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 \( 2 \times 2 \) matrices and the other sums of binomial coefficients modulo 2. For the main result, we consider the limiting measure \( \mu_\beta \) of a convergent infinite convolution of measures (Bernoulli convolutions), where \( \beta \) 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 \( \mu_\beta \). 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
\[
G_{n+2} = aG_{n+1} + bG_n \quad \text{for } n \geq 0,
\]
(1.1)
\[
G_0 = 1, \quad G_1 = a + 1 \text{ where } a, b \in \mathbb{N}, \quad a \geq b.
\]

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,
\[
n = \sum_{i=0}^{k} d_i G_i
\]
(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 \( d_kd_{k-1}\ldots d_0 \) 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 \( n = d_kG_k + n_k \), where \( 0 \leq n_k < G_k \). Generally, let \( n_{i+1} = d_iG_i + n_i \),

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”.

\( \quad \)}
Example 1. Let \(G_1 = 1, G_2 = 6, G_3 = 32, G_4 = 172, G_5 = 924, G_6 = 4964, G_7 = 26668 \ldots \). The set \(S = \{00, 01, \ldots, 0(a-b)\}\) 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.

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

\[
\Omega_{(a,b)}(n) = \{ w = w_0w_1 \ldots w_k \in A^* : w_0 > 0 \text{ and } n = \sum_{i=0}^{k} w_i G_{k-i} \}.
\]

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, \(R_{(a,b)}(n) = R^+_{(a,b)}(n) + R^-_{(a,b)}(n)\). 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

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

is a Pisot number. This follows directly from a result of A. Brauer in [3] as $a \geq b \geq 1$. We simply write $\beta$ 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 $A_N = \{x \mid x = \sum_{n=1}^{N} a_n \beta^{-n}\}$, 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 $\beta$ and $\delta_x$ denotes the unit point mass at $x$. Then these measures converge weakly to a measure $\mu_\beta$. Jessen and Wintner [13] show that any convergent infinite convolution is either purely singular or absolutely continuous. In particular, we have that the measures $\mu_\beta$ are either purely singular or absolutely continuous.

In [7], Erdős proved that for $\beta = \frac{1 + \sqrt{5}}{2}$, $\mu_\beta$ is purely singular. For further results, we refer to [15, 18, 19]. Garsia in [10], in order to study the measures $\mu_\beta$ 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 $\mu_n$. Then set

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

\[
G_{n+m} = aG_{n+m-1} + \cdots + aG_{n+1} + aG_n \text{ for } n \geq 0,
\]

\[
G_0 = 1, G_i = (a + 1)^i, \text{ for } 1 \leq i \leq m - 1, \text{ where } a \in \mathbb{N}.
\]

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

\[
\kappa_n = \sum_{0 < i < k \atop \gcd(k,i)=1} k \ln k \quad \text{and} \quad \tilde{\alpha}_n(x) = \sum_{0 < i < n \atop \gcd(n,i)=1} x^{e(n,i)}.
\]

Furthermore, let

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

where

\[
\hat{M}(x) = \frac{(a - b + 1)(1 - x)\gamma(x)(1 - 3x^2)^2}{(a + 1)(1 + x)^3(1 - (3 + 2a - 2b)x^2)^2},
\]

and

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

\[ H_{\beta(a,b)} = \frac{1}{\ln \beta(a,b)} T \left( \frac{1}{a+1} \right). \]

## 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 \textit{principal factorization} of \( W \), as in [6]. We may write

\[ W = V_1 U_1 V_2 U_2 \ldots V_J U_J Z \]

where

- \( V_1, V_2, \ldots, V_J, Z \) 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

\[ U_i = r0x_k0x_{k-1} \ldots 0x_00y \]

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

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 \( vx0r0x_k0x_{k-1} \ldots 0x_00y \) 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 \( W = 4341002451110300112121212 \) is the greedy representation of some positive integer in the numeration system generated by the pair \((5, 2)\). Note that \( W = (434)(100)(24511)(10300)(112121212) = V_1 U_1 V_2 U_2 Z \). 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

\[
\begin{align*}
(2.1) \quad & \begin{pmatrix} R^- (r_10^\ell w) \\ R^+ (r_10^\ell w) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R^- (r_20^{\ell-2}w) \\ R^+ (r_20^{\ell-2}w) \end{pmatrix} \quad \text{for } \ell \geq 3,
\end{align*}
\]

\[
\begin{align*}
(2.2) \quad & \begin{pmatrix} R^- (r_10w) \\ R^+ (r_10w) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} R^- (w) \\ R^+ (w) \end{pmatrix},
\end{align*}
\]

\[
\begin{align*}
(2.3) \quad & \begin{pmatrix} R_- (r_10w) \\ R_+ (r_10w) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} R^- (w) \\ R^+ (w) \end{pmatrix}.
\end{align*}
\]

