similar argument as is used for Lemma (1) will show that $\varphi(w)$ is simply the product of the frequencies of the factors $e_i u_i$ as the v_i contribute nothing to the frequency of $[w]$.

So, if we set $U_i = e_i u_i$, we have that

$$w = v_1 U_1 v_2 U_2 \dots v_J U_J$$

and

$$\varphi(w) = \prod_{i=1}^{N} \varphi(U_i). \tag{3.3}$$

We now focus on the frequency $\varphi(U_i)$ where the word $U_i = r0x_m0x_{m-1}\dots 0x_00y$ is the greedy representation of some positive integer in the numeration system generated by the pair (a, b). Note that we are able to find a representative in $[U_i]$ of the form

$$\nu\eta_1 \dots \eta_{m+1} \text{ where } \eta_\ell \in S \cup T,$$

and $\nu = r0x_m$ if $x_m \leq a - b$ and $\nu = r0(x_m - 1)$ if $x_m = a - b + 1$. We use the exchange rule $xs \equiv (x-1)t$ for $s \in S$ and $t \in T$ to achieve this. Next, define two functions on this representative of $[U_i]$ as follows.

$$\begin{aligned} \varphi_1(\nu\eta_1 \dots \eta_{m+1}) &= \varphi_1(\nu\eta_1 \dots \eta_m) + \varphi_2(\nu\eta_1 \dots \eta_m), \\ \varphi_2(\nu\eta_1 \dots \eta_{m+1}) &= \begin{cases} \varphi_2(\nu\eta_1 \dots \eta_m) \text{ if both } \eta_m, \eta_{m+1} \in T \text{ or } \eta_m, \eta_{m+1} \in S \\ \varphi_1(\nu\eta_1 \dots \eta_m) \text{ otherwise,} \end{cases} \\ \varphi_1(\nu) &= 1 = \varphi_2(\nu) \end{aligned} \tag{3.4}$$

where $\nu \in \{rs : s \in S\}$.

Lemma 6. *Suppose $U = r0x_m0x_{m-1}\dots 0x_00y$ such that $U \sim u_1 s \sim u_2 t, s \in S, t \in T$. Then, $\varphi(U) = \varphi(u_1 0) + \varphi(u_2 a)$.*

Proof. We simply note that $\varphi(u_1)$ counts the number of representatives in $[U]$ having length $|U|$ with a suffix belonging to the set S, and that $\varphi(u_2)$ counts the number of representatives in $[U]$ having length $|U|$ with a suffix belonging to the set T. □

Lemma 7. *Let $U = r0x_m0x_{m-1}\dots 0x_00y \sim \nu\eta_1 \dots \eta_{m+1}$, where each $x_\ell \in \{0, a-b+1\}$ for $\ell \in \{0, \dots, m\}$, and $\nu \in \{r00, r0(a-b)\}$. Then $\varphi(U) = \varphi_1(U) + \varphi_2(U)$.*

Proof. We proceed by induction on m. When $m = 1$, $U \sim \nu\eta_1$. If $\eta_1 \in S$, then $U \sim \nu 0y$ with $0 \leq y \leq a - b$ and $\nu \in \{r00, r0(a-b)\}$. If $\nu \sim r00$, then we have that

$$U \sim r000y \sim (r-1)ab0y \sim (r-1)a(b-1)a(y+b).$$

If $\nu \sim r0(a-b)$, then

$$U \sim r0(a-b)0y \sim (r-1)aa0y \sim (r-1)a(a-1)a(y+b).$$

Therefore, $\varphi(\nu\eta_1) = 3$, and similarly if $\eta_1 \in T$. From the definitions in (3.4), $\varphi_1(\nu\eta_1) = 2$ and $\varphi_2(\nu\eta_1) = 1$ so that $\varphi(\nu\eta_1) = \varphi_1(\nu\eta_1) + \varphi_2(\nu\eta_1)$.

Now suppose that $U \sim \nu\eta_1 \dots \eta_{m+1}$ with $\eta_{m+1} \in S$, and further that $\eta_{m-k+1}, \eta_{m-k+2}, \dots, \eta_{m+1}$ are all contained in the set S, and $\eta_{m-k} \in T$. We define

$$w_1 = \nu_1\eta_1 \dots \eta_m 0 \text{ and } w_2 = \nu_2\eta_1 \dots \eta_{m-k-1}[a(b-1)]^{k+1}a,$$

so that $u_1 = \nu_1\eta_1 \dots \eta_m s \sim U$ and $u_2 = \nu_2\eta_1 \dots \eta_{m-k-1}[a(b-1)]^{k+1}t \sim U$ for the appropriate choice of $s \in S$ and $t \in T$. Note that we can find such a word u_2 by the exchange rule.

