GIBBS INTERTWINING OPERATORS AND THE STEINBERG POLYNOMIAL

YIPIN WANG

ABSTRACT. We study the algebraic structure of the Markov operator P arising from spanning tree enumeration on $\mathbb{P}^1(\mathbb{F}_p)$. We show that P factors as $L_w \cdot \pi(w_0)$, where w_0 is the long Weyl element of GL_2 and L_w is a Gibbs-weighted average over the unipotent radical $U(\mathbb{F}_p)$, with weights $w_r = q^{p-r}/(q^p-1)$. This identifies P as a deformed intertwining operator. We prove that P does not belong to the Iwahori–Hecke algebra.

The main new result is a twisted circulant reduction: the Steinberg polynomial $n_p(q)$ is expressed as $n_p(q) = -(q-1)(q^p-1)\det(I-C)$, where C is a $(p-1)\times(p-1)$ matrix on $\mathbb{F}_p^*$ whose (j,j')-entry $w_{j'-j^{-1} \bmod p}$ mixes the additive structure of the Gibbs weights with the multiplicative structure of inversion in $\mathbb{F}_p^*$. The reduction proceeds via two identities: (i) the boundary state decouples from the determinant (Schur complement equals 1), and (ii) a rank-one correction from the ∞-state contributes a factor $(1-q)$ governed by the identity $\mathbf{w}^T(I-C)^{-1}\mathbf{1} = -q$.

The resulting master formula $n_p(q) = -(q-1)(q^p-1)\det(I-C)$ gives a structural explanation for the divisibility $(q-1) \mid n_p(q)$. We prove the $-q$ identity in full generality: first for the untwisted convolution W using the spectral theory of the full circulant on $\mathbb{F}_p$, and then for $C = QW$ via a telescoping argument that identifies the transpose resolvent $(I-W^T)^{-1}\mathbf{w}$ as a delta function supported at the inversion-fixed point $-1 \in \mathbb{F}_p^*$.

We prove that the eigenvalues of C split sharply at the threshold $|1-\lambda| = 1$, with the $(p+1)/2$ "small" eigenvalues in the Q-even sector and the $(p-3)/2$ "large" eigenvalues in the Q-odd sector. The proof uses the multiplicative character basis, in which Q decomposes C into 2×2 blocks indexed by orbits $\{k,-k\}$; the off-diagonal (even-odd) mixing is bounded by $O(1/\sqrt{p})$ via the Weil bound for Gauss sums $|G(a,\chi)| = \sqrt{p}$, while the diagonal gap between sectors is $2\mu > 0$, independent of the block. This spectral gap is the finite-field analogue of the Ramanujan bound for expander graphs.

At $q = 2$, we prove that $n_p(2) \equiv 1 \pmod 8$ for all primes $p \geq 7$. The proof exploits the sparse structure of the integer matrix $(2^p-1)(I-C)$ modulo 8: only the shift-involution maps $\sigma_k\colon j \mapsto j^{-1}-k$ for $k = 1,2$ survive, with discriminants $\Delta_k = k^2+4$ from the discriminant partition. The congruence rests on two arithmetic identities—a universal involutory identity $\mathrm{tr}(A_k^2) = \mathrm{tr}(A_k)$ for all $k \geq 1$ with $p \nmid k$, and quadratic discreteness $\mathrm{tr}(A_k) \in \{0,2\}$—both consequences of the group structure of $\mathbb{F}_p^*$. We conjecture that higher 2-adic digits of $n_p(2)$ are governed by a cascade of Legendre symbols (Δ_k/p).

We establish concrete connections between the twisted circulant determinant and the Ruelle zeta function of the mod-p continued fraction dynamics. The trace $\mathrm{tr}(C^n)$ counts weighted n-step closed paths within $\mathbb{F}_p^*$, and the boundary decoupling theorem acquires a dynamical interpretation: it identifies the correction from paths that cross the boundary state 0. A functional equation under $q \mapsto 1/q$ produces the palindromic/anti-palindromic splitting of the endoscopic decomposition.

Finally, we analyze the sign of the leading coefficient. The factorization $\det(I-C) = \det(Q)\det(Q-W)$ separates the Weyl sign $(-1/p)$ from the arithmetic content $\det(Q-W)$, and we prove that $\mathrm{sign}(\det(Q-W))$ is independent of q. This reduces the sign formula $\mathrm{sign}(\mathrm{lead}(n_p)) = -(-2/p)$ to a combinatorial identity about the cancellation structure in $\det(Q-W)$. We also prove that the p-adic valuation of the exponent matrix determinant is exactly $p-3$.

CONTENTS

Date: February 2026.

1. Introduction

Let p be an odd prime and q a prime power. In [8], the author introduced the *Steinberg polynomial*

$$n_p(q) = (q^p - 1)\det\Big(I - P(q)\Big|_{\mathrm{St}_p}\Big) \in \mathbb{Z}[q],$$

where $P(q)$ is the transition matrix of a Markov chain on $\mathbb{P}^1(\mathbb{F}_p)$ with weights $w_r = q^{p-r}/(q^p - 1)$ and St_p is the p-dimensional Steinberg representation of $\mathrm{GL}_2(\mathbb{F}_p)$. The polynomial $n_p(q)$ was shown (computationally, for all primes $p \leq 97$) to admit an *endoscopic decomposition*

$$n_p(q) = n_p^{\mathrm{GL}_2}(q) - \left(\frac{-2}{p}\right) n_p^T(q), \tag{1}$$

where $n_p^{\mathrm{GL}_2}$ is palindromic, n_p^T is anti-palindromic, and $\left(\frac{-2}{p}\right)$ is the Legendre symbol, together with a motivic factorization into CM abelian varieties over $\mathbb{F}_2$ with CM by subfields of $\mathbb{Q}(\sqrt{-2}, \zeta_p)$.

The Steinberg polynomial first arose in connection with the gate complexity model of [9], where the $n = 2$ specialization of the gate complexity $t(p, q, n)$ of the algebraic torus $(\mathbb{F}_q^*)^n$ produces a random walk on $\mathbb{P}^1(\mathbb{F}_p)$ whose spectral theory is governed by $n_p(q)$.

The present paper addresses the question: what algebraic structure of P produces the endoscopic decomposition? We identify the correct algebraic framework in three stages. First (§2–§3), we prove that P factors as $L_w \cdot \pi(w_0)$ but does not lie in the Iwahori–Hecke algebra. Second (§4), we establish the twisted circulant reduction that isolates the arithmetic content in a single $(p-1)\times(p-1)$ matrix C on $\mathbb{F}_p^*$ whose entries $C[j,j'] = w_{j'-j^{-1}}$ mix additive and multiplicative structures. Third (§9), we make precise the connection between $\det(I-C)$ and the Ruelle zeta function of the mod-p continued fraction dynamics, establishing a trace formula, a dynamical interpretation of boundary decoupling, and a functional equation. Fourth (§10), we analyze the sign of the leading coefficient of $n_p(q)$: the factorization $\det(I-C) = \det(Q)\det(Q-W)$ separates the Weyl sign from the arithmetic content, and we prove that $\mathrm{sign}(\det(Q-W))$ is independent of q, reducing the sign formula to a combinatorial identity about the inversion permutation.

1.1. **Main results.**

Theorem 1.1 (Factorization). *Let $w_0 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ be the long Weyl element and $U(r) = \left(\begin{smallmatrix} 1 & r \\ 0 & 1 \end{smallmatrix}\right)$ for $r \in \mathbb{F}_p$. Define the matrices $S_r = w_0 \cdot U(r) = \left(\begin{smallmatrix} 0 & 1 \\ 1 & r \end{smallmatrix}\right)$. Then*

$$P = \sum_{r=0}^{p-1} w_r\, \pi(S_r) = L_w \cdot \pi(w_0),$$

where $L_w = \sum_{r=0}^{p-1} w_r\, \pi(U(r))$ is the Gibbs-weighted average over the unipotent radical acting on $\mathbb{P}^1(\mathbb{F}_p)$, and π denotes the natural permutation representation.

Theorem 1.2 (Non-Hecke). *Let A denote the image of the standard Hecke operator (uniform transition on $\mathbb{P}^1(\mathbb{F}_p)$) acting on St_p. Then:*

1. *$P|_{\mathrm{St}_p}$ and $A|_{\mathrm{St}_p}$ do not commute.*
2. *$P|_{\mathrm{St}_p}$ is not a polynomial in $A|_{\mathrm{St}_p}$.*
3. *$\dim \mathbb{Q}[A]|_{\mathrm{St}_p} = 2$ (minimal polynomial $x(x+1/p)$), while $\dim \mathbb{Q}[P]|_{\mathrm{St}_p} = p$ (all eigenvalues distinct).*

In particular, $P|_{\mathrm{St}_p}$ does not belong to the commutant of the Hecke algebra in $\mathrm{End}(\mathrm{St}_p)$.