Proof. Consider equation (2.1). Since \( R^- (r_10^\ell w) \) counts the number of representations of the form \( (r_1-1)ab0^{\ell-2}w \), \( R^- (r_10^\ell w) = R(b0^{\ell-2}w) = R(r_20^{\ell-2}w) = R^- (r_20^{\ell-2}w) + R^+ (r_20^{\ell-2}w) \). And, since \( R^+ (r_10^\ell w) \) and \( R^+ (r_20^{\ell-2}w) \) count the number of representations fixing \( r_1 \) and \( r_2 \), respectively, \( R^+ (r_10^\ell w) = R(w) = R^+ (r_20^{\ell-2}w) \).

Similar arguments hold for the remaining equations. \( \square \)

Using the identities

\[
\begin{align*}
\left( \begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array} \right)^{d-1} \left( \begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array} \right) &= \left( \begin{array}{cc} d & d \\ 1 & 1 \end{array} \right), \\
\left( \begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array} \right)^d \left( \begin{array}{cc} 1 & 0 \\ 1 & 1 \end{array} \right) &= \left( \begin{array}{cc} d+1 & d \\ 1 & 1 \end{array} \right)
\end{align*}
\]

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

\[
\left( \begin{array}{cc} R^-(U) \\ R^+(U) \end{array} \right) = \left( \begin{array}{cc} \lceil \frac{\ell}{2} \rceil & \left\lfloor \frac{\ell}{2} \right\rfloor \\ 1 & 1 \end{array} \right) \begin{pmatrix} R^-(w) \\ R^+(w) \end{pmatrix} \quad \text{for } \ell \geq 1.
\]

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

**Proposition 1.** Let \( U = r0^{d_1}x_10^{d_2}x_2 \ldots x_k0^{d_k-1}y \), where \( x_i \in \{0, a-b+1\} \) and \( 0 \leq y \leq a-b \). Then

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

We consider now the more general case where for \( U = r0^{d_1}x_10^{d_2}x_2 \ldots x_k0^{d_k-1}y \), some \( x_i \in \{1, \ldots, a-b\} \). Denote such \( x_i \) as \( y_1, y_2, \ldots, y_j \) with \( 1 \leq j \leq m \), and rewrite

\[
U = y_{j+1}t_jy_jy_{j-1}y_{j-1} \ldots y_1t_0y_0
\]

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 \( \tilde{U} = (rt_jy_j)(y_jt_{j-1}y_{j-1}) \ldots (y_1t_0y_0) = L_jL_{j-1} \ldots L_0 \). 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 $U = r_0 x_k 0 x_{k-1} \ldots 0 x_0 y 0$ with $1 \leq r \leq a, 0 \leq y \leq a - b$, and $x_i \in \{0, 1, \ldots, a - b + 1\}$, 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 $\tilde{U} = (rt_jy_j)(y_jt_{j-1}y_{j-1}) \ldots (y_1t_0y_0) = L_jL_{j-1}\ldots L_0$. Then $R(\tilde{U}) = R(U)$.

**Proof.** We begin with the observation that since each $y_k \leq (a - b)$ for $0 \leq k \leq j$, 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 $y_k \geq 1$, 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 $U = r_0 x_k 0 x_{k-1} \ldots 0 x_0 y$ where $r = a - b + 1, x_i \in \{0, a - b + 1\}$ for $0 \leq i \leq k$, and $0 \leq y \leq a - b$. Suppose that $1 \leq z \leq a$. Then

