PAPER *Special Section on Discrete Mathematics and Its Applications*

On the Zeta Function of a Periodic-Finite-Type Shift*

Akiko MANADA[†a)], ***Member and*** **Navin KASHYAP**[††b)], ***Nonmember***

SUMMARY Periodic-finite-type shifts (PFT's) are sofic shifts which forbid the appearance of finitely many pre-specified words in a periodic manner. The class of PFT's strictly includes the class of shifts of finite type (SFT's). The zeta function of a PFT is a generating function for the number of periodic sequences in the shift. For a general sofic shift, there exists a formula, attributed to Manning and Bowen, which computes the zeta function of the shift from certain auxiliary graphs constructed from a presentation of the shift. In this paper, we derive an interesting alternative formula computable from certain "word-based graphs" constructed from the periodically-forbidden word description of the PFT. The advantages of our formula over the Manning-Bowen formula are discussed.

key words: *periodic-finite-type shift, zeta function, word-based graph, Möbius inversion formula*

1. Introduction

A sofic shift is a set of bi-infinite sequences which can be represented by some labeled directed graph, and is core to the study of constrained coding (see, for example, [6]). A classic example of sofic shifts is the class of shifts of finite type (SFT's), which arise commonly in the context of coding for data storage devices, such as CD's and DVD's.

A new class of sofic shifts, called *periodic-finite-type shifts* (PFT's), was introduced by Moision and Siegel [9], who were interested in studying the properties of distance-enhancing codes, in which the appearance of certain words is forbidden in a periodic manner. The class of PFT's strictly includes the class of SFT's, and some other interesting classes of shifts, such as constrained systems with unconstrained positions [11], and shifts arising from the time-varying maximum transition run constraint [10].

The difference between the definitions of SFT's and PFT's is small, but significant. An SFT is defined by forbidding the appearance of finitely many words at any position of a bi-infinite sequence. A PFT is also defined by forbidding the appearance of finitely many words within a bi-infinite sequence, except that these words are only forbidden to appear at positions indexed by certain pre-defined *periodic* integer sequences; see Section 2 for a formal definition. Thus, there is a notion of period inherent in the definition of a PFT that causes it to differ from an SFT.

The properties of SFT's are quite well understood (see, for example, [6]), but the same cannot be said for PFT's. The study of PFT's has, up to this point, primarily focused on finding efficient algorithms for constructing their presentations [1],[2],[5]. The work presented in this paper may be viewed as part of an effort (see also [7]) to extend some of what is known about SFT's to the larger class of PFT's.

This paper focuses on zeta functions. The zeta function of a sofic shift is a generating function for the number of periodic sequences in the shift. That is, the zeta function is an invariant which contains all information on periodic sequences. It is known that the zeta function of a sofic shift is always a rational function [8]. Indeed, the zeta function of a sofic shift can be explicitly computed from labeled directed graphs derived from a graph presenting the shift; see [6, Theorem 6.4.8].

It is well known that the zeta function can be computed in a much simpler way when the sofic shift is in fact an SFT. In this case, the zeta function is obtainable from the characteristic polynomial of only one matrix — the adjacency matrix of a graph derivable from a forbidden-word description of the SFT; see [6, Theorem 6.4.6]. In this paper, we prove an analogous result (Theorem 4.1) for a PFT. We show that the zeta function of a PFT can be computed from certain matrices directly derivable from a description of the PFT in terms of periodically-forbidden words. Moreover, the number of these matrices depends only on the period of the PFT. For example, the number of matrices needed is two when the PFT has period equal to 2.

The rest of this paper is organized as follows. We provide some of the necessary background on PFT's in Section 2. In Section 3, we introduce certain graphs $\mathcal{G}_z$ derived from binary words z, which are subsequently used in Section 4 to state and prove our formula for the zeta function of a PFT. Finally, in Section 5, we compare our formula for the zeta function of a PFT with the known formula for the zeta function of a general sofic shift.

[†]Graduate School of Information Systems, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, 182-8585, JAPAN

[††]Dept. of Electrical Communication Engineering, Indian Institute of Science, Bangalore, 560 012, INDIA

[*]The paper was presented in part at the 2009 IEEE International Symposium on Information Theory (ISIT'09), Seoul, Korea, June 28-July 3, 2009.

a) E-mail: amanada@appnet.is.uec.ac.jp

b) E-mail: nkashyap@ece.iisc.ernet.in

DOI: 10.1587/transfun.E0.A.1

2. Fundamental Background

We begin with the basic background, based on material from [2] and [6]. Let Σ be an *alphabet*, a finite set of symbols. We always assume that $|\Sigma| \geq 2$ since $|\Sigma| = 1$ gives us the trivial case. A *word* is a finite-length sequence over Σ, and the length of a word w is denoted by $|w|$. For a bi-infinite sequence $\mathbf{x} = \ldots x_{-1}x_0x_1 \ldots$ over Σ, we call a word w (with length $|w| = n$) a *subword* (or *factor*) of $\mathbf{x}$, denoted by $w \prec \mathbf{x}$, if $w = x_ix_{i+1}\ldots x_{i+n-1}$ for some integer i. We will write $w \prec_i \mathbf{x}$ when we want to emphasize the fact that w is a subword of $\mathbf{x}$ starting at the index i. For a bi-infinite sequence $\mathbf{x} = \ldots x_{-1}x_0x_1 \ldots$ and an integer $r \geq 1$, we define $\sigma^r(\mathbf{x}) = \ldots x^*_{-1}x^*_0x^*_1 \ldots$, the *r-shifted sequence* of $\mathbf{x}$, to be the bi-infinite sequence satisfying $x^*_i = x_{i+r}$ for every integer i. The sequence $\mathbf{x}$ is said to be *periodic* if $\mathbf{x} = \sigma^r(\mathbf{x})$ for some $r \geq 1$; In this case, r is called a *period* of $\mathbf{x}$, and $\mathbf{x}$ can be written as $\mathbf{x} = (x_0x_1\ldots x_{r-1})^\infty$.

Each graph $\mathcal{G}$ we focus on in this paper is a labeled directed graph; *i.e.*, a directed graph with a label assigned to each edge. We denote by $\mathcal{V}_\mathcal{G}$ the vertex set of $\mathcal{G}$. We will refer to vertices of graphs as *states*. The *adjacency matrix* $A_\mathcal{G}$ of such a graph is a $|\mathcal{V}_\mathcal{G}| \times |\mathcal{V}_\mathcal{G}|$ matrix whose rows and columns are indexed (in the same order) by the elements of $\mathcal{V}_\mathcal{G}$. For $U, V \in \mathcal{V}_\mathcal{G}$, the (U, V)-th entry of $A_\mathcal{G}$ is the number of directed edges from U to V in $\mathcal{G}$.

Given a labeled directed graph $\mathcal{G}$, where edge labels come from Σ, let $S(\mathcal{G})$ be the set of bi-infinite sequences which are generated by reading off labels along bi-infinite paths in $\mathcal{G}$. A *sofic shift* $\mathcal{S}$ is a set of bi-infinite sequences such that $\mathcal{S} = S(\mathcal{G})$ for some labeled directed graph $\mathcal{G}$. In this case, we say that $\mathcal{S}$ is *presented by* $\mathcal{G}$, or that $\mathcal{G}$ is a *presentation* of $\mathcal{S}$. A classic example of a sofic shift is a *shift of finite type* (*SFT*) $\mathcal{Y} = \mathcal{Y}_{\mathcal{F}'}$, where $\mathcal{F}'$ is a finite set of forbidden words (a *forbidden set*). The SFT $\mathcal{Y} = \mathcal{Y}_{\mathcal{F}'}$ is defined to be the set of all bi-infinite sequences $\mathbf{x} = \ldots x_{-1}x_0x_1 \ldots$ over Σ such that $\mathbf{x}$ contains no word $f' \in \mathcal{F}'$ as a subword.