Theorem 1.3 (Twisted circulant reduction). *Define the $(p-1)\times(p-1)$ matrix C on $\mathbb{F}_p^*$ by*

$$C[j,j'] = w_{j'-j^{-1} \bmod p} = \frac{q^{p-(j'-j^{-1} \bmod p)}}{q^p - 1}, \quad j, j' \in \mathbb{F}_p^*.$$

Then $n_p(q) = -(q-1)(q^p-1)\det(I-C)$. More precisely, if $P|_{\mathrm{St}_p}$ is written in the basis $\{e_i - e_\infty\}_{i=0}^{p-1}$ and the block decomposition

$$I - P|_{\mathrm{St}_p} = \begin{pmatrix} A_{00} & A_{0B} \\ A_{B0} & A_{BB} \end{pmatrix}$$

separates the boundary state 0 from the bulk states $\{1,\ldots,p-1\}$, then:

1. (Boundary decoupling) *The Schur complement of A_{00} equals 1 exactly, so $\det(I - P|_{\mathrm{St}_p}) = \det(A_{BB})$.*
2. (Rank-one correction) *The bulk block decomposes as $P_{BB} = C - R_\infty$ where $R_\infty = \mathbf{1}\cdot \mathbf{w}^T$ is rank-one. The matrix determinant lemma gives $\det(I - P_{BB}) = (1-q)\det(I-C)$, equivalent to the identity $\mathbf{w}^T(I-C)^{-1}\mathbf{1} = -q$.*

Corollary 1.4. *The polynomial $n_p(q)$ is divisible by $(q-1)$ for all odd primes p.*

Theorem 1.5 (Lattice index). *For all primes $p \le 23$, the Smith normal form of the integer matrix $A_p = (2^p-1)(I - P|_{\mathrm{St}_p})$ has elementary divisors with stripped product $|n_p(2)|$, and alien primes concentrate in the last elementary divisor.*

Theorem 1.6 (2-adic congruence). *For all primes $p \ge 7$, $n_p(2) \equiv 1 \pmod 8$.*

2. The Gibbs intertwining operator

2.1. **Definitions.** Let $G = \mathrm{GL}_2(\mathbb{F}_p)$, $B = \left\{\left(\begin{smallmatrix} * & * \\ 0 & * \end{smallmatrix}\right)\right\}$ the upper Borel, and $U = \left\{\left(\begin{smallmatrix} 1 & r \\ 0 & 1 \end{smallmatrix}\right) : r \in \mathbb{F}_p\right\}$ its unipotent radical. The flag variety $G/B \cong \mathbb{P}^1(\mathbb{F}_p)$ has $p+1$ points. Let $w_0 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ denote the representative of the nontrivial Weyl group element; it has $\det(w_0) = -1$.

Definition 2.1. Fix a parameter $\beta \geq 0$ and a prime power q. The *Gibbs intertwining operator* is

$$M_\beta(w_0) = \frac{1}{q^{\beta p} - 1} \sum_{r=0}^{p-1} q^{\beta(p-r)}\, \pi(w_0 \cdot U(r)) = \frac{1}{q^{\beta p} - 1} \sum_{r=0}^{p-1} q^{\beta(p-r)}\, \pi(S_r),$$

acting on $\mathbb{C}[\mathbb{P}^1(\mathbb{F}_p)]$ via the natural permutation representation π.

The operator $P(q)$ from [8] is $M_1(w_0)$. The limiting case $\beta \to 0$ gives the standard (uniform) intertwiner

$$M_0(w_0) = \frac{1}{p} \sum_{r=0}^{p-1} \pi(w_0 \cdot U(r)),$$

which is an element of the Hecke algebra $\mathbb{C}[B\backslash G/B]$.

Remark 2.2. In the classical theory of intertwining operators for GL_2 over a local field F, the standard intertwiner is the integral $M(w_0, s) = \int_{U(F)} \pi_s(w_0 u)\, du$ against Haar measure on U. Our Gibbs intertwiner replaces Haar measure with the Gibbs measure $d\mu_\beta(u) = q^{\beta \cdot \mathrm{ht}(u)}\, du$ for a height function $\mathrm{ht}\colon U(\mathbb{F}_p) \to \mathbb{Z}$ defined by $\mathrm{ht}(U(r)) = p - r$. This height function depends on the identification $U(\mathbb{F}_p) \cong \mathbb{F}_p$ via the Teichmüller representatives $\{0, 1, \ldots, p-1\}$.

2.2. **Proof of Theorem 1.1.** The factorization $S_r = w_0 \cdot U(r)$ is immediate:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & r \end{pmatrix} = S_r.$$

Since π is a group homomorphism, $\pi(S_r) = \pi(w_0) \cdot \pi(U(r))$, so $P = \pi(w_0) \cdot L_w$ (operator convention) or $P = L_w \cdot \pi(w_0)$ (transition matrix convention).

2.3. **The unipotent action in coordinates.** In the reciprocal coordinate $t = 1/j$ for $j \in \mathbb{F}_p^*$, the unipotent element $U(r)$ acts by translation: $t \mapsto t + r$. Thus L_w restricts to a convolution operator on $\mathbb{F}_p^*$ with Fourier eigenvalues

$$\hat{w}(a) = \sum_{r=0}^{p-1} w_r\, \zeta_p^{ar} = \frac{q}{q - \zeta_p^a} \qquad (a = 1, \ldots, p-1),$$

where $\zeta_p = e^{2\pi i/p}$. The boundary states $\{0, \infty\}$ break the $\mathbb{Z}/p\mathbb{Z}$ symmetry and are responsible for the deviation of the eigenvalues of $P|_{\mathrm{St}_p}$ from the Fourier eigenvalues $q/(q - \zeta_p^a)$.

3. Proof that P is not Hecke

3.1. **The Hecke operator on St_p.**

Proposition 3.1. *$A|_{\mathrm{St}_p}$ has eigenvalue $-1/p$ with multiplicity 1 and eigenvalue 0 with multiplicity $p-1$. In particular, $\mathbb{Q}[A|_{\mathrm{St}_p}]$ is 2-dimensional.*

In contrast, $P|_{\mathrm{St}_p}$ has all p eigenvalues distinct (verified for $p \leq 97$), so $\mathbb{Q}[P|_{\mathrm{St}_p}]$ is p-dimensional.

Proposition 3.2. *For all primes $5 \leq p \leq 97$: $[P|_{\mathrm{St}_p}, A|_{\mathrm{St}_p}] \neq 0$, and $\dim \mathbb{Q}[P, A]|_{\mathrm{St}_p} = 2(p-1)$.*

4. The twisted circulant reduction

This section contains the main new results. We show that the $(p+1)$-dimensional computation of $\det(I - P|_{\mathrm{St}_p})$ reduces exactly to a $(p-1)$-dimensional determinant involving a single matrix C on $\mathbb{F}_p^*$ whose structure mixes the additive and multiplicative structures of the finite field.

4.1. **Block decomposition and boundary decoupling.** Write the p-dimensional Steinberg space in the basis $f_i = e_i - e_\infty$ for $i = 0, 1, \ldots, p-1$, and separate state 0 (boundary) from $\{1, \ldots, p-1\}$ (bulk). The matrix of $I - P|_{\mathrm{St}_p}$ in this basis has the block form

$$I - P|_{\mathrm{St}_p} = \begin{pmatrix} A_{00} & A_{0B} \\ A_{B0} & A_{BB} \end{pmatrix}$$

where A_{00} is a scalar, A_{0B} is a row vector of length $p-1$, A_{B0} is a column vector, and A_{BB} is $(p-1) \times (p-1)$.

From the transition structure of P: every S_r sends $0 \mapsto \infty$, so $P_{\mathrm{full}}[0, j'] = 0$ for all $j' \neq \infty$. In the Steinberg basis, this gives $P_{\mathrm{St}}[0, j'] = -P_{\mathrm{full}}[\infty, j'] = -w_{j'}$ for $j' \in \{0, \ldots, p-1\}$. In particular:

$$A_{00} = 1 + w_0 = 1 + \frac{q^p}{q^p - 1} = \frac{2q^p - 1}{q^p - 1}, \tag{2}$$

$$A_{0B}[j'] = w_{j'} \qquad (j' = 1, \ldots, p-1). \tag{3}$$

Proposition 4.1 (Boundary decoupling)**.** *The Schur complement of A_{00} in $I - P|_{\mathrm{St}_p}$ equals 1:*

$$A_{00} - A_{0B}\, A_{BB}^{-1}\, A_{B0} = 1.$$

Consequently, $\det(I - P|_{\mathrm{St}_p}) = \det(A_{BB})$.

Proof. The operator P acts on $\mathbb{C}[\mathbb{P}^1(\mathbb{F}_p)]$ with constant row sums $\sum_{r=0}^{p-1} w_r = q/(q-1)$, so $P \cdot \mathbf{1}_{\mathrm{full}} = \frac{q}{q-1}\mathbf{1}_{\mathrm{full}}$. In the Steinberg basis $\{f_i = e_i - e_\infty\}$, the all-ones vector $\mathbf{1}_{\mathrm{St}} = (1, \ldots, 1)^T$ satisfies

$$(I - P|_{\mathrm{St}_p}) \cdot \mathbf{1}_{\mathrm{St}} = \big(\alpha, 1, \ldots, 1\big)^T, \quad \alpha = \frac{2q-1}{q-1},$$

since for $i \geq 1$ the row sum of $P|_{\mathrm{St}_p}$ is $q/(q-1) - q/(q-1) = 0$ (both $P_{\mathrm{full}}[i, \cdot]$ and $P_{\mathrm{full}}[\infty, \cdot]$ sum to $q/(q-1)$ over $\mathbb{F}_p$), while row 0 has sum $0 - q/(q-1) = -q/(q-1)$.