$$R^+(z0r0x_k0x_{k-1}\ldots0x_0y) = R(U) = R^+(U) + R^-(U)$$

$$R^-(z0r0x_k0x_{k-1}\ldots0x_0y) = R^-(U)$$

$$R^+(z00x_k0x_{k-1}\ldots0x_0y) = R^+(U)$$

$$R^-(z00x_k0x_{k-1}\ldots0x_0y) = R(U) = R^+(U) + R^-(U)$$

**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 $U = r_0 x_k 0 x_{k-1} \ldots 0 x_0 y$ where $1 \leq r \leq a, x_i \in \{0, a - b + 1\}$ for $0 \leq i \leq k$, and $0 \leq y \leq a - b$. We construct a tower of $k + 2$ levels $L_0, L_1, \ldots, L_{k+1}$, 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 $s = 1 \cdot 2^{k+1} + x'_k \cdot 2^k + \cdots + x'_1 \cdot 2 + x'_0$. 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 $U = r_0 x_0 x_k 0 x_{k-1} \ldots 0 x_0 y$ 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 $U = r_0 x_0 x_k 0 x_{k-1} \ldots 0 x_0 y$ 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 $\tilde{U} = L_1 L_{k-1} \ldots L_0$ be the inflated form of $U$, so that each $L_i$ is of the form $L_i = r_i 0 x_k,0 x_{k-1,i} \ldots 0 x_0,0 y_i$. If $s_i = 1 \cdot 2^{k+1} + x'_k,i \cdot 2^k + \cdots + x'_1,i \cdot 2 + x'_0,i$, then

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

**Proof.** The proof follows directly from Proposition 2 and Lemma 1. \hfill \Box

Given a positive integer $n$, with principal factorization $W = V_1 U_1 V_2 U_2 \ldots V_j U_j Z$, we may calculate $R(n)$ as follows. Let $\tilde{U}_i = L_i, L_{i,j-1} \ldots L_{i_0}$ 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

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

Then we have,

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

and

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

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

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

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 $1 \leq x \leq a$ 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), $1 \leq x \leq a$, and $0 \leq y \leq a - b$, 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

$$P(x) = \frac{1 + x}{1 - x} \left[ G(x)x^3a(a - b + 1) - x^3a(a - b + 1)^2(G_0(x) + G_a(x)) + \frac{2x^2(a - b + 1)^2}{x + 1} \right]$$

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

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

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

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

$$w = v_1e_1u_1v_2e_2u_2 \ldots v_fe_fu_f,$$

where

- each $v_i$ is $ST$-free $(1 < i \leq J)$,
- each $e_i$ is of the form $r0x$ where $r \in \{1, \ldots, a\}, x \in \{0, \ldots, a - b + 1\}$,
- $v_Jr$ 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 $r0x_0x_{m-1} \ldots 0x_00y$ where $r \in \{1, \ldots, a\}, x_\ell \in \{0, \ldots, a - b + 1\}$ for $0 \leq \ell \leq m$, and $0 \leq y \leq (a - b)$.
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 \(e_iu_i\) as the \(v_i\) contribute nothing to the frequency of \([w]\).

So, if we set \(U_i = e_iu_i\), we have that

\[
w = v_1U_1v_2U_2 \ldots v_JU_J
\]

and

\[
(3.3) \quad \varphi(w) = \prod_{i=1}^{N} \varphi(U_i).
\]

We now focus on the frequency \(\varphi(U_i)\) where the word \(U_i = r0x_m0x_{m-1} \ldots 0x_0y\) 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 \ldots \eta_{m+1} \quad \text{where} \quad \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.

\[

\varphi_1(\nu \eta_1 \ldots \eta_{m+1}) = \varphi_1(\nu \eta_1 \ldots \eta_m) + \varphi_2(\nu \eta_1 \ldots \eta_m),
\]

\[

(3.4) \quad \varphi_2(\nu \eta_1 \ldots \eta_{m+1}) = \begin{cases} 
\varphi_2(\nu \eta_1 \ldots \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 \ldots \eta_m) & \text{otherwise},
\end{cases}
\]

\[
\varphi_1(\nu) = 1 = \varphi_2(\nu)
\]

where \(\nu \in \{rs : s \in S\}\).

**Lemma 6.** Suppose \(U = r0x_m0x_{m-1} \ldots 0x_0y\) such that \(U \sim u_1s \sim u_2t, s \in S, t \in T\). Then, \(\varphi(U) = \varphi(u_10) + \varphi(u_2a)\).