A *periodic-finite-type shift* (*PFT*) is characterized by an ordered list of finite sets $\mathcal{F} = (\mathcal{F}^{(0)}, \mathcal{F}^{(1)}, \ldots, \mathcal{F}^{(T-1)})$ and a *period* T. More precisely, the PFT $\mathcal{X}_{\{\mathcal{F},T\}}$ is defined as the set of all bi-infinite sequences $\mathbf{x}$ over Σ such that for some integer $r \in \{0, 1, \ldots, T-1\}$, the r-shifted sequence $\sigma^r(\mathbf{x})$ of $\mathbf{x}$ satisfies $f \prec_i \sigma^r(\mathbf{x}) \Longrightarrow f \notin \mathcal{F}^{(i \mod T)}$ for every integer i. It is easy to see that a PFT $\mathcal{X}_{\{\mathcal{F},T\}}$ with period $T = 1$ is the SFT $\mathcal{Y}_{\mathcal{F}'}$ with $\mathcal{F}' = \mathcal{F}^{(0)}$. Thus, the class of SFT's is (strictly) included in the class of PFT's.

Any PFT $\mathcal{X}$ has a representation of the form $\mathcal{X}_{\{\mathcal{F},T\}}$ such that $\mathcal{F}^{(j)} = \emptyset$ for $1 \leq j \leq T-1$, and every word in $\mathcal{F}^{(0)}$ has the same length. An arbitrary representation $\mathcal{X}_{\{\widehat{\mathcal{F}},T\}}$ can be converted to one in the above form as follows. For a given PFT $\mathcal{X} = \mathcal{X}_{\{\widehat{\mathcal{F}},T\}}$, if $\hat{f} \in \widehat{\mathcal{F}}^{(j)}$ for some $1 \leq j \leq T-1$, then list out all words of length $j + |\hat{f}|$ which end with $\hat{f}$, add them to $\widehat{\mathcal{F}}^{(0)}$, and delete $\hat{f}$ from $\widehat{\mathcal{F}}^{(j)}$. Continue this process until $\widehat{\mathcal{F}}^{(1)} = \cdots = \widehat{\mathcal{F}}^{(T-1)} = \emptyset$. Next, find the longest word in the resulting $\widehat{\mathcal{F}}^{(0)}$, and let ℓ denote its length. Define $\mathcal{F}^{(0)} = \{f \in \Sigma^\ell : f$ starts with some word in $\widehat{\mathcal{F}}^{(0)}\}$. It is easy to check that $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ with $\mathcal{F} = (\mathcal{F}^{(0)}, \emptyset, \ldots, \emptyset)$, and every word in $\mathcal{F}^{(0)}$ has the same length ℓ. Throughout this paper, for a given PFT $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$, we always assume that $\mathcal{F}$ is in *standard form*, *i.e.*, $\mathcal{F} = (\mathcal{F}^{(0)}_\mathcal{X}, \emptyset, \ldots, \emptyset)$, and $\mathcal{F}^{(0)}_\mathcal{X}$ is a subset of Σ^ℓ for some $\ell \geq 1$.

Moision and Siegel proved that every PFT is a sofic shift, that is, each PFT has a presentation, by giving an algorithm to construct a presentation of a PFT [2], [9]. We call their algorithm the *MS algorithm*, and we refer to the presentation of a PFT $\mathcal{X}$ resulting from the algorithm as the *MS presentation* of $\mathcal{X}$, denoted by $\mathcal{G}^{(\mathrm{ms})}_\mathcal{X}$. The MS algorithm, given a PFT $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ with $\mathcal{F}$ in standard form as input, runs as follows.

Algorithm 1 : The MS Algorithm

define T sets of words $\mathcal{V}^{(0)}, \mathcal{V}^{(1)}, \ldots, \mathcal{V}^{(T-1)}$ as $\mathcal{V}^{(0)} = \Sigma^\ell \setminus \mathcal{F}^{(0)}_\mathcal{X}$ and $\mathcal{V}^{(1)} = \cdots = \mathcal{V}^{(T-1)} = \Sigma^\ell$.
for each integer $0 \leq j \leq T-1$, **and for** each pair of words $u = u_1u_2\ldots u_\ell \in \mathcal{V}^{(j)}$ and $v = v_1v_2\ldots v_\ell \in \mathcal{V}^{(j+1 \mod T)}$
    if $u_2 \ldots u_\ell = v_1 \ldots v_{\ell-1}$ **then**
        draw an edge labeled v_ℓ from u to v.
    else
        draw no edge from u to v.
return the resulting graph and name it $\mathcal{G}^{(\mathrm{ms})}_\mathcal{X}$.

We note here that Béal, Crochemore and Fici have given an algorithm, different from the MS algorithm, which also generates a presentation of a PFT [1].

The MS presentation of a PFT is an example of a *word-based graph* (*WBG*), which we define to be any labeled directed graph $\mathcal{G}$ with the following properties:

(W1) Every state in $\mathcal{G}$ is a word $w \in \Sigma^\ell$ for some $\ell \geq 1$.
(W2) The vertex set consists of T disjoint *phases* $\mathcal{V}^{(0)}, \mathcal{V}^{(1)}, \ldots, \mathcal{V}^{(T-1)}$ for some $T \geq 1$, and each phase has at most one state corresponding to $w \in \Sigma^\ell$. We denote by $w^{(i)}$ the state in $\mathcal{V}^{(i)}$ corresponding to w.
(W3) There is an edge labeled $a \in \Sigma$ from $u^{(i)} = u_1u_2\ldots u_\ell \in \mathcal{V}^{(i)}$ to $v^{(i+1 \mod T)} = v_1v_2\ldots v_\ell \in \mathcal{V}^{(i+1 \mod T)}$ if and only if $u_2\ldots u_\ell = v_1\ldots v_{\ell-1}$ and $v_\ell = a$.

Observe that WBG's are always *deterministic*, that is, distinct outgoing edges from the same state are labeled distinctly.

Given a WBG $\mathcal{G}$ and a finite-length path $\alpha : V_0 \to V_1 \to \cdots \to V_n$ in $\mathcal{G}$, let $\mathbf{I}_\mathcal{G}(\alpha)$ and $\mathbf{T}_\mathcal{G}(\alpha)$ denote the initial state (V_0) and terminal state (V_n) of α in $\mathcal{G}$, respectively. In the case when $\mathbf{I}_{\mathcal{G}}(\alpha) = \mathbf{T}_{\mathcal{G}}(\alpha) = V$, we call α a *cycle at* V. Furthermore, we denote by $\mathcal{L}_{\mathcal{G}}(\alpha)$ the sequence which is generated by reading off labels along α. We also simply say x is *generated by* α (in $\mathcal{G}$) if $x = \mathcal{L}_{\mathcal{G}}(\alpha)$. The length of a path (or cycle) α is equal to the number of edges in α, and is denoted by $|\alpha|$.

The following remark encapsulates some observations that follow easily from the definition of a WBG.

Remark 2.1: In a WBG $\mathcal{G}$ with T phases, and states belonging to Σ^ℓ,

(1) if a path $\alpha : V_0 \to V_1 \to \cdots \to V_n$ originates at $V_0 \in \mathcal{V}^{(i)}$, then for $0 \le r \le n$, we have $V_r \in \mathcal{V}^{(k)}$ if and only if $r \equiv k - i \pmod{T}$;
(2) a path α of length $|\alpha| \ge \ell$ terminates at a state corresponding to the word $u \in \Sigma^\ell$ if and only if the length-ℓ suffix of $\mathcal{L}_{\mathcal{G}}(\alpha)$ is u;
(3) there is no cycle of length n if $n \not\equiv 0 \pmod{T}$.

3. The Word-Based Graphs $\mathcal{G}_z$

The statement and proof of our main result, namely, an expression for the zeta function of a PFT $\mathcal{X}_{\{\mathcal{F},T\}}$, makes extensive use of the adjacency matrices of certain WBG's determined by $\mathcal{F}$ and T. We introduce these WBG's in this section, and record an important property of their adjacency matrices.