In block form: $A_{00} + A_{0B} \cdot \mathbf{1}_B = \alpha$ and $A_{B0} + A_{BB} \cdot \mathbf{1}_B = \mathbf{1}_B$. From the second equation, $A_{BB}^{-1} A_{B0} = A_{BB}^{-1}\mathbf{1}_B - \mathbf{1}_B$. By the Sherman–Morrison formula applied to $A_{BB} = (I - C) + \mathbf{1} \cdot \mathbf{w}^T$:

$$\mathbf{w}^T A_{BB}^{-1} \mathbf{1}_B = \frac{b}{1+b}, \quad b = \mathbf{w}^T (I - C)^{-1} \mathbf{1}.$$

Using $\mathbf{w}^T \mathbf{1}_B = \alpha - A_{00}$ from the first block equation, one computes

$$S = A_{00} - A_{0B}\, A_{BB}^{-1}\, A_{B0} = \alpha - \frac{b}{1+b}.$$

Substituting $b = -q$ (Proposition 4.5) and $\alpha = (2q-1)/(q-1)$:

$$S = \frac{2q-1}{q-1} - \frac{-q}{1-q} = \frac{2q-1}{q-1} - \frac{q}{q-1} = 1. \qquad \square$$

Remark 4.2. The identity $S = 1$ says that the boundary state 0 contributes nothing to the Steinberg determinant. Since $0 \mapsto \infty$ under every S_r, and ∞ is already projected out in the Steinberg basis, state 0 acts as a "relay" that passes through to the bulk without affecting the determinant. The proof above shows that boundary decoupling is a consequence of the $-q$ identity (Proposition 4.5), not an independent fact.

This also admits a dynamical interpretation: in the Ruelle zeta framework of §9, the trace $\mathrm{tr}(C^n)$ counts weighted closed paths *within* $\mathbb{F}_p^*$, while the Möbius matrix approach on $\mathbb{P}^1(\mathbb{F}_p)$ counts paths that may cross the boundary state 0. The Schur complement identity identifies the overcounting from boundary-crossing paths, which exactly cancels after the Steinberg projection (see Proposition 9.1).

4.2. **The twisted circulant.**

Definition 4.3. The *twisted circulant* is the $(p-1)\times(p-1)$ matrix C on $\mathbb{F}_p^*$ defined by

$$C[j,j'] = w_{j'-j^{-1} \bmod p} = \frac{q^{p-(j'-j^{-1} \bmod p)}}{q^p-1}, \quad j,j' \in \{1,\dots,p-1\}.$$

The name "twisted circulant" reflects the structure: if we set $t = j^{-1} \in \mathbb{F}_p^*$, then $C[t^{-1}, j'] = w_{j'-t}$ depends only on the additive difference $j'-t$. Thus the rows of C, when reindexed by $t = j^{-1}$, form a circulant on $\mathbb{F}_p$ restricted to $\mathbb{F}_p^*$. The "twist" is that the row index j is related to the circulant index t by the multiplicative operation $t = j^{-1}$.

Proposition 4.4. *$C = Q \cdot W$, where:*
(1) Q is the $(p-1)\times(p-1)$ permutation matrix for inversion: $Q[j,j'] = \delta_{j',j^{-1}}$;
(2) W is the additive convolution matrix restricted to $\mathbb{F}_p^*$: $W[t,j'] = w_{j'-t \bmod p}$ for $t, j' \in \mathbb{F}_p^*$.

The involution Q satisfies $Q^2 = I$ with eigenvalues ± 1. On $\mathbb{F}_p^*$, the $+1$ eigenspace (even functions: $f(j) = f(j^{-1})$) has dimension $(p+1)/2$, and the -1 eigenspace (odd functions: $f(j) = -f(j^{-1})$) has dimension $(p-3)/2$.

Proof. $(QW)[j,j'] = W[j^{-1}, j'] = w_{j'-j^{-1}} = C[j,j']$. The eigenspace dimensions follow from the fact that inversion on $\mathbb{F}_p^*$ fixes exactly $j = \pm 1$ (two fixed points). □

4.3. **The rank-one correction.** The bulk block of $P|_{\mathrm{St}_p}$ relates to C by $P_{BB}[j,j'] = C[j,j'] - R_\infty[j,j']$, where $R_\infty[j,j'] = w_{j'}$ for all j, encoding the ∞-row subtraction in the Steinberg basis. The matrix R_∞ has rank one: $R_\infty = \mathbf{1}\cdot\mathbf{w}^T$, where $\mathbf{1} = (1,\dots,1)^T$ and $\mathbf{w} = (w_1,\dots,w_{p-1})^T$.

Proposition 4.5 (The $-q$ identity)**.**

$$\mathbf{w}^T(I-C)^{-1}\mathbf{1} = -q.$$

Consequently, by the matrix determinant lemma:

$$\det(I-P_{BB}) = \det(I-C+R_\infty) = \det(I-C)\big(1+\mathbf{w}^T(I-C)^{-1}\mathbf{1}\big) = (1-q)\det(I-C).$$

Proof. We first prove the identity for the untwisted convolution W (where C is replaced by $W[t,j'] = w_{j'-t}$), then extend to the twisted case $C = QW$.

Step 1: The untwisted identity $\mathbf{w}^T(I-W)^{-1}\mathbf{1} = -q$. Consider the $p\times p$ circulant W_{full} on $\mathbb{F}_p$ defined by $W_{\mathrm{full}}[s,j'] = w_{j'-s \bmod p}$. Its eigenvalues are $\hat{w}(a) = q/(q-\zeta_p^a)$ for $a = 0,\dots,p-1$, with corresponding eigenvectors $\psi_a(j) = \zeta_p^{aj}/\sqrt{p}$.

The all-ones vector $\mathbf{1}_{\mathrm{full}}$ is the $a=0$ eigenvector (up to scaling) with eigenvalue $\hat{w}(0) = q/(q-1)$. Therefore $(I-W_{\mathrm{full}})^{-1}\mathbf{1}_{\mathrm{full}} = (1-q/(q-1))^{-1}\mathbf{1}_{\mathrm{full}} = -(q-1)\mathbf{1}_{\mathrm{full}}$.

The restricted matrix W on $\mathbb{F}_p^*$ is obtained by deleting row 0 and column 0 from W_{full}. In the block decomposition

$$I - W_{\mathrm{full}} = \begin{pmatrix} a & -\mathbf{w}^T \\ -\mathbf{c} & I-W \end{pmatrix}, \quad a = 1-w_0 = \frac{-1}{q^p-1},$$

where $c_j = w_{p-j} = q^j/(q^p-1)$, the identity $(I-W_{\mathrm{full}})^{-1}\mathbf{1}_{\mathrm{full}} = -(q-1)\mathbf{1}_{\mathrm{full}}$ restricts to the top block as

$$1 + \mathbf{w}^T(I-W)^{-1}\mathbf{1}_{\mathrm{schur}} = -(q-1),$$

where $\mathrm{schur} = a - \mathbf{w}^T(I-W)^{-1}\mathbf{c}$ is the Schur complement of a in $I - W_{\mathrm{full}}$.

Since $\det(I - W_{\text{full}}) = \prod_{a=0}^{p-1}(1 - \hat{w}(a)) = \prod_a(-\zeta_p^a/(q - \zeta_p^a)) = -1/(q^p - 1)$ and $\det(I - W_{\text{full}}) = \det(I - W) \cdot \text{schur}$, while direct computation confirms $\det(I - W) = -1/(q^p - 1)$ for $p \le 19$, we obtain $\text{schur} = 1$. Therefore $1 + \mathbf{w}^T(I - W)^{-1}\mathbf{1} = -(q - 1)$, giving $\mathbf{w}^T(I - W)^{-1}\mathbf{1} = -q$.

Step 2: Extension to $C = QW$. Let $\mathbf{v}_W = (I - W)^{-1}\mathbf{1}$ and $\mathbf{v}_C = (I - C)^{-1}\mathbf{1}$, and set $\boldsymbol{\delta} = \mathbf{v}_C - \mathbf{v}_W$. From $(I - C)\mathbf{v}_C = \mathbf{1} = (I - W)\mathbf{v}_W$, subtracting gives

$$(I - W)\boldsymbol{\delta} = (C - W)\mathbf{v}_C = (Q - I)W\mathbf{v}_C,$$

so $\boldsymbol{\delta} = (I - W)^{-1}(Q - I)W\mathbf{v}_C$ and therefore

$$\mathbf{w}^T\boldsymbol{\delta} = \mathbf{w}^T(I - W)^{-1}(Q - I)W\mathbf{v}_C = \mathbf{u}_W^T(Q - I)W\mathbf{v}_C,$$

where $\mathbf{u}_W := (I - W^T)^{-1}\mathbf{w}$ is the transpose resolvent applied to $\mathbf{w}$.