Using the inductive hypothesis and definitions in (3.4), we have

$$\begin{aligned}
\varphi(w_1) &= \varphi(\nu\eta_1\eta_2 \dots \eta_m 0) \\
&= \varphi(\nu\eta_1\eta_2 \dots \eta_m) \\
&= \varphi_1(\nu\eta_1\eta_2 \dots \eta_m) + \varphi_2(\nu\eta_1\eta_2 \dots \eta_m) \\
&= \varphi_1(\nu\eta_1\eta_2 \dots \eta_{m+1}) \\
&= \varphi_1(U)
\end{aligned}$$

and

$$\begin{aligned}
\varphi(w_2) &= \varphi(\nu\eta_1 \dots \eta_{m-k-1}[a(b-1)]^{k+1}a) \\
&= \varphi(\nu\eta_1 \dots \eta_{m-k-1}) \\
&= \varphi_1(\nu\eta_1 \dots \eta_{m-k-1}) + \varphi_2(\nu\eta_1 \dots \eta_{m-k-1}) \\
&= \varphi_1(\nu\eta_1 \dots \eta_{m-k}) \\
&= \varphi_2(\nu\eta_1 \dots \eta_{m-k+1}) \\
&= \varphi_2(\nu\eta_1 \dots \eta_{m-k+2}) \\
&\vdots \\
&= \varphi_2(\nu\eta_1 \dots \eta_{m+1}) \\
&= \varphi_2(U).
\end{aligned}$$

Therefore, by Lemma (6), we have that $\varphi(U) = \varphi_1(U) + \varphi_2(U)$. If $\eta_{m+1} \in T$, define words $w_1 = \nu_1\eta_1 \dots \eta_m a$ and $w_2 = \nu_2\eta_1 \dots \eta_{m-k-1}[0(a-b+1)]^{k+1}0$. Then, an analogous proof holds. □

We must consider now the case where for $U_i = r0x_m0x_{m-1} \dots 0x_00y$, there are some $x_i \in \{1, \dots, a-b\}$. Denote such x_i as $y_1, y_2, \dots, y_j$, where $1 \le j \le m$ and rewrite $U_i = rt_jy_jt_{j-1}y_{j-1} \dots y_1t_0y_0$. We now "inflate" U_i with a second copy of the y_k in order to calculate the frequency.

Proposition 3. *Let $U_i = r0x_m0x_{m-1} \dots 0x_00y$ and $\widetilde{U}_i = L_jL_{j-1} \dots L_0$, as in Lemma (3). The frequency $\varphi(U_i)$ is given by*

$$\prod_{l=0}^{j} \varphi(L_\ell),$$

where $\varphi(L_\ell) = \varphi_1(L_\ell) + \varphi_2(L_\ell)$.

Proof. This follows directly from the preceding lemmas. □

Let $e(k,i)$ denote the number of steps in the subtractive Euclidean algorithm applied to the pair (k,i), so that $e(i,i)=0$ and $e(k+i,i)=e(k,i)+1$. We define a labeled complete binary tree (as in [1, 11]) as follows. Following the rules given in (3.4), we start with the root labeled $(1,1)$ at level 0, and for each node labeled (a,b), we label its left successor by $(a+b,a)$ and its right successor by $(a+b,b)$. For level 1, we have one node labeled $(2,1)$. Then at level k, each node is labeled with a pair $(\varphi_1(\nu\eta_1\ldots\eta_k),\varphi_2(\nu\eta_1\ldots\eta_k))$. To arrive at the node corresponding to the frequency of $\varphi(\eta\nu_1\ldots\eta_{k+1})$, to move from level k to level $k+1$, we choose the left node if η_k and η_{k+1} are not both in S or both in T, and we choose the right node if η_k and η_{k+1} are in the same set S or T. We note that by Lemma (7), the frequency of a word U is accurately obtained via a path on this tree when $U=r0x_m0x_{m-1}\ldots0x_00y\sim\nu\eta_1\ldots\eta_{m+1}$, where each $x_\ell\in\{0,a-b+1\}$, and $\nu=r0x_m$ if $x_m\le a-b$ or $r0(x_m-1)$ if $x_m=a-b+1$. Therefore, the classes of words having a representative of this form have a generating function given by the expression

$$a(a-b+1)x^3\sum_{\substack{0<i<k\\ \gcd(k,i)=1}}x^{2e(k-i,i)}=a(a-b+1)\sum_{\substack{0<i<k\\ \gcd(k,i)=1}}x^{1+2e(k,i)}.$$

