Zeta functions for sofic subshifts

Dennis Pixton*

Definition 1. *A* **directed graph** G *consists of a pair of sets* (V, E) *and a pair of functions* $\text{head} : E \to V$ *and* $\text{tail} : E \to V$. *One says that the edge* $e \in E$ *is* **from** $\text{tail}(e)$ **to** $\text{head}(e)$. G *is finite iff* V *and* E *are finite.*

Definition 2. *A* **labeling** *of the directed graph* G *is a function* $\lambda : E \to \Sigma$, *where* Σ *is a set called the* **alphabet**.

Throughout we shall assume that G is a finite labeled graph.

Definition 3. *For* $j, k \in \mathbb{Z}$, $[j,k] = \{ n \in \mathbb{Z} : j \le n \le k \}$. *A* **finite path** *in* G *is an element* w *of* $E^{[j,k]}$ *for some* j, k *such that, for* $j \le n < k$, $\text{head}(w_n) = \text{tail}(w_{n+1})$. *An* **infinite path** *in* G *is an element* w *of* $E^{\mathbb{Z}}$ *satisfying* $\text{head}(w_n) = \text{tail}(w_{n+1})$ *for all* $n \in \mathbb{Z}$. *We denote by* $\Pi^* = \Pi^*(G)$ *the set of all finite paths in* G *with domain an interval of the form* $[0,n]$ *and by* $\Pi = \Pi(G)$ *the set of all infinite paths in* G.

Definition 4. *If* $w \in E^{[j,k]}$ *is a path in* G *with* $j \le k$ *we define* $\text{tail}(w) = \text{tail}(w_j)$ *and* $\text{head}(w) = \text{head}(w_k)$.

Definition 5. *For* $X \subset \mathbb{Z}$ *we extend* λ *to a map* $E^X \to \Sigma^X$, *also denoted* λ, *by* $\lambda(w)_n = \lambda(w_n)$ *for all* $n \in X$. *We then set* $\Lambda^* = \Lambda^*(G) = \lambda(\Pi^*)$ *and* $\Lambda = \Lambda(G) = \lambda(\Pi)$. *Such* Λ *is called a* **sofic subshift**, *and any labeled graph such that* $\Lambda = \Lambda(G)$ *is called a* **presentation** *of* Λ.

Definition 6. *The shift map* σ *is defined by* $\sigma(w)_n = w_{n+1}$. *This is defined from each of* Π *and* Λ *to itself, and also from paths in* $E^{[j,k]}$ *to paths in* $E^{[j-1,k-1]}$.

Lemma 1.

(1) $\lambda\sigma = \sigma\lambda$.
(2) *For* $x \in \Lambda$, σ *is a bijection of* $\lambda^{-1}(x)$ *onto* $\lambda^{-1}(\sigma x)$.

Definition 7. *A labeled graph is* **Shannon** *iff for any* $e, f \in E$, *if* $\text{tail}(e) = \text{tail}(f)$ *and* $\lambda(e) = \lambda(f)$ *then* $e = f$.

Theorem 1. *Any sofic subshift may be presented by a Shannon graph.*

Proof: This is just the subset construction for converting a nondeterministic automaton to a deterministic one. $\square$

Lemma 2. *If* G *is Shannon then* $|\lambda^{-1}(x)| \le |V|$ *for all* $x \in \Lambda^*$.

Proof: Take $x \in \Lambda^*$. Then $x \in \Sigma^{[0,k]}$ for some k. If $w, z \in \lambda^{-1}(x)$ then $w, z \in E^{[0,k]}$. If $\text{tail}(w_n) = \text{tail}(z_n)$ for some $n \in [0,k]$ then, since $\lambda(w_n) = x_n = \lambda(z_n)$ and G is Shannon, we have $w_n = z_n$, so $\text{head}(w_n) = \text{head}(z_n)$. Hence $\text{tail}(w_{n+1}) = \text{tail}(z_{n+1})$ if $n < k$. By induction we then prove that $\text{head}(w_0) = \text{head}(z_0) \implies w = z$. Therefore tail is an injection from $\lambda^{-1}(x)$ into V. $\square$

*These are notes I prepared for part of a seminar on sofic systems in the summer of 1991. They are largely based on work with a former student of mine, Carol Lawrence, in 1985-6.