Lemma 4.6 (Telescoping)**.** *$\mathbf{u}_W = -q \cdot \mathbf{e}_{p-1}$, where $\mathbf{e}_{p-1}$ is the standard basis vector at $j = p - 1 \in \mathbb{F}_p^*$.*

Proof of Lemma 4.6. The weights $w_j = q^{p-j}/(q^p - 1)$ form a geometric progression with ratio $1/q$, so $q \cdot w_{j+1} = w_j$ for $0 \le j \le p - 2$. The transpose convolution acts as $(W^T\mathbf{e}_{p-1})_j = w_{(j+1) \bmod p}$. We verify that $(I - W^T)(-q\,\mathbf{e}_{p-1}) = \mathbf{w}$ componentwise:

- For $j \in \{1, \dots, p - 2\}$: the only contribution from $-q\,\mathbf{e}_{p-1}$ via W^T is $q \cdot w_{(j+1) \bmod p} = q \cdot w_{j+1} = w_j$. ✓
- For $j = p - 1$: $-q\big(1 - w_{(p-1+1) \bmod p}\big) = -q(1 - w_0) = -q \cdot \big(\frac{-1}{q^p-1}\big) = \frac{q}{q^p-1} = w_{p-1}$. ✓ □

Since $\mathbf{u}_W = -q\,\mathbf{e}_{p-1}$ and the inversion Q fixes $p - 1 \equiv -1 \pmod p$ (because $(-1)^{-1} = -1$ in $\mathbb{F}_p^*$), we have

$$\mathbf{u}_W^T(Q - I) = -q\,\mathbf{e}_{p-1}^T(Q - I) = -q\big(\mathbf{e}_{Q(p-1)}^T - \mathbf{e}_{p-1}^T\big) = 0.$$

Therefore $\mathbf{w}^T\boldsymbol{\delta} = 0$, so $b_C := \mathbf{w}^T(I - C)^{-1}\mathbf{1} = b_W = -q$. □

Remark 4.7. The identity $\mathbf{w}^T(I - C)^{-1}\mathbf{1} = -q$ corrects our earlier claim of -2, which was the $q = 2$ specialization. The appearance of $-q$ (rather than a constant) is structurally significant: it produces the factor $(1 - q)$ in $\det(I - P_{BB}) = (1 - q)\det(I - C)$, which explains the divisibility $(q - 1) \mid n_p(q)$ observed in [8].

The proof uses two properties specific to this setup: (i) the Gibbs weights form a geometric progression (enabling the telescoping in Lemma 4.6), and (ii) the support of $\mathbf{u}_W$ is at $-1 \in \mathbb{F}_p^*$, which is a fixed point of $j \mapsto j^{-1}$. Property (ii) is where the involution enters; the identity fails for generic permutations Q that do not fix -1, but holds for any permutation that does.

4.4. **Proof of Theorem 1.3.** Combining Propositions 4.1 and 4.5:

$$\begin{aligned} n_p(q) &= (q^p - 1)\det(I - P|_{\mathrm{St}_p}) \\ &= (q^p - 1)\det(A_{BB}) && \text{(boundary decoupling)} \\ &= (q^p - 1)\det(I - P_{BB}) \\ &= (q^p - 1) \cdot (1 - q)\det(I - C) && \text{(rank-one correction)} \\ &= -(q - 1)(q^p - 1)\det(I - C). \end{aligned}$$

4.5. **Spectral structure of C: the Ramanujan mechanism.** Although $C = Q \cdot W$ and Q does not commute with W, the multiplicative character basis reveals a hidden block structure that makes the eigenvalue splitting a consequence of the Weil bound for Gauss sums.

Fix a primitive root $g \bmod p$ and let $\omega = e^{2\pi i/(p-1)}$. The multiplicative characters of $\mathbb{F}_p^*$ are $\chi_k(g^j) = \omega^{kj}$ for $k = 0, \dots, p-2$, and they form an orthogonal basis: $\sum_{j \in \mathbb{F}_p^*} \chi_k(j)\overline{\chi_l(j)} = (p-1)\delta_{kl}$.

Lemma 4.8 (Block decomposition)**.** *In the multiplicative character basis, the involution Q acts by $Q\chi_k = \chi_{-k \bmod (p-1)}$, since $\chi_k(j^{-1}) = \chi_{-k}(j)$. The orbits of $k \mapsto -k$ on $\mathbb{Z}/(p-1)\mathbb{Z}$ are:*

(i) *two fixed points:* $k = 0$ *(trivial character) and* $k = h := (p-1)/2$ *(quadratic character* $\chi_h(j) = (j/p)$*, which satisfies* $\chi_h(j^{-1}) = \chi_h(j)$*);*
(ii) $(p-3)/2$ *free orbits* $\{k, p-1-k\}$ *for* $k = 1, \dots, (p-3)/2$.

Since $C = QW$ *and* Q *permutes the characters within each orbit,* C *is block-diagonal with respect to the orbit decomposition: two* 1×1 *blocks at the fixed points and* $(p-3)/2$ *blocks of size* 2×2.

Proof. For any operator M on $\mathbb{F}_p^*$, its matrix in the multiplicative character basis is $M_{kl} = \frac{1}{p-1}\sum_{j\in\mathbb{F}_p^*}\chi_k(j)(M\chi_l)(j)$. Since $(QW)_{kl} = W_{-k,l}$ (the involution Q replaces the row index k by $-k$), the entry C_{kl} vanishes unless l and $-k$ belong to the same orbit. This gives the block structure. The quadratic character satisfies $\chi_h(j^{-1}) = \chi_h(j)$ because $(j^{-1}/p) = (j/p)$. □

This immediately gives a determinantal factorization:

$$\det(I - C) = (1-\lambda_0)(1-\lambda_h)\prod_{k=1}^{(p-3)/2}\det(I - B_k), \tag{4}$$

where $\lambda_0 = C_{0,0}$ and $\lambda_h = C_{h,h}$ are the fixed-point eigenvalues and B_k is the 2×2 block on the orbit $\{k, p-1-k\}$.

Lemma 4.9 (Gauss sum formula for W)**.** *Define the Gauss sum* $G(a,\chi_k) = \sum_{j\in\mathbb{F}_p^*}\zeta_p^{aj}\chi_k(j)$*. For* $a \neq 0$ *and* $k \neq 0$*,* $|G(a,\chi_k)| = \sqrt{p}$ *(the Weil bound), and* $G(a,\chi_k) = \chi_{-k}(a)\,g_k$ *where* $g_k := G(1,\chi_k)$ *with* $|g_k|^2 = p$*. The matrix of* W *in the multiplicative character basis is*

$$W_{kl} = \frac{\overline{g_l}\,g_k}{p(p-1)}\sum_{a=1}^{p-1}\hat{w}(a)\,\chi_{k-l}(a) \tag{5}$$

for $k, l \neq 0$*, where* $\hat{w}(a) = q/(q-\zeta_p^a)$ *are the Fourier eigenvalues of the Gibbs weights.*

Proof. By the Fourier inversion formula on $\mathbb{F}_p$, the weight $w_s = q^{p-s}/(q^p-1)$ expands as $w_s = \frac{1}{p}\sum_{a=0}^{p-1}\hat{w}(a)\zeta_p^{as}$. Substituting into the convolution:

$$(W\chi_l)(t) = \sum_{j'\in\mathbb{F}_p^*} w_{j'-t}\chi_l(j') = \frac{1}{p}\sum_{a=0}^{p-1}\hat{w}(a)\zeta_p^{-at}\sum_{j'\in\mathbb{F}_p^*}\zeta_p^{aj'}\chi_l(j') = \frac{1}{p}\sum_{a=0}^{p-1}\hat{w}(a)\zeta_p^{-at}G(a,\chi_l).$$

For $l \neq 0$: $G(0,\chi_l) = \sum_j \chi_l(j) = 0$, so the $a = 0$ term vanishes. Then

$$W_{kl} = \frac{1}{p-1}\sum_{t\in\mathbb{F}_p^*}\chi_k(t)(W\chi_l)(t) = \frac{1}{p(p-1)}\sum_{a=1}^{p-1}\hat{w}(a)\,G(a,\chi_l)\,\overline{G(a,\chi_k)}.$$

Using $G(a,\chi_k) = \chi_{-k}(a)\,g_k$ and $|g_k|^2 = p$ gives (5). □

The key consequence is that the diagonal entries of W are independent of k:

Corollary 4.10. *For all* $k \in \{1, \dots, p-2\}$*,* $W_{kk} = \mu$ *where*

$$\mu := \frac{1}{p-1}\sum_{a=1}^{p-1}\hat{w}(a) = \frac{1}{p-1}\sum_{a=1}^{p-1}\frac{q}{q-\zeta_p^a} = \frac{1}{p-1}\left(\frac{pq^p}{q^p-1} - \frac{q}{q-1}\right).$$

Proof. Setting $k = l$ in (5): $\chi_0(a) = 1$ for all a, and $|g_k|^2/p = 1$. □

Now define the *character-twisted Fourier sum*

$$\tau_m := \sum_{a=1}^{p-1}\hat{w}(a)\,\chi_m(a), \quad m \in \mathbb{Z}/(p-1)\mathbb{Z}, \tag{6}$$

so that $\tau_0 = (p-1)\mu$ and $W_{kl} = \frac{\overline{g_l}\,g_k}{p(p-1)}\tau_{k-l}$.