We need a few definitions first. For a word z over some alphabet, let $z^\sharp$ denote the *primitive root* of z, *i.e.*, $z^\sharp$ is the shortest word such that $z = (z^\sharp)^n$ for some integer $n \ge 1$. In particular, for a binary word z (over the alphabet $\{0, 1\}$), we will find it convenient to define L_z to be the length of $z^\sharp$, W_z to be the number of 1's in $z^\sharp$, and $N_z = |z|/L_z$. Thus, $z = (z^\sharp)^{N_z}$.

Now, let $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ be a PFT with period T and $\mathcal{F}$ in standard form. Set $Z_T = \{0,1\}^T \setminus \{0^T\}$. For $z = z_0 z_1 \dots z_{T-1} \in Z_T$, let $\sigma_T(z)$ denote the cyclically shifted sequence $z_1 z_2 \dots z_{T-1} z_0$. We say that $z, z' \in Z_T$ are *conjugate* (or *cyclically equivalent*) if $z' = \sigma_T^q(z)$ for some integer q. Conjugacy is an equivalence relation on Z_T, which partitions Z_T into *conjugacy classes*. We then construct a set Ω_T by picking from each conjugacy class of Z_T one representative sequence $z = z_0 z_1 \dots z_{T-1}$ such that $z_0 = 1$.

For each $z = z_0 z_1 \dots z_{T-1} \in \Omega_T$, let us denote by $\mathcal{G}_z$ the word-based graph with L_z phases $\mathcal{V}^{(0)}, \mathcal{V}^{(1)}, \dots, \mathcal{V}^{(L_z-1)}$ defined (for $i = 0, 1, \dots, L_z - 1$) by

$$\mathcal{V}^{(i)} = \begin{cases} \Sigma^\ell & \text{when } z_i = 0. \\ \Sigma^\ell \setminus \mathcal{F}_{\mathcal{X}}^{(0)} & \text{when } z_i = 1. \end{cases}$$

For example, $\mathcal{G}_{10^{T-1}}$ is the MS presentation $\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}$ of $\mathcal{X}$. Furthermore, let us denote by $\mathcal{H}_z$ the subgraph of $(\mathcal{G}_z)^{L_z}$ (the L_z-th power graph of $\mathcal{G}_z$) induced by $\mathcal{V}^{(0)} = \Sigma^\ell \setminus \mathcal{F}_{\mathcal{X}}^{(0)}$. More precisely, the vertex set of $\mathcal{H}_z$ is $\mathcal{V}^{(0)}$ in $\mathcal{G}_z$, and there is an edge from $u^{(0)}$ to $v^{(0)}$ in $\mathcal{H}_z$ if and only if there is a path of length L_z from $u^{(0)}$ to $v^{(0)}$ in $\mathcal{G}_z$. We denote by A_z and B_z the adjacency matrices of $\mathcal{G}_z$ and $\mathcal{H}_z$, respectively. The following lemma states an important relationship between the traces of the matrices A_z and B_z defined above.

Lemma 3.1: For the adjacency matrices A_z and B_z defined above, (a) $\mathrm{tr}(A_z^n) = 0$ if $n \not\equiv 0 \pmod{L_z}$, and (b) $\mathrm{tr}(A_z^{L_z m}) = L_z \times \mathrm{tr}(B_z^m)$ for any integer $m \ge 1$.

Proof: (a) follows directly from Remark 2.1(3). For (b), first observe that for any integer $m \ge 1$, $\mathrm{tr}(A_z^{L_z m}) = \sum_{i=0}^{L_z-1} \mathrm{tr}(B_{\mathcal{V}^{(i)}}^m)$, where $B_{\mathcal{V}^{(i)}}$ is the adjacency matrix of the subgraph of $(\mathcal{G}_z)^{L_z}$ induced by $\mathcal{V}^{(i)}$. However, we also have $\mathrm{tr}(B_{\mathcal{V}^{(0)}}^m) = \mathrm{tr}(B_{\mathcal{V}^{(1)}}^m) = \cdots = \mathrm{tr}(B_{\mathcal{V}^{(L_z-1)}}^m)$, since every cycle of length $L_z m$ in $\mathcal{G}_z$ can be viewed as a cycle at a state in $\mathcal{V}^{(i)}$ for any $0 \le i \le L_z - 1$. Hence, $\mathrm{tr}(A_z^{L_z m}) = L_z \times \mathrm{tr}(B_{\mathcal{V}^{(0)}}^m) = L_z \times \mathrm{tr}(B_z^m)$. ∎

4. The Zeta Function of a PFT

The zeta function of a sofic shift $\mathcal{S}$ is a generating function for the number of periodic sequences in $\mathcal{S}$. More precisely, for a given sofic shift $\mathcal{S}$, the zeta function $\zeta_{\mathcal{S}}(t)$ of $\mathcal{S}$ is defined to be

$$\zeta_{\mathcal{S}}(t) = \exp\left(\sum_{n=1}^{\infty} \frac{|P_n(\mathcal{S})|}{n} t^n \right), \tag{1}$$

where $P_n(\mathcal{S})$ is the set of periodic sequences in $\mathcal{S}$ of period n. In fact, there exists a formula for computing the zeta function of a sofic shift, which shows that the zeta function of a sofic shift is always rational [8],[6, Theorem 6.4.8]. In the particular case when $\mathcal{S}$ is an SFT $\mathcal{Y} = \mathcal{Y}_{\mathcal{F}'}$ with forbidden set $\mathcal{F}' \subset \Sigma^\ell$, the zeta function is simply the reciprocal of a polynomial [3]. More precisely (see [6, Theorem 6.4.6]),

$$\zeta_{\mathcal{Y}}(t) = \frac{1}{\det(I - tA_{\mathcal{G}_{\mathcal{Y}}^{(\mathrm{ms})}})}, \tag{2}$$

where $A_{\mathcal{G}_{\mathcal{Y}}^{(\mathrm{ms})}}$ is the adjacency matrix of the MS presentation of $\mathcal{Y}$, by taking $\mathcal{Y} = \mathcal{X}_{\{(\mathcal{F}'),1\}}$, and I is the identity matrix (with the same order of $A_{\mathcal{G}_{\mathcal{Y}}^{(\mathrm{ms})}}$). In this section, we prove the following theorem, which is the main result of our paper.

Theorem 4.1: For a PFT $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ with period T and $\mathcal{F}$ in standard form, the zeta function $\zeta_{\mathcal{X}}(t)$ of $\mathcal{X}$ is given by

$$\zeta_{\mathcal{X}}(t) = \prod_{z \in \Omega_T} [\det(I - tA_z)]^{(-1)^{W_z}} \times \prod_{\substack{z \in \Omega_T: \\ N_z \text{ is even}, W_z \text{ is odd}}} \det(I - t^{2L_z} {B_z}^2).$$

Theorem 4.1 clearly shows that the zeta function of a PFT is rational, and the number of graphs (or adjacency matrices) needed to compute the zeta function of a PFT depends only on its period T. Furthermore, the case when N_z is even and W_z is odd can happen only when period T is even. Therefore, when T is odd, we can compute the zeta function using only the adjacency matrices A_z:

$$\zeta_{\mathcal{X}}(t) = \prod_{z \in \Omega_T} [\det(I - tA_z)]^{(-1)^{W_z}}. \tag{3}$$

Note that when $T = 1$, then $\Omega_T = \{1\}$ and $A_1 = A_{\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}}$. Hence, (3) reduces to $\zeta_{\mathcal{X}}(t) = [\det(I - tA_{\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}})]^{-1}$, showing that (2) is a special case of Theorem 4.1.

The remainder of this section is devoted to a proof of Theorem 4.1. The main idea of the proof is to express the number, $|P_n(\mathcal{X})|$, of period-n sequences in $\mathcal{X}$ in terms of the traces of the nth powers of the matrices A_z. This takes some development, which we kick off with two easy lemmas.