LEMMA 3. *If there is K so that $|\lambda^{-1}(x)| \le K$ for all $x \in \Lambda^*$ then $|\lambda^{-1}(x)| \le K$ for all $x \in \Lambda$.*

PROOF: Suppose that there are $K+1$ elements in $\lambda^{-1}(x)$ for some $x \in \Lambda$. Write these as $w^0, \ldots, w^K$. Then there is k such that the finite paths $z^j = w^j|[-k,k]$ are all distinct. Then the elements $\sigma^{-k}(z^j)$ of Λ^* are all distinct, contradicting Lemma 2. □

Throughout this note we assume that λ is **finite to one** onto Λ.

Note that G with the identity labeling is Shannon, and the corresponding sofic subshift is just Π. Hence the various definitions and lemmas that follow apply to Π as well as to Λ.

DEFINITION 8. *A point $x \in \Lambda$ is* **periodic** *iff there is $p > 0$ so that $\sigma^p(x) = x$. Such p is called a* **period** *of x. Denote by $\mathrm{Per}_p(\Lambda)$ the set of all periodic points of Λ with period p, and by $\mathrm{Per}(\Lambda)$ the set of all periodic points of Λ.*

LEMMA 4.

(1) *For $n, p > 0$, x has period n iff x has least period p and $p \mid n$.*
(2) *x is periodic with period n iff for all $k \in \mathbb{Z}$, $x_{k+n} = x_k$.*
(3) *$\lambda(\mathrm{Per}_p(\Pi)) \subset \mathrm{Per}_p(\Lambda)$.*

LEMMA 5. *Suppose $x \in \mathrm{Per}_p(\Lambda)$.*

(1) *For some n, $\lambda^{-1}(x) \subset \mathrm{Per}_{np}(\Pi)$.*
(2) *The least period of each $y \in \lambda^{-1}(x)$ is a multiple of the least period of x.*

PROOF:

(1) From part (2) of lemma 1, σ^p is a bijection of $\lambda^{-1}(x)$ onto itself, and, since this is finite, σ^{pn} is the identity on $\lambda^{-1}(x)$ for some $n \in \mathbb{N}$.
(2) If w has least period q then $\sigma^q(x) = \sigma^q(\lambda(w)) = \lambda(\sigma^q(w)) = \lambda(w) = x$ so q is a period of x. Hence the least period of x divides q.

□

LEMMA 6. *If $x \in \Lambda$ is periodic then for any r the map $t_r : w \mapsto \mathrm{tail}(w_r)$ is an injection of $\lambda^{-1}(x)$ into V.*

PROOF: Replacing x with $\sigma^r(x)$, we may assume that $r = 0$. Suppose $z, w \in \lambda^{-1}(x)$, $\mathrm{tail}(z_0) = \mathrm{tail}(w_0)$, and $z \ne w$. Let q be a common period for z and w. For each $m > 1$ define $y^m \in E^{\mathbb{Z}}$ by

$$y_j^m = \begin{cases} z_j & \text{if } 0 \le j < (m-1)q \\ w_j & \text{if } (m-1)q \le j < mq \end{cases}$$

and by $y_{j+mqt}^m = y_j^m$ otherwise. Since $\mathrm{head}(z_{kq-1}) = \mathrm{head}(z_{-1}) = \mathrm{tail}(z_0) = \mathrm{tail}(w_0)$, and similarly with w and z interchanged, we have $y^m \in \Pi$. Also, $\lambda(y^m) = x$ is clear. Since $y^m \ne y^n$ for $m \ne n$ we have contradicted the assumption that $\lambda^{-1}(x)$ is finite. □

DEFINITION 9. *If X is a set then $\mathcal{P}_n(X) = \{\, Y \subset X : |Y| = n \,\}$.*

For $n \in \mathbb{N}$ define labeled graphs $G^n = (V^n, E^n, \lambda)$ as follows. First, $V^n = \mathcal{P}_n(V)$. Next, an element W of E^n will be a subset of E with the following properties:

(1) $W \in \mathcal{P}_n(E)$;
(2) the sets $\text{head}(W) = \{\, \text{head}(e) : e \in W \,\}$ and $\text{tail}(W) = \{\, \text{tail}(e) : e \in W \,\}$ are in V^n;
(3) $\lambda(e) = \lambda(f)$ for all $e, f \in W$.