**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 \(T\). \(\square\)

**Lemma 7.** Let \(U = r0x_m0x_{m-1} \ldots 0x_0y \sim \nu \eta_1 \ldots \eta_{m+1}\), where each \(x_\ell \in \{0, a-b+1\}\) for \(\ell \in \{0, \ldots, m\}\), 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

\[
U \sim r000y \sim (r-1)a0y \sim (r-1)a(b-1)a(y+b).
\]

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

\[
U \sim r0(a-b)0y \sim (r-1)a00y \sim (r-1)a(a-1)a(y+b).
\]

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 \( U \sim \nu \eta_1 \ldots \eta_{m+1} \) with \( \eta_{m+1} \in S \), and further that \( \eta_{m-k+1}, \eta_{m-k+2}, \ldots, \eta_{m+1} \) are all contained in the set \( S \), and \( \eta_{m-k} \in T \). We define
\[
w_1 = \nu_1 \eta_1 \ldots \eta_m 0 \text{ and } w_2 = \nu_2 \eta_1 \ldots \eta_{m-k-1} [a(b-1)]^{k+1} a,
\]
so that \( u_1 = \nu_1 \eta_1 \ldots \eta_m s \sim U \) and \( u_2 = \nu_2 \eta_1 \ldots \eta_{m-k-1} [a(b-1)]^{k+1} t \sim U \) 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
\[
\varphi(w_1) = \varphi(\nu \eta_1 \eta_2 \ldots \eta_m 0) = \varphi(\nu \eta_1 \eta_2 \ldots \eta_m) = \varphi_1(\nu \eta_1 \eta_2 \ldots \eta_m) + \varphi_2(\nu \eta_1 \eta_2 \ldots \eta_m) = \varphi_1(\nu \eta_1 \eta_2 \ldots \eta_{m+1}) = \varphi_1(U)
\]
and
\[
\varphi(w_2) = \varphi(\nu \eta_1 \ldots \eta_{m-k-1} [a(b-1)]^{k+1} a) = \varphi(\nu \eta_1 \ldots \eta_{m-k-1}) = \varphi_1(\nu \eta_1 \ldots \eta_{m-k-1}) + \varphi_2(\nu \eta_1 \ldots \eta_{m-k-1}) = \varphi_1(\nu \eta_1 \ldots \eta_{m-k}) = \varphi_2(\nu \eta_1 \ldots \eta_{m-k+1}) = \varphi_2(\nu \eta_1 \ldots \eta_{m-k+2})
\]
\[
\vdots
\]
\[
= \varphi_2(\nu \eta_1 \ldots \eta_{m+1})
\]
\[
= \varphi_2(U).
\]

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

We must consider now the case where for \( U_i = r_0 x_m 0 x_{m-1} \ldots 0 x_0 y \), there are some \( x_i \in \{1, \ldots, a-b\} \). Denote such \( x_i \) as \( y_1, y_2, \ldots, y_j \), where \( 1 \leq j \leq m \) and rewrite \( U_i = r t_{j} y_{j} t_{j-1} y_{j-1} \ldots y_1 t_0 y_0 \). We now “inflate” \( U_i \) with a second copy of the \( y_k \) in order to calculate the frequency.

**Proposition 3.** Let \( U_i = r_0 x_m 0 x_{m-1} \ldots 0 x_0 y \) and \( \tilde{U}_i = L_j L_{j-1} \ldots L_0 \), as in Lemma (3). The frequency \( \varphi(U_i) \) is given by
\[
\prod_{l=0}^{j} \varphi(L_l),
\]
where \( \varphi(L_l) = \varphi_1(L_l) + \varphi_2(L_l) \).

**Proof.** This follows directly from the preceding lemmas. \( \square \)
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 \((\phi_1(\nu_1 \ldots \nu_k), \phi_2(\nu_1 \ldots \nu_k))\). To arrive at the node corresponding to the frequency of \(\phi(\nu_1 \ldots \nu_{k+1})\), to move from level \(k\) to level \(k + 1\), we choose the left node if \(\eta_k\) and \(\eta_{k+1}\) are not both in \(S\) or both in \(T\), and we choose the right node if \(\eta_k\) and \(\eta_{k+1}\) 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 \(U = r0x_m0x_{m-1}\ldots0x_0y \sim \nu_1 \ldots \nu_m\), where each \(x_\ell \in \{0, a - b + 1\}\), and \(\nu = r0x_m\) if \(x_m \leq a - b\) 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