Let $\mathscr{A}_k=\{(n_0,\ldots,n_j):\prod_{\ell=0}^{j}n_\ell=k,j\ge0,n_\ell>1\}$, and set $\alpha_k(x)$ to be the generating function of all words of frequency k having the form $U=r0x_m0x_{m-1}\ldots0x_00y$ where $r\in\{1,\ldots,a\}$, $x_\ell\in\{0,\ldots,a-b+1\}$ for $0\le\ell\le m$, and $0\le y\le(a-b)$. Now we have in mind $\widetilde{U}=L_jL_{j-1}\ldots L_0$, the "inflated" version of U as in Proposition (3), so that $\varphi(U)=\varphi(\widetilde{U})$. If $\varphi(U)=k$, then $\prod_{\ell=0}^{j}\varphi(L_\ell)=k$, and we can associate to U an element of $\mathscr{A}$ so that $\varphi(L_\ell)=n_\ell$ for each $0\le\ell\le j$. Denote by $\alpha_{(n_0,\ldots,n_j)}(x)$ the part of $\alpha_k(x)$ obtained from the tuple $(n_0,\ldots,n_j)\in\mathscr{A}$. Then by Proposition (3), we have that

$$\begin{aligned}\alpha_{(n_0,\ldots,n_j)}(x)&=\frac{1}{x^j}\sum_{\substack{0<i<n_j\\ \gcd(n_j,i)=1}}ax^{1+2e(n_j,i)}\sum_{\substack{0<i<n_{j-1}\\ \gcd(n_{j-1},i)=1}}(a-b)x^{1+2e(n_{j-1},i)}\ldots\sum_{\substack{0<i<n_0\\ \gcd(n_0,i)=1}}(a-b)(a-b+1)x^{1+2e(n_0,i)}\\&=a(a-b+1)(a-b)^jx\prod_{\ell=0}^{j}\sum_{\substack{0<i<n_\ell\\ \gcd(n_\ell,i)=1}}x^{2e(n_\ell,i)}.\end{aligned}$$

Thus

$$\begin{aligned}\alpha_k(x)&=\sum_{\{j:(n_0,\ldots,n_j)\in\mathscr{A}_k\}}a(a-b+1)(a-b)^jx\prod_{\ell=0}^{j}\sum_{\substack{0<i<n_\ell\\ \gcd(n_\ell,i)=1}}x^{2e(n_\ell,i)}\\&=\sum_{\{j:(n_0,\ldots,n_j)\in\mathscr{A}_k\}}a(a-b+1)(a-b)^jx\prod_{\ell=0}^{j}\widetilde{\alpha}_{n_\ell}(x),\end{aligned}$$

if we define $\widetilde{\alpha}_n(x) = \sum_{\substack{0<i<n\\ \gcd(n,i)=1}} x^{2e(n,i)}$.

We now discuss the generating functions for the v_i in the factorization

$$w = (v_1e_1)u_1(v_2e_2)u_2\dots(v_Je_J)u_Jz.$$

First, we observe that if e_i ends in $a-b+1$, then we may rewrite $e_iu_i \sim e_i'u_i'$ so that e_i' ends in $a-b$ and $u_i' = \eta_1\dots\eta_{n_{k_i}}$, with $\eta_\ell \in S\cup T$ for $1 \le \ell \le k_i$. Consequently, we have an equivalent factorization of w given by $w \sim (v_1e_1')u_1'(v_2e_2')u_2'\dots(v_Je_J')u_J'z$ where $e_i'u_i' = e_iu_i$ if e_i ends in a letter less than $a-b+1$ and is as just described if e_i ends in $a-b+1$.

Now observe that the subwords v_ie_i' are representatives of relational prefix classes. This leads us to the generating functions for the v_i. Since v_1 can be preceded by a string of 0's or a's while the other v_i can not, the generating function for v_1 is given by

$$\frac{P(x)}{a(a-b+1)x^3} = \frac{\gamma(x)}{a(1-x^2)}G(x),$$

while the generating function for $v_i, 2 \le i \le N$, is given by

$$g(x) = \frac{(1-x)P(x)}{(1+x)a(a-b+1)x^3} = \frac{\gamma(x)}{a(1+x)^2}G(x).$$

Taking into account the generating function for v_1, we obtain a refinement of the function f_k as

$$f_k(x) = \frac{\gamma(x)}{a(1-x^2)}G(x)^2\ell_k(x), \text{ for } k \ge 2$$

where $\ell_k(x)$ is the generating function for classes of words of the form $w = U_1v_2U_2\dots v_JU_J$ as in the secondary factorization.