The above defines head and tail, and $\lambda(W)$ is just $\lambda(e)$ for any $e \in W$. Note that (2) is equivalent to the requirement that tail and head are injective on W.

Lemma 7.

(1) *G^n is empty for $n > |V|$.*
(2) *G^1 is isomorphic to G.*
(3) *$\Lambda(G^n) \subset \Lambda(G)$.*
(4) *If G is Shannon then so is G^n.*

If $W \in E^n$ there is a bijection $\phi_W : \text{tail}(W) \to \text{head}(W)$ defined by

$$\overbrace{\text{tail}(W) \xrightarrow{\ \text{tail}\ } W \xrightarrow{\ \text{head}\ } \text{head}(W)}^{\phi_W}$$

Also, if $\langle W, Z \rangle$ is a path of length 2 in G^n we have a bijection $\psi_{WZ} : W \to Z$ defined by

$$\overbrace{W \xrightarrow{\ \text{head}\ } \text{head}\,W = \text{tail}(Z) \xrightarrow{\ \text{tail}\ } Z}^{\psi_{WZ}}$$

Finally, for a finite path $W = \langle W_j, W_{j+1} \ldots, W_k \rangle$ we define $\phi_W : \text{tail}(W) \to \text{head}(W)$ and $\psi_W : W_j \to W_k$ by composition: $\phi_W = \phi_{W_k} \circ \cdots \circ \phi_{W_j}$ and $\psi_W = \psi_{W_{k-1} W_k} \circ \cdots \circ \psi_{W_j W_{j+1}}$.

Lemma 8. *There is an injection $\alpha : \Pi(G^n) \to \mathcal{P}_n(\Pi(G))$ such that, for all $W \in \Pi(G^n)$,*

(1) *$W_k = \{\, w_k : w \in \alpha(W) \,\}$ for all $k \in \mathbb{Z}$;*
(2) *$\alpha(\sigma(W)) = \sigma(\alpha(W))$;*
(3) *$\lambda(W) = \lambda(w)$ for all $w \in \alpha(W)$.*

Proof: Take $W \in \Pi(G^n)$. For $k \in \mathbb{Z}$ define Z_k as the restriction of W to $[0, k]$ if $k \geq 0$, or to $[k, 0]$ if $k < 0$. Given $w_0 \in W_0$, define a path $w \in E^{\mathbb{Z}}$ by

$$w_k = \begin{cases} \psi_{Z_k}(w_0) & \text{if } k > 0 \\ \psi_{Z_k}^{-1}(w_0) & \text{if } k < 0. \end{cases}$$

It is immediate from the definition of ψ that $\text{head}(w_k) = \text{tail}(w_{k+1})$, so w is in $\Pi(G)$. We define $\alpha(W) = \{\, w : w_0 \in W_0 \,\}$. Then the listed properties are clear. The elements of $\alpha(W)$ are distinct since they are distinct in their 0^{th} positions, so $\alpha(W) \in \mathcal{P}_n(\Pi(G))$. Finally, if $W \neq Z$ then $W_k \neq Z_k$ for some k, and it follows from (1) that $\alpha(W) \neq \alpha(Z)$. Hence α is injective. $\square$

Lemma 9. *α is a bijection from $\operatorname{Per}_p(\Pi(G^n))$ onto $\mathcal{L}_p^n = \{ L \in \mathcal{P}_n(\operatorname{Per}(\Pi)) : \sigma^p(L) = L$ and $\lambda(w) = \lambda(z)$ for all $w, x \in L \}$.*