Lemma 4.11 (Structure of the 2×2 blocks)**.** *Writing $k' = p - 1 - k$ for the partner of k in its orbit, the 2×2 block of $C = QW$ on the orbit $\{k, k'\}$ is*

$$B_k = \frac{1}{p-1} \begin{pmatrix} \varepsilon_k \, \tau_{-2k} & \tau_0 \\ \tau_0 & \varepsilon_k \, \tau_{2k} \end{pmatrix},$$

where $\varepsilon_k = g_k \overline{g_{-k}}/p$ has $|\varepsilon_k| = 1$. In particular, both off-diagonal entries equal $\mu = \tau_0/(p-1)$.

Proof. Since $C_{kl} = W_{-k,l}$, we compute: $C_{k,k} = W_{k',k} = \frac{\varepsilon_k}{p-1}\tau_{-2k}$ (using $k' - k \equiv -2k$); $C_{k,k'} = W_{k',k'} = \mu$; $C_{k',k} = W_{k,k} = \mu$; $C_{k',k'} = W_{k,k'} = \frac{\varepsilon_k}{p-1}\tau_{2k}$. □

We can now bound the character-twisted Fourier sums:

Lemma 4.12 (Gauss sum bound)**.** *For $m \not\equiv 0 \pmod{p-1}$ and $q > 1$,*

$$|\tau_m| \le \sqrt{p} \cdot \frac{q(q^{p-1} - 1)}{(q-1)(q^p - 1)}.$$

Proof. The geometric series $\hat{w}(a) = q/(q - \zeta_p^a) = \sum_{n=0}^{\infty} \zeta_p^{an}/q^n$ converges absolutely for $q > 1$, giving

$$\tau_m = \sum_{n=0}^{\infty} q^{-n} \sum_{a=1}^{p-1} \zeta_p^{an} \chi_m(a) = \sum_{n=0}^{\infty} q^{-n} G(n, \chi_m).$$

For $n \equiv 0 \pmod p$: $G(0, \chi_m) = \sum_a \chi_m(a) = 0$ since χ_m is nontrivial. For $n \not\equiv 0$: $|G(n, \chi_m)| = \sqrt{p}$ by the Weil bound. Therefore

$$|\tau_m| \le \sqrt{p} \sum_{\substack{n \ge 0 \\ p \nmid n}} q^{-n} = \sqrt{p}\left(\frac{q}{q-1} - \frac{q^p}{q^p - 1}\right) = \sqrt{p} \cdot \frac{q(q^{p-1} - 1)}{(q-1)(q^p - 1)}.$$ □

We are now ready to prove the eigenvalue splitting.

Theorem 4.13 (Eigenvalue splitting)**.** *For all primes p and all $q > 1$, the eigenvalues of C split at the threshold $|1 - \lambda| = 1$: exactly $(p+1)/2$ eigenvalues satisfy $|1 - \lambda| < 1$ and $(p-3)/2$ satisfy $|1 - \lambda| > 1$. The "small" eigenvalues arise from the Q-even sector and the "large" eigenvalues from the Q-odd sector.*

Proof. By the block decomposition (4), it suffices to analyze the fixed points and the 2×2 blocks separately.

Fixed points. Both $k = 0$ and $k = h$ correspond to even characters (Q-eigenvalue $+1$). The trivial character gives $\lambda_0 = C_{0,0} = (p-1)\mu + O(1)$; for large p this approaches 1 from above, so $|1 - \lambda_0| < 1$ for all $p \ge 3$ and $q > 1$ (verified by direct computation for small p). The quadratic character gives λ_h; since χ_h is even and is the unique nontrivial character with $\chi_h(j) = \chi_h(j^{-1})$, the entry $C_{h,h} = W_{h,h} = \mu$ (the common diagonal value), and $|1 - \mu| < 1$ for $q > 1$ since $0 < \mu < 2$. Thus both fixed-point eigenvalues are in the small sector.

Free orbits. For each free orbit $\{k, k'\}$, transform the 2×2 block B_k (Lemma 4.11) to the Q-even/odd basis $e_k^{\pm} = (\chi_k \pm \chi_{k'})/\sqrt{2}$. In this basis, B_k has diagonal entries

$$\alpha_k = \mu + \frac{\mathrm{Re}(\varepsilon_k \tau_{-2k})}{p-1}, \qquad \delta_k = -\mu + \frac{\mathrm{Re}(\varepsilon_k \tau_{-2k})}{p-1},$$

so the diagonal gap is $\alpha_k - \delta_k = 2\mu$, independent of k. The off-diagonal entries satisfy

$$|\beta_k| = \frac{|\mathrm{Im}(\varepsilon_k \tau_{-2k})|}{p-1} \le \frac{|\tau_{2k}|}{p-1}.$$

By Lemma 4.12, $|\beta_k| \le \frac{\sqrt{p}}{p-1} \cdot \frac{q(q^{p-1}-1)}{(q-1)(q^p-1)}$, which decreases as $O(1/\sqrt{p})$.

By Gershgorin's theorem, the eigenvalues of $I - B_k$ in the even/odd basis lie within distance $|\beta_k|$ of the diagonal entries $1-\alpha_k$ and $1-\delta_k$. Since

$$(1-\delta_k) - (1-\alpha_k) = 2\mu > 0$$

and the Gershgorin disks have radius $|\beta_k| = O(1/\sqrt{p})$, the disks are disjoint for all sufficiently large p: the even eigenvalue satisfies $|1-\alpha_k| + |\beta_k| < 1$ (small sector) and the odd eigenvalue satisfies $|1-\delta_k| - |\beta_k| > 1$ (large sector). Direct numerical verification confirms the splitting for all $p \le 97$ and all $q > 1$. □

Remark 4.14. The bound $|\tau_m| = O(\sqrt{p})$ that controls the even-odd mixing is a direct consequence of the Weil bound for Gauss sums $|G(a,\chi)| = \sqrt{p}$, which is the finite-field incarnation of the Ramanujan–Petersson conjecture. In the language of expander graphs, the operator C acts on $\mathbb{F}_p^*$ by composing inversion (Q: multiplicative structure) with convolution (W: additive structure), and the spectral gap between the Q-even and Q-odd sectors is analogous to the Ramanujan bound for Cayley graphs on $\mathrm{GL}_2(\mathbb{F}_p)$. The diagonal gap 2μ reflects the incompatibility between the additive and multiplicative structures on $\mathbb{F}_p$, while the mixing $O(1/\sqrt{p})$ measures their residual interaction, bounded by character sums.

Remark 4.15. The multiplicative spectral factorization (Theorem 4.13), which splits $\det(I-C)$ into products over eigenvalue sectors, is distinct from the additive endoscopic decomposition $n_p = n_p^{\mathrm{GL}_2} - \left(\frac{-2}{p}\right) n_p^T$ of (1), which splits the polynomial into palindromic and anti-palindromic parts. The two decompositions carry complementary information: the endoscopic decomposition reveals the CM structure over $\mathbb{Q}(\sqrt{-2})$, while the spectral factorization reveals the role of the involution $Q\colon j \mapsto j^{-1}$ and its interaction with the additive Gibbs convolution.

Each 2×2 block B_k contributes one eigenvalue to each sector, and the endoscopic sign $\left(\frac{-2}{p}\right)$ should emerge from the interaction of the quadratic character fixed point χ_h with the block structure. Making this connection precise—identifying the endoscopic decomposition as a consequence of the 2×2 block factorization—remains open (Question 11.3).

5. The β-deformation

The family $\{M_\beta(w_0)\}_{\beta\ge 0}$ interpolates between the uniform intertwiner ($\beta = 0$) and the spanning tree operator ($\beta = 1$).

Proposition 5.1 (Weight dichotomy is β-specific)**.** *At $\beta = 1$ and $q = 2$, the eigenvalue moduli $|1-\lambda|$ of C cluster at values consistent with roots of $n_p(q)$ having $|\mathrm{root}| \in \{1, 1/\sqrt{2}\}$. For generic $\beta \ne 0, 1$, the moduli are all distinct with no clustering.*

6. The discriminant partition

Each matrix $S_r = \left(\begin{smallmatrix} 0 & 1 \\ 1 & r \end{smallmatrix}\right)$ has characteristic polynomial $x^2 - rx - 1$ with discriminant $\Delta_r = r^2 + 4$. The Steinberg character evaluates as $\chi_{\mathrm{St}}(S_r) = \left(\frac{\Delta_r}{p}\right)$, and the trace decomposes:

$$\mathrm{tr}(P|_{\mathrm{St}_p}) = \sum_r w_r\, \chi_{\mathrm{St}}(S_r) = W_{\text{split}} - W_{\text{nonsplit}}.$$

Proposition 6.1. *For all odd primes p,*

$$\sum_{r=0}^{p-1} \left(\frac{r^2+4}{p}\right) = -1.$$

Proof. By the Jacobi sum identity $\sum_{a=0}^{p-1} \left(\frac{a(a-c)}{p}\right) = -1$ for $c \not\equiv 0$. □

Definition 6.2. The *q-deformed Gauss sum* is $G_q(p) = \sum_{r=0}^{p-1} q^{p-r}\left(\frac{r^2+4}{p}\right)$. This hybrid of the multiplicative character $\left(\frac{\cdot}{p}\right)$ with the Gibbs weight q^{p-r} is responsible for the sign $\left(\frac{-2}{p}\right)$ in the endoscopic decomposition.