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

Let \( \mathcal{A}_k = \{(n_0, \ldots, n_j) : \prod_{\ell=0}^j n_\ell = k, j \geq 0, n_\ell > 1\} \), and set \(\alpha_k(x)\) to be the generating function of all words of frequency \(k\) having the form \(U = r0x_m0x_{m-1}\ldots0x_0y\) where \(r \in \{1, \ldots, a\}\), \(x_\ell \in \{0, \ldots, a - b + 1\}\) for \(0 \leq \ell \leq m\), and \(0 \leq y \leq (a - b)\). Now we have in mind \(\tilde{U} = L_jL_{j-1}\ldots L_0\), the “inflated” version of \(U\) as in Proposition (3), so that \(\phi(U) = \phi(\tilde{U})\).

If \(\phi(U) = k\), then \(\prod_{\ell=0}^j \phi(L_\ell) = k\), and we can associate to \(U\) an element of \(\mathcal{A}\) so that \(\phi(L_\ell) = n_\ell\) for each \(0 \leq \ell \leq j\). Denote by \(\alpha_{(n_0, \ldots, n_j)}(x)\) the part of \(\alpha_k(x)\) obtained from the tuple \((n_0, \ldots, n_j) \in \mathcal{A}\). Then by Proposition (3), we have that

\[
\alpha_{(n_0, \ldots, n_j)}(x) = \frac{1}{x^j} \sum_{0 < i < n_j \atop \gcd(n_j, i) = 1} ax^{1+2e(n_i, i)} \sum_{0 < i < n_{j-1} \atop \gcd(n_{j-1}, i) = 1} (a - b)x^{1+2e(n_{j-1}, i)} \ldots \sum_{0 < i < n_0 \atop \gcd(n_0, i) = 1} (a - b)(a - b + 1)x^{1+2e(n_0, i)}
\]

\[
= a(a - b + 1)(a - b)^j x \prod_{\ell=0}^j \sum_{0 < i < n_\ell \atop \gcd(n_\ell, i) = 1} x^{2e(n_\ell, i)}. \]

Thus

\[
\alpha_k(x) = \sum_{\{j: (n_0, \ldots, n_j) \in \mathcal{A}_k\}} a(a - b + 1)(a - b)^j x \prod_{\ell=0}^j \sum_{0 < i < n_\ell \atop \gcd(n_\ell, i) = 1} x^{2e(n_\ell, i)}
\]

\[
= \sum_{\{j: (n_0, \ldots, n_j) \in \mathcal{A}_k\}} a(a - b + 1)(a - b)^j x \prod_{\ell=0}^j \tilde{\alpha}_{n_\ell}(x),
\]
if we define \( \tilde{\alpha}_n(x) = \sum_{0 < i < n} x^{2e(n,i)} \).

We now discuss the generating functions for the \( v_i \) in the factorization

\[
 w = (v_1 e_1)u_1(v_2 e_2)u_2 \ldots (v_J e_J)u_J z.
\]

First, we observe that if \( e_i \) ends in \( a - b + 1 \), then we may rewrite \( e_i u_i \sim e'_i u'_i \) so that \( e'_i \) ends in \( a - b \) and \( u'_i = \eta_1 \ldots \eta_k \), with \( \eta_\ell \in S \cup T \) for \( 1 \leq \ell \leq k \). Consequently, we have an equivalent factorization of \( w \) given by \( w \sim (v_1 e'_1)u'_1(v_2 e'_2)u'_2 \ldots (v_J e'_J)u'_J z \) where \( e'_i u'_i = e_i u_i \) 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 \( v_i e'_i \) 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

\[
 P(x) \frac{a(a - b + 1)x^3}{a(1 - x^2)G(x)},
\]

while the generating function for \( v_i, 2 \leq i \leq N \), 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

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

where \( \ell_k(x) \) is the generating function for classes of words of the form \( w = U_1 v_2 U_2 \ldots v_J U_J \) as in the secondary factorization.