Proof: If $\sigma^p(W) = W$ then $\sigma^p(\alpha(W)) = \alpha(W)$ by part (2) of lemma 8, so $\alpha(W) \in \mathcal{L}_p^n$. On the other hand, if $L \in \mathcal{L}_p^n$ we define $W_k = \{ w_k : w \in L \}$. By Lemma 6 each $\operatorname{tail}(W_k)$ has n elements. Hence $|W_k| = n$. Clearly $\operatorname{head}(W_k) = \operatorname{tail}(W_{k+1})$, so $\operatorname{head}(W_k)$ has n elements and $W \in \Pi(G^n)$. Finally, $W_{k+p} = \{ w_{k+p} : w \in L \} = \{ \sigma^p(w)_k : w \in L \} = \{ z_k : z \in \sigma^p(L) \} = W_k$ (using $\sigma^p(L) = L$). Hence $W \in \operatorname{Per}_p(\Pi(G^n))$, and $\alpha(W) = L$ is clear. □

We shall need to attach signs to various bijections. We denote by $\mathcal{S}_N$ the symmetric group on N symbols, which we think of as the group of bijections of $[1, N]$ onto itself. We denote by sgn the sign homomorphism from $\mathcal{S}_N$ into $\{1, -1\}$.

Definition 10. *If $f : X \to X$ is a bijection, $N = |X|$, and $\tau : [1, N] \to X$ is any bijection, we define $\operatorname{sgn}(f) = \operatorname{sgn}(\tau^{-1} f \tau)$.*

Lemma 10. *$\operatorname{sgn}(f)$ is independent of the choice of bijection τ.*

Proof: Let ρ be another bijection of $[1, N]$ onto X. Then $\pi = \tau^{-1}\rho \in \mathcal{S}_N$ and $\rho^{-1} f \rho = \pi^{-1}(\tau^{-1} f \tau)\pi$. The lemma follows since conjugate permutations have the same sign. □

Lemma 11. *If X is a nonempty finite set and $f : X \to X$ is a bijection then*

$$\sum_{\substack{L \subset X \\ f(L) = L}} (-1)^{|L|} \operatorname{sgn}(f|L) = 0.$$

Proof: First we note that the formula is true if f is transitive. In this case the only invariant subsets are $\emptyset$ and X itself, and $X \neq \emptyset$. Also, $\operatorname{sgn}(f|\emptyset) = 1$ and $\operatorname{sgn}(f|X)$ is the sign of an N-cycle, which is $(-1)^{N+1}$. Hence the sum reduces to $(-1)^0 \cdot 1 + (-1)^N \cdot (-1)^{N+1} = 0$.

We proceed by induction on $N = |X|$. The starting case, $N = 1$, follows since f is transitive. So we assume that the formula is true for bijections of sets of cardinality less than N.

We may assume that f is not transitive. Hence we can write X as the union of two disjoint, non-empty, f-invariant sets Y and Z. We let $g = f|Y$ and $h = f|Z$. Then the formula applies to both g and h. Moreover, $L \subset X$ is f-invariant iff it is the union of a g-invariant set $A \subset Y$ and an h-invariant set $B \subset Z$. In this case we have $\operatorname{sgn}(f|L) = \operatorname{sgn}(g|A)\operatorname{sgn}(h|B)$. Therefore

$$\begin{aligned}
\sum_{\substack{L \subset X \\ f(L) = L}} (-1)^{|L|} \operatorname{sgn}(f|L) &= \sum_{\substack{A \subset Y \\ g(A) = A}} \sum_{\substack{B \subset Z \\ h(B) = B}} (-1)^{|A|+|B|} \operatorname{sgn}(g|A)\operatorname{sgn}(h|B) \\
&= \left(\sum_{\substack{A \subset Y \\ g(A) = A}} (-1)^{|A|} \operatorname{sgn}(g|A) \right) \left(\sum_{\substack{B \subset Z \\ h(B) = B}} (-1)^{|B|} \operatorname{sgn}(h|B) \right) \\
&= 0 \cdot 0 = 0.
\end{aligned}$$

□

We define $N_p = N_p(\Lambda) = |\operatorname{Per}_p(\Lambda)|$. For $L \in \mathcal{L}_p^n$ we define $\operatorname{sgn}(L) = \operatorname{sgn}(\sigma^p|L)$.