Lemma 4.2: Let $\mathcal{G}$ be a WBG, and consider two states $u^{(i)} = u_1u_2\ldots u_\ell$ and $v^{(j)} = v_1v_2\ldots v_\ell$ (for some i and j) in $\mathcal{G}$. For any pair of cycles C and $\widetilde{C}$, with $|C| = |\widetilde{C}|$, at $u^{(i)}$ and $v^{(j)}$, respectively, if $\mathcal{L}_{\mathcal{G}}(C) = \mathcal{L}_{\mathcal{G}}(\widetilde{C})$, then both u and v are copies of the same word in Σ^ℓ.

Proof: Let m be an integer such that $m|C| \geq \ell$. Since $(\mathcal{L}_{\mathcal{G}}(C))^m = (\mathcal{L}_{\mathcal{G}}(\widetilde{C}))^m$ and $|(\mathcal{L}_{\mathcal{G}}(C))^m| = m|C| \geq \ell$, both $\mathbf{T}_{\mathcal{G}}(C^m)$ and $\mathbf{T}_{\mathcal{G}}(\widetilde{C}^m)$ are the length-ℓ suffix of $(\mathcal{L}_{\mathcal{G}}(C))^m$ by Remark 2.1(2). Observe that $\mathbf{T}_{\mathcal{G}}(C^m) = \mathbf{T}_{\mathcal{G}}(C) = u^{(i)}$ and $\mathbf{T}_{\mathcal{G}}(\widetilde{C}^m) = \mathbf{T}_{\mathcal{G}}(\widetilde{C}) = v^{(j)}$. ∎

Lemma 4.3: Let $\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}$ be the MS presentation of a PFT $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ with period T. When $n \equiv 0 \pmod{T}$, a periodic sequence $\mathbf{x} = (x_0x_1\ldots x_{n-1})^\infty$ is in $P_n(\mathcal{X})$ iff $x_0x_1\ldots x_{n-1} = \mathcal{L}_{\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}}(C)$ for some cycle C of length n in $\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}$.

Proof: If $x_0x_1\ldots x_{n-1} = \mathcal{L}_{\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}}(C)$ for some cycle C of length n in $\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}$, then clearly $\mathbf{x} = (x_0x_1\ldots x_{n-1})^\infty$ is in $P_n(\mathcal{X})$ since $\mathbf{x} = (\mathcal{L}_{\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}}(C))^\infty$.

Conversely, suppose that $\mathbf{x} = (x_0x_1\ldots x_{n-1})^\infty$ is in $P_n(\mathcal{X})$. Let m be an integer such that $mn \geq \ell$. For a bi-infinite path α in $\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}$ that generates $\mathbf{x}$, a finite-length subpath β of α generating $(x_0x_1\ldots x_{n-1})^m$ terminates at $w^{(i)}$ for some $0 \leq i \leq T-1$, where w is the length-ℓ suffix of $(x_0x_1\ldots x_{n-1})^m$, by Remark 2.1(2). From $w^{(i)}$, there must be a path $\gamma : w^{(i)} \to V_1 \to \cdots \to V_n$ generating $x_0x_1\ldots x_{n-1}$. Observe that the terminal state V_n of path γ is the terminal state of path $\beta\gamma$ as well, which implies (again by Remark 2.1(2)) $V_n = \mathbf{T}_{\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}}(\beta\gamma) = w^{(j)}$ for some $0 \leq j \leq T-1$. However, since $n \equiv 0 \pmod{T}$ by assumption, $i = j$ holds from Remark 2.1(1). This shows that γ is a cycle at $w^{(i)}$ generating $x_0x_1\ldots x_{n-1}$. ∎

Lemma 4.3 is not true when $n \not\equiv 0 \pmod{T}$ since, by (3) in Remark 2.1, there is no cycle of length n in $\mathcal{G}_{\mathcal{X}}^{(\mathrm{ms})}$. However, the following proposition shows that we can generate $\mathbf{x} \in P_n(\mathcal{X})$ using a cycle C in the MS presentation $\mathcal{G}_{\mathcal{X}_d}^{(\mathrm{ms})}$ of $\mathcal{X}_d$, where $\mathcal{X}_d = \mathcal{X}_{\{\mathcal{F}_d,d\}}$ is the PFT with period $d = \gcd(n,T)$ and $\mathcal{F}_d = (\mathcal{F}_{\mathcal{X}}^{(0)}, \emptyset, \ldots, \emptyset)$.

Proposition 4.4: Let $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ be a PFT with period T and $\mathcal{F}$ in standard form (*i.e.*, $\mathcal{F}^{(0)} = \mathcal{F}_{\mathcal{X}}^{(0)} \subset \Sigma^\ell$ and $\mathcal{F}^{(1)} = \cdots = \mathcal{F}^{(T-1)} = \emptyset$). If $\gcd(n,T) = d$, then $\mathbf{x} \in P_n(\mathcal{X})$ if and only if $\mathbf{x} \in P_n(\mathcal{X}_d)$, where $\mathcal{X}_d = \mathcal{X}_{\{\mathcal{F}_d,d\}}$ is the PFT with period d and $\mathcal{F}_d = (\mathcal{F}_{\mathcal{X}}^{(0)}, \emptyset, \ldots, \emptyset)$.

Proof: Clearly, $\mathbf{x} \in P_n(\mathcal{X}_d)$ implies $\mathbf{x} \in P_n(\mathcal{X})$ by definition. For the converse, suppose that $\mathbf{x} \notin P_n(\mathcal{X}_d)$. Then, for any $0 \leq r \leq d-1$, we have $f_r \prec_{\hat{i}_r} \sigma^r(\mathbf{x})$ for some $f_r \in \mathcal{F}_{\mathcal{X}}^{(0)}$ and integer $\hat{i}_r \equiv 0 \pmod{d}$, that is, $f_r \prec_{i_r} \mathbf{x}$ for $i_r = \hat{i}_r + r \equiv r \pmod{d}$. Since $\gcd(n,T) = d$ by assumption, there exists an integer $m \geq 1$ such that $mn \equiv d \pmod{T}$. As $\mathbf{x}$ is a periodic bi-infinite sequence of period n, for any $0 \leq r \leq d-1$, $\mathbf{x}$ contains f_r at indices $i_r + smn$, $s = 0, 1, \ldots, T/d-1$. This implies that for each $r' \in \{0, 1, \ldots, T-1\}$, we have $f_{r'} \prec_{j'_r} \sigma^{r'}(\mathbf{x})$ for some $f_{r'} \in \mathcal{F}_{\mathcal{X}}^{(0)}$ and integer $j_{r'} \equiv 0 \pmod{T}$. Therefore, $\mathbf{x} \notin \mathcal{X}$, and in particular, $\mathbf{x} \notin P_n(\mathcal{X})$. ∎

Thus, for the PFT's $\mathcal{X}$ and $\mathcal{X}_d$ defined in Proposition 4.4, when $\gcd(n,T) = d$, as $n \equiv 0 \pmod{d}$, we have from Lemma 4.3 that $\mathbf{x} \in P_n(\mathcal{X}) = P_n(\mathcal{X}_d)$ if and only if $\mathbf{x} = (\mathcal{L}_{\mathcal{G}_{\mathcal{X}_d}^{(\mathrm{ms})}}(C))^\infty$ for some cycle C in the MS presentation $\mathcal{G}_{\mathcal{X}_d}^{(\mathrm{ms})}$ of $\mathcal{X}_d$. That is, for any n, every periodic sequence $\mathbf{x} \in P_n(\mathcal{X})$ can be generated by a cycle in the MS presentation of some PFT.