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

\[
 \ell_k(x) = \sum_{d \mid k, d \neq 1, k} \alpha_d(x) \ell_{\frac{k}{d}}(x) g(x) + \alpha_k(x), k \geq 2
\]

(3.5)

\[
 \ell_1(x) = 1.
\]

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

\[
 \mathcal{A}(x, s) = \sum_{k=2}^{\infty} k^s \alpha_k(x) g(x),
\]

(3.6)

\[
 \mathcal{L}(x, s) = 1 + \sum_{k=2}^{\infty} k^s \ell_k(x) g(x).
\]

Because of (3.5), we have

\[
 \mathcal{L}(x, s) = \frac{1}{1 - \mathcal{A}(x, s)}.
\]
So that we are able to evaluate \( H(x) \), we consider \( \frac{\partial \Phi}{\partial s} \).

\[
\frac{\partial \Phi}{\partial s} (x, 1) = \frac{\gamma(x)}{a(1 - x^2)g(x)} G(x)^2 \frac{\partial L}{\partial s} (x, 1)
= \frac{\gamma(x)}{a(1 - x^2)g(x)} G(x)^2 \frac{1}{(1 - A(x, 1))^2} \frac{\partial A}{\partial s} (x, 1)
= \frac{1 - ax + (1 + a - 2b)x^2}{1 - x} G(x)^2 L(x, 1) \frac{\partial A}{\partial s} (x, 1).
\]

(3.8)

Next, using (3.1), we have

\[
\Phi(x, 1) = \frac{1}{1 - (a + 1)x} = f_1(x) + \sum_{k=2}^\infty kf_k(x)
= \frac{1 + x}{1 - x} G(x) + \frac{\gamma(x)}{a(1 - x^2)g(x)} G(x)^2 (L(x, 1) - 1),
\]

which gives that

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

Inserting this expression into (3.8), we obtain

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

Now

\[
A(x, s) = \sum_{k=2}^\infty k^s \alpha_k(x)g(x) = \frac{a(a - b + 1) x g(x)}{a - b} \sum_{j=1}^\infty \left( \sum_{n=2}^\infty (a - b)n^s \tilde{\alpha}_n(x) \right)^j
= a(a - b + 1) x g(x) \sum_{N=2}^\infty N^s \tilde{\alpha}_N(x) \frac{1}{1 - \sum_{n=2}^\infty (a - b)n^s \tilde{\alpha}_n(x)},
\]

so that

\[
\frac{\partial A}{\partial s} (x, 1) = \frac{a(a - b + 1) x g(x)}{\left( 1 - \sum_{n=2}^\infty (a - b)n \tilde{\alpha}_n(x) \right)^2} \sum_{N=1}^\infty \kappa_N x^{2N}.
\]
From [11], we have that \( \sum_{k=2}^{\infty} k \tilde{\alpha}_k(x) = \frac{2x^2}{1 - 3x^2} \), so that
\[
\frac{\partial A}{\partial s}(x, 1) = \frac{(a - b + 1)x \gamma(x)(1 - 3x^2)^2}{(1 + x)(1 - ax + (1 + a - 2b)x^2)(1 - (3 + 2a - 2b)x^2)^2} \sum_{N=1}^{\infty} \kappa_N x^{2N}.
\]

Let
\[
\hat{M}(x) = \frac{(a - b + 1)(1 - x) \gamma(x)(1 - 3x^2)^2}{(a + 1)(1 + x)^3(1 - (3 + 2a - 2b)x^2)^2}.
\]

Then
\[
H(x) = \frac{x}{(1 - x)^2} \left( \ln(a + 1) - \hat{M} \left( \frac{x}{a + 1} \right) \sum_{N=1}^{\infty} \kappa_N \left( \frac{x}{a + 1} \right)^{2N} \right)
= \frac{x}{(1 - x)^2} T \left( \frac{x}{a + 1} \right).
\]

Since \( T \left( \frac{x}{a+1} \right) \) has radius of convergence greater than 1, the Cauchy Integral Formula and the Residue Theorem give us
\[
H(A_n) = \frac{1}{2\pi i} \oint_{|z|=\frac{1}{2}} \frac{H(z)}{z^{n+1}} dz
= \frac{1}{2\pi i} \oint_{|z|=1+\epsilon} \frac{H(z)}{z^{n+1}} dz - \text{Res} \left( \frac{H(z)}{z^{n+1}}, z = 1 \right)
= \frac{1}{2\pi i} \oint_{|z|=1+\epsilon} \frac{H(z)}{z^{n+1}} dz + n T \left( \frac{1}{a+1} \right) - \frac{1}{a+1} T' \left( \frac{1}{a+1} \right).
= n T \left( \frac{1}{a+1} \right) + \mathcal{O}(1)
\]

Therefore,
\[
H_\beta = \lim_{N \to \infty} \frac{H(A_N)}{N \ln \beta} = \frac{1}{\ln \beta} T \left( \frac{1}{a+1} \right).
\]

This completes the proof of Theorem 1.