Lemma 12.

$$N_p = \sum_{n=1}^{|V|} (-1)^{n+1} \sum_{L \in \mathcal{L}_p^n} \operatorname{sgn}(L).$$

Proof: For a given $x \in \mathrm{Per}_p(\Lambda)$ let $X = \lambda^{-1}(x)$ and let $f = \sigma^p|X$. Then $|X| \le |V|$ by Lemma 6 and $X \ne \emptyset$ so Lemma 11 applies. Writing the term for $L = \emptyset$ on the left side and the other terms on the right gives

$$1 = \sum_{n=1}^{|V|} (-1)^{n+1} \sum_{L \in \mathcal{L}_p^n \cap \lambda^{-1}(x)} \operatorname{sgn}(L).$$

The formula for N_p now follows by summing this for all $x \in \mathrm{Per}_p(\Lambda)$. $\square$

Now we attach signs to the edges of G^n. We choose for each $S \in V^n$ a bijection $\tau_S : [1,n] \to S$. Then if $W \in E^n$ we define $\operatorname{sgn}(W) = \operatorname{sgn}(\tau_T^{-1} \phi_W \tau_S)$ where $S = \operatorname{tail}(W)$ and $T = \operatorname{head}(W)$. This definition depends on the choices of the τ's. For a finite path $W = \langle W_j, \ldots, W_k \rangle$ we define $\operatorname{sgn}(W) = \operatorname{sgn}(W_k)\operatorname{sgn}(W_{k-1}) \cdots \operatorname{sgn}(W_j)$. If $S_r = \operatorname{tail}(W_r)$ and $T_r = \operatorname{head}(W_r)$ then, since $S_r = T_{r-1}$,

$$\begin{aligned}
\operatorname{sgn}(W) &= \operatorname{sgn}(\tau_{T_k}^{-1} \phi_{W_k} \tau_{S_k}) \operatorname{sgn}(\tau_{T_{k-1}}^{-1} \phi_{W_{k-1}} \tau_{S_{k-1}}) \cdots \operatorname{sgn}(\tau_{T_j}^{-1} \phi_{W_j} \tau_{S_j}) \\
&= \operatorname{sgn}(\tau_{T_k}^{-1} \phi_{W_k} \tau_{S_k} \tau_{T_{k-1}}^{-1} \phi_{W_{k-1}} \tau_{S_{k-1}} \cdots \tau_{T_j}^{-1} \phi_{W_j} \tau_{S_j}) \\
&= \operatorname{sgn}(\tau_{T_k}^{-1} \phi_{W_k} \phi_{W_{k-1}} \cdots \phi_{W_j} \tau_{S_j}) \\
&= \operatorname{sgn}(\tau_{T_k}^{-1} \phi_W \tau_{S_j}).
\end{aligned}$$

In particular, if $\operatorname{head}(W) = \operatorname{tail}(W)$ then $\operatorname{sgn}(W) = \operatorname{sgn}(\phi_W)$.

We define

$$N_p^n = \sum_{W \in \mathrm{Per}_p(\Pi(G^n))} \operatorname{sgn}(\langle W_0, \ldots, W_{p-1} \rangle).$$

Lemma 13.

$$N_p = \sum_{n=1}^{|V|} (-1)^{n+1} N_p^n.$$

Proof: This is just a combination of Lemma 12 and Lemma 9, since the bijection α preserves signs. Specifically, with $Z = \langle W_0, \ldots, W_{p-1} \rangle$ and $V = \langle W_0, \ldots, W_p \rangle$ the following diagram of bijections commutes:

$$\begin{array}{ccccc}
\operatorname{tail}(W_0) & \xrightarrow{\ \phi_Z\ } & \operatorname{head}(W_{p-1}) & = & \operatorname{tail}(W_0) \\
\Big| \scriptstyle\text{tail} & & & & \Big| \scriptstyle\text{tail} \\
W_0 & \xrightarrow{\ \psi_V\ } & W_p & = & W_0 \\
\Big| \scriptstyle t_0 & & & & \Big| \scriptstyle t_0 \\
\alpha(W) & & \xrightarrow{\ \sigma^p\ } & & \alpha(W).
\end{array}$$