By (3.3), $\ell_k(x)$ satisfies the recurrence

$$\begin{aligned} \ell_k(x) &= \sum_{\substack{d\mid k\\ d\neq 1,k}} \alpha_d(x)\ell_{\frac{k}{d}}(x)g(x) + \alpha_k(x), k \ge 2 \\ \ell_1(x) &= 1. \end{aligned} \tag{3.5}$$

As in [11], we introduce the following two Dirichlet generating functions

$$\begin{aligned} \mathcal{A}(x,s) &= \sum_{k=2}^{\infty} k^s\alpha_k(x)g(x), \\ \mathcal{L}(x,s) &= 1 + \sum_{k=2}^{\infty} k^s\ell_k(x)g(x). \end{aligned} \tag{3.6}$$

Because of (3.5), we have

$$\mathcal{L}(x,s) = \frac{1}{1-\mathcal{A}(x,s)}. \tag{3.7}$$

So that we are able to evaluate $H(x)$, we consider $\frac{\partial\Phi}{\partial s}$.

$$\begin{aligned}
\frac{\partial\Phi}{\partial s}(x,1) &= \frac{\gamma(x)}{a(1-x^2)g(x)}G(x)^2\frac{\partial\mathcal{L}}{\partial s}(x,1) \\
&= \frac{\gamma(x)}{a(1-x^2)g(x)}G(x)^2\frac{1}{(1-\mathcal{A}(x,1))^2}\frac{\partial\mathcal{A}}{\partial s}(x,1) \qquad (3.8)\\
&= \frac{1-ax+(1+a-2b)x^2}{1-x}G(x)^2\mathcal{L}(x,1)^2\frac{\partial\mathcal{A}}{\partial s}(x,1).
\end{aligned}$$

Next, using (3.1), we have

$$\begin{aligned}
\Phi(x,1) = \frac{1}{1-(a+1)x} &= f_1(x) + \sum_{k=2}^{\infty} k f_k(x) \\
&= \frac{1+x}{1-x}G(x) + \frac{\gamma(x)}{a(1-x^2)g(x)}G(x)^2(\mathcal{L}(x,1)-1),
\end{aligned}$$

which gives that

$$G(x)\mathcal{L}(x,1) = \frac{1-x}{(1+x)(1-(a+1)x)}.$$

Inserting this expression into (3.8), we obtain

$$\frac{\partial\Phi}{\partial s}(x,1) = \frac{(1-x)(1-ax+(1+a-2b)x^2)}{(1+x)^2(1-(a+1)x)^2}\frac{\partial\mathcal{A}}{\partial s}(x,1).$$

Now

$$\begin{aligned}
\mathcal{A}(x,s) = \sum_{k=2}^{\infty} k^s\alpha_k(x)g(x) &= \frac{a(a-b+1)xg(x)}{a-b}\sum_{j=1}^{\infty}\left(\sum_{n=2}^{\infty}(a-b)n^s\widetilde{\alpha}_n(x)\right)^j \\
&= a(a-b+1)xg(x)\sum_{N=2}^{\infty}N^s\widetilde{\alpha}_N(x)\frac{1}{1-\displaystyle\sum_{n=2}^{\infty}(a-b)n^s\widetilde{\alpha}_n(x)},
\end{aligned}$$

so that

$$\frac{\partial\mathcal{A}}{\partial s}(x,1) = \frac{a(a-b+1)xg(x)}{\left(1-\displaystyle\sum_{n=2}^{\infty}(a-b)n\widetilde{\alpha}_n(x)\right)^2}\sum_{N=1}^{\infty}\kappa_N x^{2N}.$$

From [11], we have that $\sum_{k=2}^{\infty} k\widetilde{\alpha}_k(x) = \dfrac{2x^2}{1-3x^2}$, so that

$$\frac{\partial \mathcal{A}}{\partial s}(x,1) = \frac{(a-b+1)x\gamma(x)(1-3x^2)^2}{(1+x)(1-ax+(1+a-2b)x^2)(1-(3+2a-2b)x^2)^2}\sum_{N=1}^{\infty}\kappa_N x^{2N}.$$

Let

$$\widehat{M}(x) = \frac{(a-b+1)(1-x)\gamma(x)(1-3x^2)^2}{(a+1)(1+x)^3(1-(3+2a-2b)x^2)^2}.$$