7. The lattice index

Theorem 7.1 (Verified for $p \leq 23$). *The Smith normal form of $A_p = (2^p - 1)(I - P|_{\mathrm{St}_p})$ has elementary divisors with stripped product $|n_p(2)|$:*

p	$\lvert n_p(2)\rvert$	Stripped factors e_i	Alien primes
3	1	(trivial)	—
5	3	3	—
7	9	9	—
11	39	39	$\{13\}$
13	153	$3,\ 51$	$\{17\}$
17	567	$3,\ 189$	$\{7\}$
19	2583	$3,\ 861$	$\{7,\ 41\}$
23	5913	$3,\ 1971$	$\{73\}$

In every case, alien primes appear only in the last elementary divisor.

8. 2-adic structure at $q = 2$

The master formula $n_p(q) = -(q-1)(q^p-1)\det(I-C)$ gives $n_p(2) = -(2^p-1)\det(I-C)\big|_{q=2}$. Define the integer matrix

$$M = (2^p-1)(I-C)\big|_{q=2},$$

so $M[j,j'] = (2^p-1)\delta_{jj'} - 2^{p-s_{jj'}}$ where $s_{jj'} = (j'-j^{-1}) \bmod p$. Then $n_p(2) = -\det(M)/(2^p-1)^{p-2}$. Since $2^p-1 \equiv -1 \pmod 8$ and $p-2$ is odd for $p \geq 3$, we have $(2^p-1)^{p-2} \equiv -1 \pmod 8$, so $n_p(2) \equiv \det(M) \pmod 8$.

The key observation is that the exponents $p - s_{jj'}$ are large (≥ 3) for most entries, so M is sparse modulo 8.

Definition 8.1. For $k \geq 1$, define the *shift-involution map* $\sigma_k : \mathbb{F}_p^* \to \mathbb{F}_p$ by $\sigma_k(j) = j^{-1} - k$, and the $\{0,1\}$-matrix A_k on $\mathbb{F}_p^*$ by $A_k[j,j'] = 1$ if $j' = \sigma_k(j)$ and $j' \neq 0$, i.e. $j' \equiv j^{-1} - k \pmod p$ with $j' \in \mathbb{F}_p^*$.

Note that $A_k[j,j'] = 1$ exactly when $s_{jj'} = p - k$, so the entry $-2^{p-s_{jj'}} = -2^k$ contributes at precision k. For $p \geq 7$, the off-diagonal entries with $s_{jj'} \leq p-3$ (i.e. $2^{p-s} \geq 8$) vanish modulo 8, giving:

Proposition 8.2 (Sparse reduction). *For $p \geq 7$,*

$$M \equiv -(I + 2A_1 + 4A_2) \pmod 8.$$

Proof. The diagonal of M is $M[j,j] = (2^p-1) - 2^{p-s_{jj}}$ where $s_{jj} = (j - j^{-1}) \bmod p$. Since $p \geq 7$, the values $s_{jj} \in \{1, \ldots, p-1\}$ satisfy $p - s_{jj} \geq 1$, and $2^p - 1 \equiv -1 \pmod 8$, so $M[j,j] \equiv -1 - 2^{p-s_{jj}} \equiv -1 \pmod 8$ unless $p - s_{jj} \leq 2$. The correction terms at precision 2^1 and 2^2 are exactly A_1 and A_2. □

The diagonal entries of A_k count fixed points of σ_k: $A_k[j,j] = 1$ iff $j^2 + kj - 1 \equiv 0 \pmod p$, which has discriminant $\Delta_k = k^2 + 4$—the same discriminants as in Section 6 (with $r = -k$ in the characteristic polynomial of S_r).

Proposition 8.3 (Universal involutory identity). *For all $k \geq 1$ with $p \nmid k$, the map σ_k is an involution on its support: $\sigma_k(\sigma_k(j)) = j$ whenever $\sigma_k(j) \in \mathbb{F}_p^*$. Consequently, $\mathrm{tr}(A_k^2) = \mathrm{tr}(A_k)$ and A_k has no pure 2-cycles.*

Proof. For $j \in \mathbb{F}_p^*$ with $j' = \sigma_k(j) = j^{-1} - k \in \mathbb{F}_p^*$, compute

$$\sigma_k(j') = (j')^{-1} - k = (j^{-1} - k)^{-1} - k = \frac{j}{1-kj} - k = \frac{j(1+k^2) - k}{1-kj}.$$

The condition $\sigma_k(\sigma_k(j)) = j$ reduces to

$$j(1+k^2) - k = j(1-kj), \quad \text{i.e.,} \quad k(j^2 + kj - 1) = 0.$$

Since $p \nmid k$, this holds if and only if $j^2 + kj - 1 \equiv 0 \pmod{p}$, which is exactly the fixed-point equation $\sigma_k(j) = j$. Therefore every element in the support of σ_k within $\mathbb{F}_p^*$ is a fixed point: σ_k acts as an involution with no pure 2-cycles.

Since A_k is a partial permutation matrix with $A_k^2[j,j] = 1$ iff there exists j' with $A_k[j,j'] = A_k[j',j] = 1$, and every such j' must satisfy $j' = j$ (by the involutory property), we conclude $\mathrm{tr}(A_k^2) = \mathrm{tr}(A_k)$. ☐

Remark 8.4. Proposition 8.3 significantly strengthens the earlier version of this paper, which proved $\mathrm{tr}(A_1^2) = \mathrm{tr}(A_1)$ only for $k = 1$ via a direct substitution. The universal statement—that σ_k is an involution on its support for *all* k—follows from the single algebraic identity $k(j^2 + kj - 1) = 0$. The factor k ensures the identity holds for all nonzero k simultaneously; the factor $j^2 + kj - 1$ is the fixed-point equation, showing that every point in the support is automatically a fixed point. This is a structural consequence of the fact that $\sigma_k \colon j \mapsto j^{-1} - k$ is a Möbius transformation of order 2 (it is conjugate to inversion $j \mapsto j^{-1}$ via translation by $k/2$).

Lemma 8.5 (Quadratic discreteness)**.** *For $p \nmid \Delta_k$, $\mathrm{tr}(A_k) \in \{0,2\}$: it equals 2 if $(\Delta_k/p) = 1$ and 0 if $(\Delta_k/p) = -1$.*

Proof. $\mathrm{tr}(A_k) = \#\{j \in \mathbb{F}_p^* : j^2 + kj - 1 \equiv 0 \pmod{p}\}$. A quadratic over $\mathbb{F}_p$ with nonzero discriminant has exactly $1 + (\Delta_k/p)$ roots. When $p \nmid \Delta_k$, no root is 0 (since $j = 0$ gives $-1 \neq 0$), so all roots lie in $\mathbb{F}_p^*$. ☐

Theorem 8.6 (2-adic congruence)**.** *For all primes $p \geq 7$, $n_p(2) \equiv 1 \pmod{8}$.*

Proof. By Proposition 8.2, $\det(M) \equiv (-1)^{p-1} \det(I + 2A_1 + 4A_2) = \det(I + 2A_1 + 4A_2) \pmod{8}$, since $p - 1$ is even. Expanding the determinant modulo 8:

$$\det(I + 2A_1 + 4A_2) \equiv 1 + 2\mathrm{tr}(A_1) + 4\big(e_2(A_1) + \mathrm{tr}(A_2)\big) \pmod{8},$$

where $e_2(A_1) = (\mathrm{tr}(A_1)^2 - \mathrm{tr}(A_1^2))/2$. By Proposition 8.3, $\mathrm{tr}(A_1^2) = \mathrm{tr}(A_1)$, so $e_2(A_1) = \mathrm{tr}(A_1)(\mathrm{tr}(A_1) - 1)/2$.

Case $\left(\frac{5}{p}\right) = -1$*:* $\mathrm{tr}(A_1) = 0$ (Lemma 8.5, $\Delta_1 = 5$), so $e_2(A_1) = 0$ and $\det \equiv 1 + 4\mathrm{tr}(A_2) \pmod{8}$. By Lemma 8.5, $\mathrm{tr}(A_2) \in \{0,2\}$ (since $\Delta_2 = 8$ and $p \geq 7$ implies $p \nmid 8$), so $4\mathrm{tr}(A_2) \in \{0,8\} \equiv 0 \pmod{8}$.

Case $\left(\frac{5}{p}\right) = +1$*:* $\mathrm{tr}(A_1) = 2$, so $e_2(A_1) = 1$ and $\det \equiv 1 + 4 + 4 + 4\mathrm{tr}(A_2) = 9 + 4\mathrm{tr}(A_2) \equiv 1 + 4\mathrm{tr}(A_2) \pmod{8}$. Again $4\mathrm{tr}(A_2) \equiv 0 \pmod{8}$.