For a WBG $\mathcal{G}$, let $\mathcal{C}_n(w^{(i)})_{\mathcal{G}}$ be the set of periodic sequences $\mathbf{x} = (x_0x_1\ldots x_{n-1})^\infty$ which can be generated by a cycle C of length n at $w^{(i)}$ in $\mathcal{G}$. In other words, $\mathbf{x} = (x_0x_1\ldots x_{n-1})^\infty \in \mathcal{C}_n(w^{(i)})_{\mathcal{G}}$ iff there exists a cycle C at $w^{(i)}$ satisfying $\mathcal{L}_{\mathcal{G}}(C) = x_0x_1\ldots x_{n-1}$. By convention, $\mathcal{C}_n(w^{(i)})_{\mathcal{G}} = \emptyset$ if $w^{(i)}$ is not a state in $\mathcal{G}$. Putting together Lemma 4.2, Lemma 4.3 and Proposition 4.4, we have the following corollary.

Corollary 4.5: Let $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ be a PFT with period T and $\mathcal{F}$ in standard form. Given an integer $n \geq 1$, suppose $\gcd(n,T) = d$, and consider the PFT $\mathcal{X}_d = \mathcal{X}_{\{\mathcal{F}_d,d\}}$ with period d and $\mathcal{F}_d = (\mathcal{F}_{\mathcal{X}}^{(0)}, \emptyset, \ldots, \emptyset)$. Then, we have that

$$|P_n(\mathcal{X})| = |P_n(\mathcal{X}_d)| = \sum_{w \in \Sigma^\ell} \left| \bigcup_{i=0}^{d-1} \mathcal{C}_n(w^{(i)})_{\mathcal{G}_{\mathcal{X}_d}^{(\mathrm{ms})}} \right|. \tag{4}$$

Proof: The first equality is obvious from Proposition 4.4. Also, as $\mathcal{X}_d$ has period d and $n \equiv 0 \pmod d$, it follows from Lemma 4.3 that

$$|P_n(\mathcal{X}_d)| = \left| \bigcup_{w \in \Sigma^\ell} \bigcup_{i=0}^{d-1} \mathcal{C}_n(w^{(i)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} \right|.$$

Furthermore, $\mathcal{C}_n(w^{(i)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} \cap \mathcal{C}_n(\widehat{w}^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} = \emptyset$ if $w \neq \widehat{w}$ by Lemma 4.2, which shows the second equality. ∎

For (4), we have from the inclusion-exclusion principle that

$$\left| \bigcup_{i=0}^{d-1} \mathcal{C}_n(w^{(i)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} \right| = \sum_{J \neq \emptyset, J \subseteq [d]} (-1)^{|J|-1} \left| \bigcap_{j \in J} \mathcal{C}_n(w^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} \right|, \qquad (5)$$

where $[d] = \{0, 1, \ldots, d-1\}$. Therefore, our goal is to count $\left| \bigcap_{j \in J} \mathcal{C}_n(w^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} \right|$ for each non-empty set $J \subseteq [d]$. To do this, we first show that the intersection $\bigcap_{j \in J} \mathcal{C}_n(w^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}}$ can be replaced a single set $\mathcal{C}_n(w^{(q)})_{\mathcal{G}_z}$ determined in some WBG $\mathcal{G}_z$. Note that the following lemma is stated for an arbitrary PFT with period T, so that it applies in particular to the PFT's $\mathcal{X}_d$ defined above, upon setting $T = d$. The pieces of notation Ω_T and L_z used in the lemma were defined in Section 3.

Lemma 4.6: Let $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ be a PFT with period T and $\mathcal{F}$ in standard form. For each $z = z_0 z_1 \ldots z_{T-1} \in \Omega_T$ and each integer $0 \leq q \leq L_z - 1$, define

$$J(z,q) = \{(q-i) \bmod T : 0 \leq i \leq T-1,\ z_i = 1\}. \qquad (6)$$

The mapping $(z,q) \mapsto J(z,q)$ is a bijection between pairs (z,q) as above and non-empty sets $J \subseteq [T]$. Furthermore, setting $J = J(z,q)$, we have

$$\mathcal{C}_n(w^{(q)})_{\mathcal{G}_z} = \bigcap_{j \in J} \mathcal{C}_n(w^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}}, \qquad (7)$$

for any integer $n \equiv 0 \pmod T$, and any word $w \in \Sigma^\ell$.

Proof: It is easy to verify that the mapping $(z,q) \mapsto J(z,q)$ is a bijection as stated above, so we focus on proving (7). For clarity, we use the notation $\mathcal{V}^{(i)}_{\mathrm{ms}}$ and $\mathcal{V}^{(i)}_{\mathcal{G}_z}$ (for some i) to denote the phase $\mathcal{V}^{(i)}$ in $\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}$ and the phase $\mathcal{V}^{(i)}$ in $\mathcal{G}_z$, respectively.

Consider $\mathbf{x} = (x_0 x_1 \ldots x_{n-1})^\infty \in \mathcal{C}_n(w^{(q)})_{\mathcal{G}_z}$. Then there exists a cycle $C : V_0 = w^{(q)} \to V_1 \to \cdots \to V_n = w^{(q)}$ of length n at $w^{(q)} \in \mathcal{V}^{(q)}_{\mathcal{G}_z}$ such that $\mathcal{L}_{\mathcal{G}_z}(C) = x_0 x_1 \ldots x_{n-1}$. Recall from Remark 2.1(1) that $V_r \in \mathcal{V}^{(i')}_{\mathcal{G}_z}$, where $0 \leq i' \leq L_z - 1$, if and only if $r \equiv i' - q \pmod{L_z}$. Since $\mathcal{V}^{(i')}_{\mathcal{G}_z} = \Sigma^\ell \setminus \mathcal{F}^{(0)}_{\mathcal{X}}$ iff $z_{i'} = 1$ for some $0 \leq i' \leq L_z - 1$, we have that V_r cannot be a word in $\mathcal{F}^{(0)}_{\mathcal{X}}$ if and only if $r \equiv i' - q \pmod{L_z}$ for some $0 \leq i' \leq L_z - 1$ satisfying $z_{i'} = 1$. Furthermore, as $z_{i'} = z_{i'+L_z} = \cdots = z_{i'+(N_z-1)L_z}$ for each $0 \leq i' \leq L_z - 1$, we infer that V_r cannot be a word in $\mathcal{F}^{(0)}_{\mathcal{X}}$ if and only if $r \equiv i - q \pmod T$ for some $0 \leq i \leq T-1$ satisfying $z_i = 1$.

Now, for $J = \{j_1, j_2, \ldots, j_{|J|}\} = J(z,q)$, consider $\mathbf{x}' = (x'_0 x'_1 \ldots x'_{n-1})^\infty \in \bigcap_{j \in J} \mathcal{C}_n(w^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}}$. Then, for each $j \in J$ and a state $w^{(j)} \in \mathcal{V}^{(j)}_{\mathrm{ms}}$, there exists a cycle $C^{(j)} : w^{(j)} \to V^{(j)}_1 \to \cdots \to V^{(j)}_{n-1} \to w^{(j)}$ of length n at $w^{(j)}$ such that $\mathcal{L}_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}}(C^{(j)}) = x'_0 x'_1 \ldots x'_{n-1}$. For the cycle $C^{(j)}$, we have, from Remark 2.1(1), that $V^{(j)}_r \in \mathcal{V}^{(0)}_{\mathrm{ms}}$ if and only if $r \equiv -j \pmod T$. That is, $V^{(j)}_r$ cannot be a word in $\mathcal{F}^{(0)}_{\mathcal{X}}$ if and only if $r \equiv -j \pmod T$. Since $\mathcal{L}_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}}(C^{(j_1)}) = \mathcal{L}_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}}(C^{(j_2)}) = \cdots = \mathcal{L}_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}}(C^{(j_{|J|})})$, we have that for each cycle $C^{(j)}$, $j \in J$, $V^{(j)}_r$ cannot be a word in $\mathcal{F}^{(0)}_{\mathcal{X}}$ if and only if $r \equiv -j_k \pmod T$ for some $j_k \in J$. Since $J = J(z,q)$ is as defined in (6), we find that $V^{(j)}_r$ cannot be a word in $\mathcal{F}^{(0)}_{\mathcal{X}}$ iff $r \equiv i - q \pmod T$ for some $0 \leq i \leq T-1$ satisfying $z_i = 1$.