Then

$$\begin{aligned} H(x) &= \frac{x}{(1-x)^2}\left(\ln(a+1) - \widehat{M}\left(\frac{x}{a+1}\right)\sum_{N=1}^{\infty}\kappa_N\left(\frac{x}{a+1}\right)^{2N}\right) \\ &= \frac{x}{(1-x)^2}\mathcal{T}\left(\frac{x}{a+1}\right). \end{aligned}$$

Since $\mathcal{T}\left(\frac{x}{a+1}\right)$ has radius of convergence greater than 1, the Cauchy Integral Formula and the Residue Theorem give us

$$\begin{aligned} H(A_n) &= \frac{1}{2\pi i}\oint_{|z|=\frac{1}{2}} \frac{H(z)}{z^{n+1}}dz \\ &= \frac{1}{2\pi i}\oint_{|z|=1+\epsilon} \frac{H(z)}{z^{n+1}}dz - \text{Res}\ \left(\frac{H(z)}{z^{n+1}}, z=1\right) \\ &= \frac{1}{2\pi i}\oint_{|z|=1+\epsilon} \frac{H(z)}{z^{n+1}}dz + n\mathcal{T}\left(\frac{1}{a+1}\right) - \frac{1}{a+1}\mathcal{T}'\left(\frac{1}{a+1}\right). \\ &= n\mathcal{T}\left(\frac{1}{a+1}\right) + \mathcal{O}(1) \end{aligned}$$

Therefore,

$$H_\beta = \lim_{N\to\infty}\frac{H(A_N)}{N\ln\beta} = \frac{1}{\ln\beta}\mathcal{T}\left(\frac{1}{a+1}\right).$$

This completes the proof of Theorem 1.

4. Computations and Bounds

Though we now have a formula for computing H_β, the series $\sum_{N=1}^{\infty}\kappa_N(a+1)^{-2N}$ converges too slowly for efficient computation. We make note here that the definition of κ_n in this paper differs slightly from the definitions in both [1] and [11] by a factor of $\ln 2$, and so the statements regarding the results from these papers have been adjusted accordingly. In [1], Alexander and Zagier show that $2\cdot 3^{N-1}\ln(N+1) < \kappa_N < 2\cdot 3^{N-1}N\ln\phi$, where ϕ is the golden ratio. However, by rearranging the series, they put useful bounds on the terms of the series. They show that for $\mu_n = \frac{1}{2}(\kappa_{n+1} - 3\kappa_n)$,

$$\ln\frac{3}{2} < \frac{\mu_n}{3^{n-1}} < \frac{2}{3}.$$

In [11], Grabner, Kirschenhofer, and Tichy give a different rearrangement that produces sharper bounds. In this rearrangement, a factor of 3^n is eliminated, producing a series that converges much faster. We will use the arrangement of the series given by Grabner, Kirschenhofer, and Tichy to give more precise estimates for $H_{\beta_{(a,b)}}$.

Let $\nu_1 = \kappa_1 = 2\ln 2, \nu_2 = \kappa_2 - 6\kappa_1 = 6\ln 3 - 12\ln 2$, and set $\nu_{n+2} = 9\kappa_n - 6\kappa_{n+1} + \kappa_{n+2}$, for $n \geq 1$. Then the terms of this sequence are the coefficients of the function $(1-3x)^2 \sum_{n=1}^{\infty} \kappa_n x^n$. In [11], it is shown that the ν_n can be bounded by

$$-0.00104665\ldots = (-0.00151\ldots)(\ln 2) \leq \nu_n \leq \frac{2}{15} = 0.1333\ldots, \text{ for } n \geq 3. \tag{4.1}$$

Using the rearrangement, we have that

$$\begin{aligned} H_{\beta_{(a,b)}} \ln \beta_{(a,b)} &= \ln(a+1) - \widehat{M}\left((a+1)^{-1}\right) \sum_{N=1}^{\infty} \kappa_N (a+1)^{-2N} \\ &= \ln(a+1) - \widetilde{M}\left((a+1)^{-1}\right) \sum_{N=1}^{\infty} \nu_N (a+1)^{-2N}, \end{aligned} \tag{4.2}$$

where

$$\widetilde{M}(x) = \frac{(a-b+1)(1-x)\gamma(x)}{(a+1)(1+x^3)(1-(3+2a-2b)x^2)^2}.$$