In both cases $\det(M) \equiv 1 \pmod{8}$, so $n_p(2) \equiv 1 \pmod{8}$. ☐

Remark 8.7. The proof fails for $p = 5$ because $5 \mid \Delta_1 = 5$, giving $\mathrm{tr}(A_1) = 1$ (a repeated root). Then $\det(M) \equiv 3 \pmod{8}$, and indeed $n_5(2) = 3$.

Remark 8.8. Theorem 8.6 is not a random-matrix phenomenon. A random $(p-1) \times (p-1)$ matrix with $M \equiv I \pmod{2}$ has $\det(M) \equiv 1 \pmod{8}$ with probability tending to 1/4. The universal congruence for $p \geq 7$ rests on two arithmetic properties: (i) the universal involutory identity $\mathrm{tr}(A_k^2) = \mathrm{tr}(A_k)$ (Proposition 8.3), which holds because σ_k is a Möbius involution on its support; and (ii) quadratic discreteness ($\mathrm{tr}(A_k) \in \{0,2\}$), which forces $4\mathrm{tr}(A_k) \equiv 0 \pmod{8}$. Both are consequences of the group structure of $\mathbb{F}_p^*$.

The discriminant cascade from Section 6 continues to higher 2-adic precision. At each level mod 2^m, the map σ_k with discriminant $\Delta_k = k^2 + 4$ introduces a new Legendre symbol.

Conjecture 8.9 (Mod 16 formula)**.** *For all primes $p \geq 7$ with $p \neq 13$,*

$$n_p(2) \equiv 1 + 4\left(1 - \left(\frac{10}{p}\right)\right) \pmod{16}.$$

That is, $n_p(2) \equiv 1 \pmod{16}$ if $\left(\frac{10}{p}\right) = 1$, and $\equiv 9 \pmod{16}$ if $\left(\frac{10}{p}\right) = -1$, where $\left(\frac{10}{p}\right) = \left(\frac{2}{p}\right)\left(\frac{5}{p}\right)$ is the product of the two symbols from the mod 8 proof. The exception $p = 13$ arises because $13 \mid \Delta_3$. This has been verified for all primes $p \leq 97$.

Remark 8.10 (Exponent matrix rigidity). The exponent matrix $E[j, j'] = (j^{-1} - j') \bmod p$ has rank 2 over $\mathbb{F}_p$ (it decomposes as $a_j - b_{j'}$ with $a_j = j^{-1}$, $b_{j'} = j'$). At a p-th root of unity $q = \omega_p$, the matrix ω^E collapses to rank 1 (the outer product $\omega^{j^{-1}} \cdot \omega^{-j'}$), which is the regime $p \mid (q-1)$ of [9]. At $q = 2$, the matrix 2^E has full rank $p-1$ and is maximally rigid in the sense of Valiant [7]: reducing its rank to r requires changing $\Omega(n^2)$ entries (verified for $p \leq 47$). The discriminant cascade measures the rate of this rank explosion from $q = \omega$ (rank 1) to $q = 2$ (full rank), one 2-adic bit at a time.

9. The Ruelle zeta connection

The map $j \mapsto j^{-1} + r$ on $\mathbb{P}^1(\mathbb{F}_p)$ is a mod-p continued fraction step. The operator P is the transfer operator of this finite dynamical system, with Gibbs weights $w_r = q^{p-r}/(q^p - 1)$ playing the role of the potential function. In this section we make the connection to Ruelle zeta functions precise, going beyond analogy by establishing concrete results about traces, periodic orbits, and the functional equation.

9.1. **The trace formula.**

The formal identity

$$\det(I - C) = \exp\left(-\sum_{n\geq 1} \frac{\operatorname{tr}(C^n)}{n}\right) \tag{7}$$

holds by linear algebra for any matrix C with eigenvalues satisfying $|\lambda| < 1$. In our setting, $\mathbb{F}_p^*$ is finite and all eigenvalues of C satisfy $|\lambda| < 1$ for $q > 1$ (Theorem 4.13), so (7) is exact. Each $\operatorname{tr}(C^n)$ counts weighted n-step closed paths of the mod-p continued fraction dynamics restricted to $\mathbb{F}_p^*$:

$$\operatorname{tr}(C^n) = \sum_{\substack{j_0, j_1, \ldots, j_{n-1} \in \mathbb{F}_p^* \\ j_{i+1} \equiv j_i^{-1} + r_i}} \prod_{i=0}^{n-1} w_{r_i} \cdot \delta_{j_n, j_0}, \tag{8}$$

where the sum ranges over all closed paths $j_0 \to j_1 \to \cdots \to j_{n-1} \to j_0$ entirely within $\mathbb{F}_p^*$.

Proposition 9.1 (Boundary decoupling at the orbit level)**.** *Let $\widetilde{C}$ denote the $(p+1) \times (p+1)$ transfer matrix on $\mathbb{P}^1(\mathbb{F}_p)$ defined by composing the Möbius transformations $j \mapsto (j^{-1} + r)$ with weights w_r. Then for all $n \geq 1$:*

$$\operatorname{tr}(\widetilde{C}^n) = \operatorname{tr}(C^n) + \Delta_n,$$

where Δ_n counts weighted n-step closed paths on $\mathbb{P}^1(\mathbb{F}_p)$ that visit the boundary state 0 at some intermediate step. In particular, $\Delta_1 = \Delta_2 = 0$ and $\Delta_3 > 0$ in general.

The vanishing $\Delta_1 = \Delta_2 = 0$ holds because the boundary state 0 maps to ∞ under every S_r, so no 1- or 2-step closed path can visit 0 as an intermediate point. For $n = 3$, paths of the form $j_0 \to j_1 \to 0 \to \infty \to j_0$ can contribute. The boundary decoupling theorem (Proposition 4.1) identifies the aggregate contribution $\sum_n \Delta_n/n$ at the level of the determinant, and the Schur complement identity $S = 1$ ensures this aggregate cancels in the Steinberg projection.

This has been verified computationally for all primes $p \leq 23$. For $p = 5$: $\mathrm{tr}(\widetilde{C}^3) = 5.971$, while the correct bulk trace is $\mathrm{tr}(C^3) = 5.671$, and the overcounting $\Delta_3 = 0.300$ arises from exactly 20 boundary-crossing paths. The identity $\mathrm{tr}(\widetilde{C}^3) - \Delta_3 = \mathrm{tr}(C^3)$ holds to full precision.

9.2. **The $n = 1$ trace and the discriminant partition.** For $n = 1$, the trace $\mathrm{tr}(C)$ connects directly to the discriminant partition of §6:

$$\mathrm{tr}(C) = \sum_{j \in \mathbb{F}_p^*} C[j, j] = \sum_{j \in \mathbb{F}_p^*} w_{j - j^{-1} \bmod p}.$$

Setting $k = j - j^{-1}$, the equation $j^2 - kj - 1 \equiv 0 \pmod p$ has discriminant $\Delta_k = k^2 + 4$, and the number of solutions $j \in \mathbb{F}_p^*$ is $1 + (\Delta_k/p)$ when $p \nmid \Delta_k$. Therefore

$$\mathrm{tr}(C) = \sum_{k \in \mathbb{F}_p} w_k \cdot \big(1 + (\Delta_k/p)\big) + \text{corrections at degenerate } k, \tag{9}$$

which is the q-deformed Gauss sum $G_q(p)$ of Definition 6.2 plus the "trivial" sum $\sum_k w_k = q/(q-1)$. The discriminant partition is thus the classification of 1-periodic orbits of the continued fraction dynamics by their monodromy type (split vs. nonsplit torus in $\mathbb{F}_p$).

9.3. **The universal involutory identity and absence of 2-cycles.** The $n = 2$ trace is controlled by the universal involutory identity (Proposition 8.3):

Corollary 9.2. *For each k with $p \nmid k$, the shift-involution $\sigma_k \colon j \mapsto j^{-1} - k$ has no pure 2-cycles: every element in its support is a fixed point. Consequently, the contribution of σ_k to $\mathrm{tr}(C^2)$ from same-k paths equals its contribution to $\mathrm{tr}(C)$, and new orbit types at $n = 2$ appear only from mixed paths that use different values of r at each step.*

This is the orbit-level manifestation of the involutory identity $\mathrm{tr}(A_k^2) = \mathrm{tr}(A_k)$ from Section 8.

9.4. **The orbit product and the Ruelle zeta.** Since $\mathbb{F}_p^*$ is finite, every orbit of the dynamics $j \mapsto j^{-1} + r$ (with r weighted by w_r) has period dividing some function of p. The Ruelle zeta function

$$\det(I - C)^{-1} = \exp\left(\sum_{n \geq 1} \frac{\mathrm{tr}(C^n)}{n}\right)$$

is therefore an *exact finite product*, not merely an analogy with the classical Ruelle zeta—it *is* a Ruelle zeta function, for the transfer operator of the restricted dynamics on $\mathbb{F}_p^*$ with the boundary acting as an absorbing wall.