Hence, we can see (by considering the edge structure of WBG's) $x_0 x_1 \ldots x_{n-1} = \mathcal{L}_{\mathcal{G}_z}(C)$ for some cycle C at $w^{(q)} \in \mathcal{V}^{(q)}_{\mathcal{G}_z}$ if and only if for any $j \in J$, there exists a cycle $C^{(j)}$ at $w^{(j)} \in \mathcal{V}^{(j)}_{\mathrm{ms}}$ such that $x_0 x_1 \ldots x_{n-1} = \mathcal{L}_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}}}(C^{(j)})$. This clearly shows that (z,q) and $J = J(z,q)$ satisfy (7) for any integer $n \equiv 0 \pmod T$ and any $w \in \Sigma^\ell$, as required. ∎

We are now in a position to give the key idea in the proof of our zeta function result, namely, that we can explicitly determine $|P_n(\mathcal{X})|$ for a PFT $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ using the adjacency matrices of the WBG's $\mathcal{G}_z$.

Lemma 4.7: Let $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ be a PFT with period T and $\mathcal{F}$ in standard form. For an integer $n \geq 1$, suppose that $\gcd(n,T) = d$ and consider the PFT $\mathcal{X}_d = \mathcal{X}_{\{\mathcal{F}_d,d\}}$ with period d and $\mathcal{F}_d = (\mathcal{F}^{(0)}_{\mathcal{X}}, \emptyset, \ldots, \emptyset)$. Then,

$$|P_n(\mathcal{X})| = |P_n(\mathcal{X}_d)| = \sum_{z \in \Omega_d} (-1)^{dW_z/L_z - 1} \operatorname{tr}(A_z^n).$$

Proof: We would like to show, from Corollary 4.5 and (5) that

$$\sum_{w \in \Sigma^\ell} \sum_{J \subseteq [d]} (-1)^{|J|-1} \left| \bigcap_{j \in J} \mathcal{C}_n(w^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} \right| = \sum_{z \in \Omega_d} (-1)^{dW_z/L_z - 1} \operatorname{tr}(A_z^n). \qquad (8)$$

Pick $w \in \Sigma^\ell$ arbitrarily. Applying Lemma 4.6 to the PFT $\mathcal{X}_d$, we see that the mapping that takes a pair (z,q), with $z \in \Omega_d$ and $0 \leq q \leq L_z - 1$, to

$$J(z,q) = \{(q-i) \bmod d : 0 \le i \le d-1,\ z_i = 1\} \quad (9)$$

gives us a one-to-one correspondence between such pairs (z,q) and non-empty sets $J \subseteq [d]$, such that

$$\sum_{J \subseteq [d]} (-1)^{|J|-1} \left| \bigcap_{j \in J} \mathcal{C}_n(w^{(j)})_{\mathcal{G}^{(\mathrm{ms})}_{\mathcal{X}_d}} \right| = \sum_{z \in \Omega_d} \sum_{q=0}^{L_z-1} (-1)^{|J(z,q)|-1} \left| \mathcal{C}_n(w^{(q)})_{\mathcal{G}_z} \right|.$$

Now, it is clear from (9) that $|J(z,q)|$ is equal to the number of 1's in z, which in turn is equal to $(d/L_z)W_z$. Therefore, the left-hand side of (8) can be expressed as

$$\sum_{w \in \Sigma^\ell} \sum_{z \in \Omega_d} \sum_{q=0}^{L_z-1} (-1)^{(d/L_z)W_z-1} \left| \mathcal{C}_n(w^{(q)})_{\mathcal{G}_z} \right|. \quad (10)$$

To simplify the above expression, we make use of the fact that WBG's are deterministic. In particular, $\mathcal{G}_z$ is deterministic, so that for any state $w^{(q)}$ in $\mathcal{G}_z$, $|\mathcal{C}_n(w^{(q)})_{\mathcal{G}_z}|$ is equal to the number of cycles of length n at $w^{(q)}$ in $\mathcal{G}_z$, which is the $(w^{(q)}, w^{(q)})$-th entry of A_z^n. Hence,

$$\sum_{w \in \Sigma^\ell} \sum_{q=0}^{L_z-1} \left| \mathcal{C}_n(w^{(q)})_{\mathcal{G}_z} \right| = \mathrm{tr}(A_z^n).$$

Thus, the expression in (10) evaluates to the right-hand side of (8), which proves the lemma. ■

We use the above lemma to derive our expression for the zeta function of a PFT $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$. As is evident from (1), we need to evaluate the sum $\sum_{n=1}^{\infty} \frac{|P_n(\mathcal{X})|}{n} t^n$. We re-write this sum as

$$\sum_{d|T} \sum_{n:\gcd(n,T)=d} \frac{|P_n(\mathcal{X})|}{n} t^n. \quad (11)$$

Using Lemma 4.7, the sum above can be expressed as

$$\sum_{d|T} \sum_{z \in \Omega_d} \sum_{n:\gcd(n,T)=d} (-1)^{dW_z/L_z-1} \frac{\mathrm{tr}\left(A_z^n\right)}{n} t^n. \quad (12)$$

Observe from the definition of $\mathcal{G}_z$ that for $z \in \Omega_T$, $\mathcal{G}_z = \mathcal{G}_{\hat{z}}$ if and only if $\hat{z}$ can be represented as $\hat{z} = (z^\sharp)^k$, for some positive integer k. Therefore, for $z \in \Omega_T$, A_z appears in (12) if and only if $d = kL_z$ for some $k|N_z$. Thus, (12) can be expressed as

$$\sum_{z \in \Omega_T} \sum_{k|N_z} \sum_{n:\gcd(n,T)=kL_z} (-1)^{kW_z-1} \frac{\mathrm{tr}\left(A_z^n\right)}{n} t^n. \quad (13)$$

We will use the Möbius inversion formula of elementary number theory to put the innermost sum above in a different form; see, for example, [4]. This sum is of the form $\sum_{n:gcd(n,T)=kL_z} g(n)$ which, by a change of variable, can be put in the form $\sum_{n:gcd(n,\widetilde{T})=1} f(n)$, where $\widetilde{T} = T/kL_z = N_z/k$, and $f(n) = g(kL_z n)$. Now, for $r \in \mathbb{N}$, define $F(r)$ and $G(r)$ as

$$F(r) = \sum_{n:\gcd(n,\widetilde{T})=\frac{\widetilde{T}}{r}} f(n), \quad G(r) = \sum_{n:n \in \left(\frac{\widetilde{T}}{r}\right)\mathbb{N}} f(n).$$

Then,

$$G(\widetilde{T}) = \sum_{n \in \mathbb{N}} f(n) = \sum_{r|\widetilde{T}} \sum_{n:\gcd(n,\widetilde{T})=\frac{\widetilde{T}}{r}} f(n) = \sum_{r|\widetilde{T}} F(r).$$

So, the Möbius inversion formula gives us

$$\sum_{n:\gcd(n,\widetilde{T})=1} f(n) = F(\widetilde{T}) = \sum_{r|\widetilde{T}} \mu(r) G\left(\frac{\widetilde{T}}{r}\right) = \sum_{r|\widetilde{T}} \mu(r) \sum_{n \in r\mathbb{N}} f(n) = \sum_{r|\widetilde{T}} \mu(r) \sum_{m=1}^{\infty} f(rm),$$

where $\mu(\cdot)$ is the Möbius function. This allows us to write (13) as

$$\sum_{z \in \Omega_T} \sum_{k|N_z} \sum_{r|\frac{N_z}{k}} \mu(r) \sum_{m=1}^{\infty} (-1)^{kW_z-1} \frac{\mathrm{tr}\left(A_z^{rkL_z m}\right)}{rkL_z m} t^{rkL_z m}.$$

Using the change of variable $s = rk$ (so that the sum over pairs (k,r) is now a sum over pairs (s,r)), the above may be rewritten as

$$\sum_{z \in \Omega_T} \sum_{s|N_z} \sum_{r|s} \mu(r)(-1)^{(s/r)W_z-1} \sum_{m=1}^{\infty} \frac{\mathrm{tr}(A_z^{sL_z m})}{sL_z m} t^{sL_z m} = \sum_{z \in \Omega_T} \sum_{s|N_z} \beta(z,s) \sum_{m=1}^{\infty} \frac{\mathrm{tr}(A_z^{sL_z m})}{sL_z m} t^{sL_z m}, \quad (14)$$

where we have defined $\beta(z,s) = \sum_{r|s} \mu(r)(-1)^{(s/r)W_z-1}$. For $\beta(z,s)$, we have the following lemma.