If (4.2) is truncated after n terms, the error E_n can be bounded using (4.1). We have

$$-0.00151(\ln 2)\frac{\widetilde{M}((a+1)^{-1})(a+1)^{-2(n+1)}}{1-(a+1)^{-2}} \leq E_n \leq \frac{2}{15}\frac{\widetilde{M}((a+1)^{-1})(a+1)^{-2(n+1)}}{1-(a+1)^{-2}}.$$

By computing 21 values of the coefficients κ_N, the following numerical values for $H_{\beta_{(a,b)}}$ are obtained. Since the error is controlled, the digits obtained are exact.

a	b	$H_{\beta_{(a,b)}}$	a	b	$H_{\beta_{(a,b)}}$
2	1	0.907239671946427	7	1	0.985758606618629
3	1	0.954793492781010	7	2	0.989976326292009
3	2	0.997142593457004	7	3	0.993150242017613
4	1	0.971203731303039	7	4	0.995514201769271
4	2	0.990977711532539	7	5	0.997251333028757
4	3	0.990977711532539	7	6	0.998505846704918
5	1	0.978770306059569	8	1	0.987684168154986
5	2	0.989445577788253	8	2	0.990626303869694
5	3	0.995467621876660	8	3	0.993001403430024
5	4	0.998564900299892	8	4	0.994911765281798
6	1	0.983023920211140	8	5	0.996441971239801
6	2	0.989461539079558	8	6	0.997662188044302
6	3	0.993808941080083	8	7	0.998630796133262
6	4	0.996661307600303			
6	5	0.998448622874396			

5. A Generalization and Final Remarks

All results in this paper may be generalized in a straightforward way to *confluent* numeration systems. Since the proofs and calculations are very similar to those given thus far in the paper, only some key results and formulas will be given in this section. Introduced and studied by Frougny in [9], she shows that confluent numeration systems are precisely those with a sequence base given by the linear recurrence

$$\begin{aligned} G_{n+m} &= aG_{n+m-1} + \cdots + aG_{n+1} + bG_n \text{ for } n \geq 0, \\ G_0 &= 1, G_i = (a+1)^i, \text{ for } 1 \leq i \leq m-1, \text{ where } a, b \in \mathbb{N}, a \geq b. \end{aligned} \tag{5.1}$$

It is clear that the dominant root of the characteristic polynomial, call it $\beta_{m,a,b}$, for this recurrence is a Pisot number, see [3]. The sets S_m and T_m, containing the forbidden subwords, are $S_m = \{0^{m-1}0, 0^{m-1}1, \ldots, 0^{m-1}(a-b)\}$ and $T_m = \{a^{m-1}a, a^{m-1}(a-1), \ldots, a^{m-1}b\}$. As before, suppose a word w is the representation of an integer n in base G_k, defined by recurrence (5.1). If we let $s \in S_m$, and x be such that $1 \leq x \leq a$, then any occurrence of a subword of the form xs in w may be replaced by the word $(x-1)t$, for some $t \in T_m$, to obtain an equivalent representation of n, and vice versa.

In the principal factorization of a greedy representation W of a positive integer n, all points remain the same except for the form of the U_i. Each U_i is of the form

$$U_i = r0^{m-1}x_k 0^{m-1}x_{k-1} \ldots 0^{m-1}x_0 0^{m-1}y$$

with $1 \leq r \leq a, 0 \leq y \leq a-b$, and $x_i \in \{0, 1, \ldots, a-b+1\}$.

The following is a restatement of Proposition 1 in this general setting. When $a = b = 1$, Theorem 2.6 of Kocábová, Masáková, and Pelantová in [14] is recovered, where a formula for the number of representations of an integer in the m-bonacci base is given.

Proposition 4. *Let $U = r0^{d_1}x_1 0^{d_2}x_2 \dots x_k 0^{d_k-1}y$, where $x_i \in \{0, a-b+1\}$ and $0 \le y \le a-b$. Then*

$$R(U) = \begin{pmatrix} 1 & 1 \end{pmatrix} \left(\prod_{i=1}^{k} \begin{pmatrix} \left\lceil \frac{d_j+1}{m} \right\rceil & \left\lceil \frac{d_j}{m} \right\rceil \\ 1 & 1 \end{pmatrix} \right) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The next proposition is a generalization of Proposition 2. When $a = b = 1$, we recover Corollary 1 in [6] where the numeration system is m-bonacci.