Hence the maps $\phi_Z : \operatorname{tail}(W_0) \to \operatorname{tail}(W_0)$ and $\sigma^p : \alpha(W) \to \alpha(W)$ are conjugate, so they have the same signs. $\square$

Definition 11. *For a sequence $s = \langle c_1, c_2, \dots \rangle$ of numbers we define the* **zeta function** *of c as the formal power series*

$$\zeta_c(t) = \exp\left(\sum_{p=1}^{\infty} \frac{c_p}{p}\right).$$

Using the sequence N of numbers of periodic points, we define the zeta function of Λ as $\zeta_\Lambda = \zeta_N$.

Lemma 14. *For sequences a and b:*

(1) $\zeta_{a+b} = \zeta_a \zeta_b$.
(2) $\zeta_{-a} = \zeta_a^{-1}$.
(3) *If there is a number m such that $a_p = m^p$ for all p then $\zeta_a(t) = (1 - tm)^{-1}$.*
(4) *If there is a square matrix M such that $a_p = \operatorname{tr}(M^p)$ for all p then $\zeta_a(t) = (\det(I - tM))^{-1}$.*

Proof: (1) and (2) are trivial. (3) follows from the series

$$\sum_{p=1}^{\infty} \frac{(mt)^p}{p} = -\ln(1 - mt),$$

which is obtained by integrating the geometric series. Finally, (4) follows since $\operatorname{tr}(M^p) = \sum m_i^p$ and $\det(I - tM) = \prod(1 - tm_i)$ where m_i is the list of eigenvalues of M, repeated according to multiplicity. □

We define matrices M_n indexed by the elements of V^n as follows:

$$(M_n)_{ST} = \sum\{\operatorname{sgn}(W) : W \in E^n,\ \operatorname{head}(W) = S,\ \operatorname{tail}(W) = T\}.$$

Lemma 15.

$$N_p^n = \operatorname{tr}(M_n^p).$$

Proof: This follows immediately from the following:

$$(M_n^p)_{ST} = \sum\{\operatorname{sgn}(W) : W \in \Pi^*(G^n) \text{ of length } p,\ \operatorname{head}(W) = S,\ \operatorname{tail}(W) = T\}.$$

And this formula follows by induction. It is true by definition for $p = 1$. If we take $W = \langle W_0, \dots, W_p \rangle$ of length $p+1$ from S to T and we define $V = \langle W_0, \dots, W_{p-1} \rangle$ and $R = \operatorname{head}(V)$ then $\operatorname{sgn}(W) = \operatorname{sgn}(V)\operatorname{sgn}(W_p)$. Conversely, if $V = \langle V_0, \dots, V_{p-1} \rangle$ is a path of length p from S to some vertex R and Z is any edge from R to T, we define $W = \langle V_0, \dots, V_{p-1}, Z \rangle$ from S to T and we have $\operatorname{sgn}(W) = \operatorname{sgn}(V)\operatorname{sgn}(Z)$. Thus

$$\begin{aligned}
(M_n^{p+1})_{ST} &= \sum_R (M_n^p)_{SR} (M_n)_{RS} \\
&= \sum_R \sum_{V \text{ from } S \text{ to } R} \operatorname{sgn}(V) \sum_{Z \text{ from } R \text{ to } T} \operatorname{sgn}(Z) \\
&= \sum_R \sum_{\substack{V \text{ from } S \text{ to } R \\ \text{and} \\ Z \text{ from } R \text{ to } T}} \operatorname{sgn}(V)\operatorname{sgn}(Z) \\
&= \sum_{W \text{ from } S \text{ to } T} \operatorname{sgn}(W)
\end{aligned}$$

as desired. □

THEOREM 2. *The zeta function of a sofic subshift is a rational function.*

PROOF: In fact, from Lemmas 13, 15, and 14 we have the formula

$$\zeta_\Lambda(t) = \zeta_N(t) = \prod_{n=1}^{|V|} \zeta_{N^n}(t)^{(-1)^{n+1}} = \prod_{n=1}^{|V|} (\det(I - tM_n))^{(-1)^n}.$$

□