Lemma 4.8: For any $z \in \Omega_T$ and $s|N_z$, we have

$$\beta(z,s) = \begin{cases} (-1)^{W_z-1} & \text{if } s=1 \\ -1-(-1)^{W_z-1} & \text{if } s=2 \\ 0 & \text{otherwise} \end{cases}$$

Proof: The expressions for $s=1$ and $s=2$ can be readily verified. So, we assume $s \ge 3$ from now on. The key ingredient in the proof is the standard Möbius function fact that, for any positive integer k, we have $\sum_{r|k} \mu(r) = 1$ if $k=1$, and $\sum_{r|k} \mu(r) = 0$ if $k > 1$.

First consider the case when $s \ge 3$ is odd. Then, for any $r|s$, we see that s/r is also odd, and hence, $(-1)^{(s/r)W_z-1} = (-1)^{W_z-1}$. We then have

$$\beta(z,s) = (-1)^{W_z-1} \sum_{r|s} \mu(r) = 0$$

as $s > 1$.

Next, consider the case when $s = 2^u$, with $u \geq 2$. In this case, any r that divides s is of the form $2^{u'}$ for some $u' \leq u$. It follows that for any $r|s$, $\mu(r) = 0$ unless $r = 1$ or $r = 2$. For $r = 1, 2$ we see that s/r is even, so that $(-1)^{(s/r)W_z - 1} = -1$. Thus,

$$\beta(z, s) = (-1)\,[\mu(1) + \mu(2)] = 0.$$

Finally, consider the case when $s = 2^u t$ with $u \geq 1$ and $t \geq 3$ odd. We split the divisors r of s into three groups: r odd, $r \equiv 2 \pmod 4$, and $r \equiv 0 \pmod 4$. For r's in the last group, $\mu(r) = 0$ by definition. For odd divisors r of s, we have s/r even and hence, $(-1)^{(s/r)W_z - 1} = -1$. Note also that r is an odd divisor of s iff $r|t$.

For divisors $r \equiv 2 \pmod 4$, it is easily checked that $s/r \equiv s/2 \pmod 2$. Hence, for such divisors, $(-1)^{(s/r)W_z - 1} = (-1)^{(s/2)W_z - 1}$. We also observe that $r \equiv 2 \pmod 4$ is a divisor of s iff $r = 2r'$ with $r'|t$.

Putting the above observations together, we find that when $s = 2^u t$ with $u \geq 1$ and $t \geq 3$ odd,

$$\beta(z, s) = (-1) \sum_{r|t} \mu(r) + (-1)^{(s/2)W_z - 1} \sum_{r'|t} \mu(2r').$$

The first sum is 0, since $t > 1$. To evaluate the second sum, we note that $\mu(2r') = -\mu(r')$ for any odd r'. Hence, $\sum_{r'|t} \mu(2r') = -\sum_{r'|t} \mu(r') = 0$. This proves that $\beta(z, s) = 0$ in this case as well. ∎

We are now in a position to prove our main result, namely, Theorem 4.1.

Proof of Theorem 4.1: We first simplify the expression in (14) using Lemma 4.8. For z's such that N_z is odd or W_z is even (in this case, by the lemma, $\beta(z, 2) = 0$ as well), we observe that

$$\begin{aligned}
&\sum_{s|N_z} \beta(z, s) \sum_{m=1}^{\infty} \frac{\mathrm{tr}(A_z^{sL_z m})}{sL_z m}\, t^{sL_z m} \\
&\quad = (-1)^{W_z - 1} \sum_{m=1}^{\infty} \frac{\mathrm{tr}(A_z^{L_z m})}{L_z m}\, t^{L_z m} \\
&\quad = (-1)^{W_z - 1} \sum_{n=1}^{\infty} \frac{\mathrm{tr}(A_z^{n})}{n}\, t^{n},
\end{aligned}$$

as $\mathrm{tr}(A_z^n) = 0$ if $n \not\equiv 0 \pmod{L_z}$, by Lemma 3.1(a). And for z's such that N_z is even and W_z is odd, we have

$$\begin{aligned}
&\sum_{s|N_z} \beta(z, s) \sum_{m=1}^{\infty} \frac{\mathrm{tr}(A_z^{sL_z m})}{sL_z m}\, t^{sL_z m} \\
&\quad = (-1)^{W_z - 1} \sum_{m=1}^{\infty} \frac{\mathrm{tr}(A_z^{L_z m})}{L_z m}\, t^{L_z m} \\
&\qquad + (-2) \sum_{m=1}^{\infty} \frac{\mathrm{tr}(A_z^{2L_z m})}{2L_z m}\, t^{2L_z m} \\
&\quad = (-1)^{W_z - 1} \sum_{n=1}^{\infty} \frac{\mathrm{tr}(A_z^{n})}{n}\, t^{n} - \sum_{m=1}^{\infty} \frac{\mathrm{tr}(B_z^{2m})}{m}\, t^{2L_z m},
\end{aligned}$$

where the last equality comes from Lemma 3.1(b).

The theorem now follows by plugging these expressions into (14), and then using the fact (see e.g., [6, Theorem 6.4.6]) that for any square matrix A and positive integer $k \geq 1$,

$$\exp\left(\sum_{n=1}^{\infty} \frac{\mathrm{tr}(A^{kn})}{n}\, t^{kn} \right) = \left[\det(I - t^k A^k)\right]^{-1}. \qquad (15)$$

∎

We noted previously that $\zeta_{\mathcal{X}}(t)$ has a compact expression, given in (3), when $\mathcal{X}$ has odd period T. This indicates that an easier derivation of the formula may exist in the case of odd T. Indeed, we provide below a derivation in this case that also starts from Lemma 4.7, but avoids the subsequent use of Möbius inversion.

Proof of Theorem 4.1 when T is odd: Recall from Lemma 4.7 that

$$|P_n(\mathcal{X})| = \sum_{z \in \Omega_d} (-1)^{dW_z/L_z - 1} \mathrm{tr}(A_z^n)$$

holds when $\gcd(n, T) = d$. It is obvious that for any $z \in \Omega_d$, there exists a (unique) $\hat{z} \in \Omega_T$ such that $z^\sharp = \hat{z}^\sharp$. The key claim is to show that for any $\hat{z} \in \Omega_T$ such that $\mathrm{tr}(A_{\hat{z}}^n) \neq 0$, there exists a (unique) $z \in \Omega_d$ such that $z^\sharp = \hat{z}^\sharp$.

Indeed, suppose that we can show the claim. Then, it follows that $W_z = W_{\hat{z}}$ and $\mathcal{G}_z = \mathcal{G}_{\hat{z}}$ for these $z \in \Omega_d$ and $\hat{z} \in \Omega_T$, and therefore,

$$\sum_{z \in \Omega_d} (-1)^{W_z - 1} \mathrm{tr}(A_z^n) = \sum_{\hat{z} \in \Omega_T} (-1)^{W_{\hat{z}} - 1} \mathrm{tr}(A_{\hat{z}}^n).$$

Since T is odd by assumption, d/L_z is always odd for any $z \in \Omega_d$. Therefore, dW_z/L_z is even if and only if W_z is even, which implies that $(-1)^{dW_z/L_z - 1}$ is equal to $(-1)^{W_z - 1}$. Thus, we have

$$\begin{aligned}
\sum_{z \in \Omega_d} (-1)^{dW_z/L_z - 1} \mathrm{tr}(A_z^n) &= \sum_{z \in \Omega_d} (-1)^{W_z - 1} \mathrm{tr}(A_z^n) \\
&= \sum_{\hat{z} \in \Omega_T} (-1)^{W_{\hat{z}} - 1} \mathrm{tr}(A_{\hat{z}}^n),
\end{aligned}$$

for any $n \geq 1$, and hence,

$$\sum_{n=1}^{\infty} \frac{|P_n(\mathcal{X})|}{n} t^n = \sum_{\hat{z} \in \Omega_T} (-1)^{W_{\hat{z}} - 1} \sum_{n=1}^{\infty} \frac{\mathrm{tr}(A_{\hat{z}}^n)}{n} t^n.$$

The theorem then clearly follows from (15).