Proposition 5. *Let $U = r0^{m-1}x_k 0^{m-1}x_{k-1} \dots 0^{m-1}x_0 0^{m-1}y$ where $1 \le r \le a, x_i \in \{0, a-b+1\}$ for $0 \le i \le k$, and $0 \le y \le a-b$. Further, let $s = 1 \cdot 2^{k+1} + x'_k \cdot 2^k + \dots + x'_1 \cdot 2 + x'_0$. Then*

$$R(U) = \sum_{j=0}^{s} \left[\binom{2s-j}{j} \pmod 2 \right].$$

We now define the sets $A_{m,N} = \{x | x = \sum_{n=1}^{N} a_n \beta^- n\}$, and the measures $\mu_{m,N} = (a+1)^{-N} \sum_{x \in A_{m,N}} r(x)\delta_x$, where $r(x)$ is the number of representations of x of length N in base $\beta_{m,a,b}$. For calculating the Garsia entropy in this generalized setting, the first key generating function is $G_m(x)$, the generating function for the ST_m-free words. It is given by

$$G_m(x) = \frac{x^m - 1}{(a-2b+1)x^{m+1} + (2b-2a-1)x^m + (a+1)x - 1}.$$

Next, we give the generating function, $P_m(x)$, of the analogous relational prefix classes. Here it is assumed, as before, that $a > b$. Let

$$\begin{aligned}\gamma_m(x) = a &+ (4 + 8a + 4a^2 - 8b - 8ab + 4b^2)x^{2m-1} \\ &- (2 + 6a + 2a^2 - 2b - 2ab)x^m + (2 + 4a + 2a^2 - 2b - 2ab)x^{m+1} \\ &- (6 + 11a + 6a^2 - 14b - 14ab + 8b^2)x^{2m} + (2 + 4a + 2a^2 - 6b - 6ab + 4b^2)x^{2m+1}.\end{aligned}$$

Then

$$P_m(x) = \frac{(a-b+1)x^{m+1}(1+x)\gamma_m(x)}{(1-x)(1-x^m)^2} G_m(x).$$

Using $P_m(x)$, we obtain a refinement of f_k, which is defined in the same manner as before, since the definition of $F_N(k)$ does not depend on m. We have that

$$f_k(x) = \frac{(1+x)\gamma_m(x)}{a(1-x)(1-x^m)^2} G_m(x)^2 \ell_k(x), \text{ for } k \ge 2.$$

Additionally, we have from $P_m(x)$ the function $g_m(x)$ which is the generating function for the $v_i, (i \geq 2)$ in the secondary factorization. We have that

$$g_m(x) = \frac{\gamma_m(x)}{a(1-x^m)^2} G_m(x).$$

By similar methods to those used to find $\frac{\partial \Phi}{\partial s}(x,1)$, we may obtain an analogous result for $\frac{\partial \Phi_m}{\partial s}(x,1)$. Then

$$\frac{\partial \Phi_m}{\partial s}(x,1) = M(x) \sum_{N=1}^{\infty} \kappa_N x^{mN},$$

where

$$M(x) = \frac{(a-b+1)x(1-x)(1-3x^m)^2\gamma_m(x)}{(1+x)(1-(a+1)x)^2(1-x^m)^2(1-(3+2a-2b)x^m)^2}.$$

Let

$$\widehat{M}_m(x) = \frac{(a-b+1)(1-x)(1-3x^m)^2\gamma_m(x)}{(a+1)(1+x)(1-x^m)^2(1-(3+2a-2b)x^m)^2}.$$

Then the generating function

$$H_m(x) = \sum_{N=0}^{\infty} H(A_{m,N})x^N = \frac{x}{(1-x)^2}\left(\ln(a+1) - \widehat{M}\left(\frac{x}{a+1}\right)\sum_{N=1}^{\infty}\kappa_N\left(\frac{x}{a+1}\right)^{mN}\right).$$

Now, let

$$\mathcal{T}_m(x) = \ln(a+1) - \widehat{M}\left(\frac{x}{a+1}\right)\sum_{N=1}^{\infty}\kappa_N\left(\frac{x}{a+1}\right)^{mN}.$$

If we let

$$H(A_{m,n}) = -\sum_{x \in A_{m,n}} p(x)\ln p(x),$$

where $p(x) = \frac{r(x)}{(a+1)^N}$ is the weight assigned to x by $\mu_{m,n}$, and

$$H_{m,\beta} = \lim_{N\to\infty}\frac{H(A_N)}{N\ln\beta_{m,a,b}},$$

then

$$H_{m,\beta} = \frac{1}{\ln\beta_{m,a,b}}\mathcal{T}_m\left(\frac{1}{a+1}\right).$$