4. Computations and Bounds

Though we now have a formula for computing \( H_\beta \), 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 \( \kappa_n \) 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 \( \phi \) 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 \( \nu_n \) can be bounded by

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

Using the rearrangement, we have that

\[
H_{\beta(a,b)} \ln \beta(a,b) = \ln(a + 1) - \widehat{M}( (a + 1)^{-1} ) \sum_{N=1}^{\infty} \kappa_N (a + 1)^{-2N}
\]

(4.2)

\[
= \ln(a + 1) - \widehat{M}( (a + 1)^{-1} ) \sum_{N=1}^{\infty} \nu_N (a + 1)^{-2N},
\]

where

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

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

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

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

\[ G_{n+m} = aG_{n+m-1} + \cdots + aG_{n+1} + bG_n \quad \text{for } n \geq 0, \]

\[ G_0 = 1, G_i = (a+1)^i, \quad \text{for } 1 \leq i \leq m-1, \quad \text{where } a, b \in \mathbb{N}, a \geq b. \]

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

\[ U_i = r0^{m-1}x_k0^{m-1}x_{k-1} \cdots 0^{m-1}x_00^{m-1}y \]

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

<table>
<thead>
<tr>
<th>( a )</th>
<th>( b )</th>
<th>( H_{\beta_{m,a,b}} )</th>
<th>( a )</th>
<th>( b )</th>
<th>( H_{\beta_{m,a,b}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>0.907239671946427</td>
<td>7</td>
<td>1</td>
<td>0.985758606618629</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.954793492781010</td>
<td>7</td>
<td>2</td>
<td>0.98997632692009</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0.997142593457004</td>
<td>7</td>
<td>3</td>
<td>0.993150242017613</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.971203731303039</td>
<td>7</td>
<td>4</td>
<td>0.995514201769271</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0.990977711532539</td>
<td>7</td>
<td>5</td>
<td>0.997251333028757</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.990977711532539</td>
<td>7</td>
<td>6</td>
<td>0.998505846704918</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>0.978770306059569</td>
<td>8</td>
<td>1</td>
<td>0.98764168154986</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.98944557788253</td>
<td>8</td>
<td>2</td>
<td>0.990626303869694</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0.995467621876660</td>
<td>8</td>
<td>3</td>
<td>0.998001403430024</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>0.998564900299892</td>
<td>8</td>
<td>4</td>
<td>0.994911765281798</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.983023920211140</td>
<td>8</td>
<td>5</td>
<td>0.996441971239801</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>0.989461539079558</td>
<td>8</td>
<td>6</td>
<td>0.997662188044302</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>0.993808941080083</td>
<td>8</td>
<td>7</td>
<td>0.998630796133262</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>0.996661307600303</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>0.998448622874396</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 $U = r^0x_10^d\cdot x_2\ldots x_k0^{d_k}1y$, where $x_i \in \{0, a-b+1\}$ and $0 \leq y \leq a-b$. Then

$$R(U) = (1 \ 1) \left( \prod_{i=1}^{k} \left( \begin{array}{cc} \left[ \frac{d_j+1}{m} \right] \\ 1 \end{array} \right) \right) \left( \begin{array}{c} 0 \\ 1 \end{array} \right).$$

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 $U = r^0x_10^{m-1}x_2\ldots x_k0^{m-1}x_00^{m-1}1y$ 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

$$R(U) = \sum_{j=0}^{s} \left( \begin{array}{c} 2s-j \\ j \end{array} \right) \mod 2.$$ 

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

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

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

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

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

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

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

Let

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

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

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

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

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

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

$$G_{n+m} = a_1 G_{n+m-1} + a_2 G_{n+m-2} + \cdots + a_{m-1} G_{n+1} + a_m G_n$$

for $n \geq 0$,

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

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


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