Thus, we are done if we can prove the claim. Suppose that $\mathrm{tr}(A_{\hat{z}})^n \neq 0$ for $\hat{z} \in \Omega_T$. Then, we have $n \equiv 0 \pmod{L_{\hat{z}}}$ by Lemma 3.1(a). Furthermore, $L_{\hat{z}}$ divides T by the definition of $L_{\hat{z}}$. Therefore, $L_{\hat{z}} |\gcd(n, T) = d$, that is, $|\hat{z}^\sharp|$ divides d. It implies that there exists a (unique) $z \in \Omega_d$ such that $z^\sharp = \hat{z}^\sharp$, as desired. ■

The argument above is much simpler than the one used to prove Theorem 4.1 in full generality. Unfortunately, we have not succeeded, up to this point, in extending this argument to the case when T is even, as case-by-case analysis for $(-1)^{dW_z/L_z-1}$ is needed when T has an even divisor d.

5. Comparison between Theorem 4.1 and the known formula

We conclude this paper with a comparison between our formula and the known formula, attributed to Manning and Bowen, for the zeta function of a sofic shift (see [6, Theorem 6.4.8]).

Given a sofic shift $\mathcal{S}$ and its deterministic presentation $\mathcal{G}$ with r states, the Manning-Bowen (M-B) formula also requires us to generate labeled directed graphs $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_t$ (for some $1 \leq t \leq r$) from $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G})$, and compute $\det(I - tA_{\mathcal{G}_i})$, $1 \leq i \leq t$, where $A_{\mathcal{G}_i}$ is the adjacency matrix of $\mathcal{G}_i$. For graph $\mathcal{G}_i$, each state V in $\mathcal{G}_i$ is a subset of $\mathcal{V}_\mathcal{G}$ with cardinality i; thus the graph consists of $\binom{r}{i}$ vertices. Let $a(V)$ be the terminal state of the labeled edge a from V in $\mathcal{G}$, which is uniquely determined, if it exists, since $\mathcal{G}$ is deterministic. For $a \in \Sigma$, an edge labeled a (resp. $-a$) is assigned from $V = \{V_1, V_2, \ldots, V_i\}$ to $U = \{U_1, U_2, \ldots, U_i\}$ if and only if U is obtainable from an even (resp. odd) permutation of $\{a(V_1), a(V_2), \ldots, a(V_i)\}$. Thus, computation of the M-B formula requires a deterministic presentation of the PFT, and further requires construction of auxiliary graphs that are (arguably) not as conceptually simple as our WBG's. On the other hand, our formula of Theorem 4.1 can be computed in a straightforward manner directly from a periodically-forbidden word description of the PFT.

In addition, our formula has a nice recursive property, which is not evident in the M-B formula. Let T and $\widetilde{T}$ be integers such that $T|\widetilde{T}$. Now, let $\mathcal{X} = \mathcal{X}_{\{\mathcal{F},T\}}$ be a PFT in standard form, which we extend to a PFT $\widetilde{\mathcal{X}}$ of period $\widetilde{T}$, by setting $\mathcal{F}^{(j)} = \emptyset$ for $T \leq j \leq \widetilde{T} - 1$. From Theorem 4.1, it is straightforward to verify that the zeta function of $\widetilde{\mathcal{X}}$ factors as

$$\zeta_{\widetilde{\mathcal{X}}}(t) = \zeta_\mathcal{X}(t) \times [\text{additional terms}],$$

where the additional terms are obtained from sequences z in $\Omega_{\widetilde{T}} \setminus \{z^{\widetilde{T}/T} : z \in \Omega_T, N_z \text{ is even }\}$. Thus, if $\zeta_\mathcal{X}(t)$ is already known, then only the additional terms need to be computed. This is particularly useful in situations when an SFT with a known zeta function is extended to a PFT with period greater than 1.

Finally, for T odd, our formula has a nice elegant form, as given in (3). Also, for $T = 2$, the zeta function expression given in Theorem 4.1 (upon observing that $A_{11} = B_{11}$ since $\mathcal{G}_{11} = \mathcal{H}_{11}$) reduces to $\det(I + tA_{11})/\det(I - tA_{10})$. The M-B formula is not so transparent in these cases.

References

[1] M.-P. Béal, M. Crochemore and G. Fici, "Presentations of constrained systems with unconstrained positions," IEEE Trans. Inform. Theory, vol. 51, no.5, pp.1891–1900, May 2005.
[2] M.-P. Béal, M. Crochemore, B.E. Moision and P.H. Siegel, "Periodic finite-type shift spaces," IEEE Trans. Inform. Theory, vol. 57, no.6, pp.3677–3691, June 2011 .
[3] R. Bowen and O.E. Lanford, "Zeta functions of restrictions of the shift transformation," Proc. Symp. Pure Math. A.M.S., vol.14, pp. 43–50, 1970.
[4] P.J. Cameron, Combinatorics: Topics, Techniques, Algorithms, Cambridge University Press, 1995.
[5] D.P.B. Chaves and C. Pimentel, "An algorithm for finding the Shannon cover of a periodic shift of finite type," preprint.
[6] D. Lind and B.H. Marcus, An Introduction to Symbolic Dynamics and Coding, Cambridge University Press, 1995.
[7] A. Manada and N. Kashyap, "A comparative study of periods in a periodic-finite-type shift," SIAM J. Discrete Math., vol. 23, no.3, pp.1507–1524, Oct. 2009.
[8] A. Manning, "Axiom A diffeomorphisms have rational zeta functions," Bull. London Math. Soc., vol.3, pp.215–220, 1971.
[9] B.E. Moision and P.H. Siegel, "Periodic-finite-type shift spaces," Proc. 2001 IEEE Int. Symp. Inform. Theory (ISIT'01), Washington DC, p. 65, June 24–29, 2001.
[10] T.L. Poo and B.H. Marcus, "Time-varying maximum transition run constraints," IEEE Trans. Inform. Theory, vol.52, no.10, pp.4464–4480, Oct. 2006.
[11] J.C. de Souza, B.H. Marcus, R. New and B.A. Wilson, "Constrained systems with unconstrained positions," IEEE Trans. Inform. Theory, vol.48, no.4, pp.866–879, April 2002.

Akiko Manada received the Ph.D. degree in Mathematics from Queen's University, Kingston, Canada, in 2009. She then worked at Claude Shannon Institute at University College Dublin, Ireland, as a postdoctoral fellow from 2009 through 2011. Since February 2012, she has been an assistant professor at Graduate School of Information Systems, the University of Electro-Communications, Tokyo, Japan. Her research interests are discrete mathematics and its applications towards coding theory.

Navin Kashyap received the M.S. degree in Mathematics and the Ph.D. degree in Electrical Engineering from the University of Michigan, Ann Arbor, in 2001. Following a postdoctoral appointment at the University of California San Diego, in 2004, he joined the faculty of the Dept. of Mathematics and Statistics at Queen's University, Kingston, Ontario. Since January 2011, he has been an Associate Professor in the Dept. of Electrical Communication Engineering at the Indian Institute of Science, Bangalore. His research interests lie primarily in the application of combinatorial and probabilistic methods in information and coding theory.