Remark 1. A further generalization seems possible using a recurrence of the form

$$G_{n+m} = a_1G_{n+m-1} + a_2G_{n+m-2} + \cdots + a_{m-1}G_{n+1} + a_mG_n \text{ for } n \geq 0,$$

$$G_0 = 1, G_i = \sum_{k=1}^{i} a_kG_{i-k} + 1, \text{ for } 1 \leq i \leq m-1, \text{ where } a_1 \geq a_2 \geq \dots a_m \geq 1.$$

Since the combinatorics are different in the cases $a = b$ and $a > b$, a natural concern in this more general setting would be a combination of these cases, as in the example $a_1 > a_2 = a_3 > a_4$. However, it seems quite possible that the equality $a_2 = a_3$ does not significantly change the counting, much the same as the counting in the generalization (5.1) is very similar as in the case for (1.1).

6. Acknowledgments

The author would like to thank Professor Peter Grabner of Graz University of Technology for the in-depth explanations of his work along with his invaluable comments and suggestions.

References

[1] J.C. Alexander, D.B. Zagier, The entropy of certain infinitely convolved Bernoulli measures, *J. London Math. Soc.*, **44** (1991), p. 121–134.

[2] J. Berstel, An exercise on Fibonacci representations, A tribute to Aldo de Luca, (see http://www.mat.uniroma1.it/mat_cms/pres_docente.php? id_docente=50)*RAIRO, Theor. Inform. Appl.*, **35** (2002), p. 491–498.

[3] A. Brauer, On algebraic equations with all but one root in the interior of the unit circle, *Math. Nachr.*, **4** (1951), p. 250–257.

[4] L. Carlitz, Fibonacci representations, *Fibonacci Quarterly*, **6(4)** (1968), p.193–220.

[5] M. Edson, L.Q. Zamboni, On the representation of numbers in the Fibonacci base, *Theoret. Comput. Sci.,* **326** (2004), p. 241–260.

[6] M. Edson, L.Q. Zamboni, On the number of partitions of an integer in the m-bonacci base. *Ann. Inst. Fourier (Grenoble)* **56** (2006), no. 7, p. 2271–2283.

[7] P. Erdős, On the smoothness properties of a family of Bernoulli convolutions, *Amer. J. Math.*, **92** (1939), p. 974–976.

[8] A.S. Fraenkel, Systems of numeration, *Amer. Math. Monthly*, **92** (1985), p. 105–114.

[9] C. Frougny, Confluent linear numeration systems, *Theoret. Comput. Sci.*, **106(2)** (1992), p. 183–219.

[10] A.M. Garsia, Entropy and singularity of infinite convolutions, *Pacific J. Math.*, **13** (1963), p. 1159–1169.

[11] P.J. Grabner, P. Kirschenhofer, R.F. Tichy, Combinatorial and arithmetical properties of linear numeration systems, *Combinatorica*, **22** (2002), p. 245–267.

[12] L.J. Guibas, A.M.Odlyzko, String overlaps, pattern matching, and nontransitive games, *J. Comb. Theory, Ser. A*, **30** (1981), p. 183–208.

[13] B. Jessen, A. Wintner, Distribution functions and the Riemann zeta function, *Trans. Amer. Math. Soc.*, **38** (1938), p. 48–88.

[14] P. Kocábová, Z. Masáková, and E. Pelantová, Ambiguity in the m-bonacci numeration system, *Discrete Math. Theor. Comput. Sci.* **9** (2007), no. 2, 109–123.

[15] S.P. Lalley, Random series in powers of algebraic integers: Hausdorff dimension of the limit distribution, *J. London Math. Soc.*, **57** (1998), p. 629–654.

[16] M. Lamberger, *Probabilistic and Dynamical Properties of Number Systems* (Doctoral dissertation), Graz University of Technology, 2002.

[17] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, Cambridge, 2005.

[18] Y. Peres, W. Schlag, B. Solomyak, Sixty years of Bernoulli convolutions, *Fractal Geometry and Stochastics, II*, *Progr. Probab.* **46** (1999) p. 39–65.

[19] B. Solomyak, On the random series $\sum \pm\lambda^n$ (an Erdős problem), *Ann. of Math. (2)*, **142(3)** (1995), p. 611–625.

Department of Mathematics and Statistics, Murray State University, Murray, KY 42071, USA

E-mail address: marcia.edson@murraystate.edu

Institute of Analysis and Computational Number Theory (Math A), Graz University of Technology, Steyrergasse 30/II, 8010 Graz, Austria (current address)

E-mail address: edson@math.tugraz.at