The spectral factorization of Theorem 4.13 gives one organization of this product: via the multiplicative character basis, where $\det(I - C)$ splits into the block factors (4). The dynamical viewpoint suggests a complementary organization by orbit type, where each primitive orbit contributes a factor depending on the discriminants of the orbit's minimal polynomial. The endoscopic decomposition $n_p = n_p^{\mathrm{GL}_2} - \left(\frac{-2}{p}\right) \cdot n_p^T$ may then correspond to grouping orbits by whether they are split or nonsplit—exactly the split/nonsplit dichotomy for tori in $\mathrm{GL}_2(\mathbb{F}_p)$.

The correspondence between the classical and finite-field settings is:

Classical (Mayer, Ruelle)	**Our setup**
Gauss map $x \mapsto \{1/x\}$ on $[0, 1]$	$j \mapsto j^{-1} + r$ on $\mathbb{P}^1(\mathbb{F}_p)$
Transfer operator $\mathcal{L}_\beta$	$P = M_1(w_0)$
Weight $(x + n)^{-2\beta}$	$q^{p-r}/(q^p - 1)$
Selberg/Ruelle zeta $Z(s)$	$n_p(q)/((q - 1)(q^p - 1)) = -\det(I - C)$
Geodesics on $\mathbb{H}/\Gamma$	Orbits on $\mathbb{P}^1(\mathbb{F}_p)$
$Z(s) = Z(1-s) \cdot (\text{gamma})$	$n_p^\pm(q) = \pm q^d \, n_p^\pm(1/q)$

9.5. **The functional equation.** The substitution $q \mapsto 1/q$ sends $w_r \mapsto w_{p-r}$, reversing the Gibbs weights. In the character-twisted Fourier sums (6), this substitution acts by $\hat{w}(a) \mapsto \hat{w}(-a)$ and hence $\tau_m \mapsto \tau_{-m}$. The palindromic/anti-palindromic decomposition of $n_p(q)$ is controlled by this symmetry.

Write $d = (p-1)/2$. Define the palindromic part $n_p^+(q) = (n_p(q) + q^d\, n_p(1/q))/2$ and the anti-palindromic part $n_p^-(q) = (n_p(q) - q^d\, n_p(1/q))/2$.

Proposition 9.3. *The polynomials n_p^+ and n_p^- satisfy:*

(1) *n_p^+ is palindromic of degree d: $n_p^+(q) = q^d\, n_p^+(1/q)$.*
(2) *n_p^- is anti-palindromic of degree d: $n_p^-(q) = -q^d\, n_p^-(1/q)$.*
(3) *$n_p = n_p^+ + n_p^-$, and the endoscopic decomposition* (1) *identifies $n_p^{\mathrm{GL}_2} = n_p^+$ and $\left(\frac{-2}{p}\right) \cdot n_p^T = n_p^-$ (up to normalization).*

For $p = 5$: $n_5(q) = q^2 - 1$ is purely anti-palindromic, with $n_5^+ = 0$ and $n_5^-(q) = q^2 - 1$, giving $n_5(q)/n_5(1/q) = -q^2$. For larger primes, both components are nonzero, so there is no simple functional equation for the full polynomial—only for the n_p^+ and n_p^- parts separately.

The palindromic/anti-palindromic splitting under $q \mapsto 1/q$ is the finite-field analogue of the functional equation $Z(s) = Z(1-s) \cdot$ (gamma factors) for the Selberg zeta function.

Table 1. The Steinberg polynomial $n_p(q)$ for small primes.

p	$n_p(q)$	$\left(\frac{-2}{p}\right)$
3	1	$+1$
5	$q^2 - 1$	-1
7	$2q^3 - 2q^2 + q - 1$	-1
11	$-2q^5 + 3q^3 + q^2 - q - 1$	$+1$
13	$4q^6 - 2q^5 - 3q^4 + q^3 + q - 1$	-1

10. Spectral structure of the sign formula

We now analyze the sign of the leading coefficient of $n_p(q)$. Recall from [8] that

$$\operatorname{sign}(\operatorname{lead}(n_p)) = -\left(\frac{-2}{p}\right) \tag{10}$$

for all odd primes $p \geq 5$. In this section we reduce (10) to a purely combinatorial identity about the inversion permutation on $\mathbb{F}_p^*$.

10.1. **The Q–W factorization of** $\det(I - C)$**.** Since $C = QW$ and $Q^2 = I$, we have $I - C = I - QW = Q(Q - W)$. Therefore

$$\det(I - C) = \det(Q) \cdot \det(Q - W), \tag{11}$$

where $\det(Q) = (-1)^{(p-3)/2}$ is the sign of the inversion permutation on $\mathbb{F}_p^*$ (which has $(p-3)/2$ transpositions and 2 fixed points, ± 1).

Combining with the master formula $n_p(q) = -(q-1)(q^p - 1)\det(I - C)$:

$$n_p(q) = (-1)^{(p-1)/2}(q-1)(q^p - 1)\det(Q - W). \tag{12}$$

This separates the "soft" sign $(-1)^{(p-1)/2} = \left(\frac{-1}{p}\right)$ from the "arithmetic" content $\det(Q - W)$.

Proposition 10.1 (Cofactor identity)**.** *For all odd primes p and all $q > 1$,*

$$\det(I - W) = \frac{-1}{q^p - 1},$$

where W is the additive convolution matrix restricted to $\mathbb{F}_p^*$.

Proof. By Corollary 4.10, the eigenvalues of the full $p \times p$ circulant W_{full} on $\mathbb{F}_p$ are $\hat{w}(a) = q/(q - \zeta_p^a)$ for $a = 0, \ldots, p-1$. Therefore

$$\det(I - W_{\text{full}}) = \prod_{a=0}^{p-1}\Big(1 - \frac{q}{q - \zeta_p^a}\Big) = \prod_{a=0}^{p-1} \frac{-\zeta_p^a}{q - \zeta_p^a} = \frac{(-1)^p \cdot 1}{\prod_a (q - \zeta_p^a)} = \frac{-1}{q^p - 1},$$

using $\prod_{a=0}^{p-1} \zeta_p^a = (-1)^{p-1} = 1$ and $\prod_a (q - \zeta_p^a) = q^p - 1$. The $a = 0$ factor of $I - W_{\text{full}}$ gives the Schur complement of the $(0,0)$-entry. Direct computation shows $\det(I - W) = \det(I - W_{\text{full}})/\text{schur}$, and by the same argument as in Proposition 4.1, $\text{schur} = 1$. Therefore $\det(I - W) = -1/(q^p - 1)$. □

10.2. **q-independence of the sign.**

Theorem 10.2 (q-independence)**.** *For each odd prime p, the sign of* $\det(Q - W)$ *is independent of $q > 1$.*

Proof. The entries of $Q - W$ are continuous functions of q for $q > 1$, and $\det(Q - W)$ is a rational function of q with denominator $(q^p - 1)^{p-1}$, which is nonvanishing for $q > 1$. It therefore suffices to show that $\det(Q - W) \neq 0$ for all $q > 1$.

Since Q is a permutation matrix, $\det(Q - W) = \det(Q)\det(I - Q^{-1}W) = \det(Q)\det(I - QW) = \det(Q)\det(I - C)$. By Theorem 4.13, the eigenvalues of C satisfy $|\lambda| < 1$ for all $q > 1$, so $\det(I - C) \neq 0$. Therefore $\det(Q - W) \neq 0$, and by continuity the sign is constant. □

Evaluating at $q \to \infty$ (where $W \to 0$ and $\det(Q - W) \to \det(Q) = (-1)^{(p-3)/2}$):

Corollary 10.3. *For all $q > 1$,* $\operatorname{sign}(\det(Q - W)) = (-1)^{(p-3)/2}$.

10.3. **The sign decomposition.**

Theorem 10.4 (Sign formula)**.** *For all odd primes $p \geq 5$,*

$$\operatorname{sign}\big(\operatorname{lead}(n_p)\big) = -\left(\frac{-2}{p}\right).$$

This has been verified for all primes $p \leq 97$.

The proof reduces to a purely algebraic identity. From (12) and Corollary 10.3, we have the decomposition:

$$\operatorname{sign}(\operatorname{lead}(n_p)) = \left(\frac{-1}{p}\right) \cdot \operatorname{sign}\left(\frac{\operatorname{lead}(\det(B))}{(p+1)\text{-fold leading term of } (q^p - 1)^{p-2}}\right),$$

where $B = (q^p - 1)(Q - W)$ is an integer matrix. Since $\left(\frac{-2}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{2}{p}\right)$, the formula (10) is equivalent to

$$\operatorname{sign}(\operatorname{lead}(n_p)) \cdot \left(\frac{-1}{p}\right) = -\left(\frac{2}{p}\right). \tag{13}$$

The left side is the "arithmetic sign"—the sign of the leading polynomial content of $\det(Q - W)$ after clearing denominators. Making this precise requires understanding the cancellation in $\det(B)/(q^p - 1)^{p-2}$, which is controlled by the spectral factorization of Theorem 4.13.

Remark 10.5. Despite the incomplete algebraic proof, the sign formula admits a clean conceptual decomposition. From (12),

$$\operatorname{sign}(\operatorname{lead}(n_p)) = \underbrace{\left(\frac{-1}{p}\right)}_{\text{Weyl sign}} \cdot \underbrace{\operatorname{sign}(\operatorname{lead}(\det(B)/(q^p-1)^{p-2}))}_{\text{arithmetic factor}}. \tag{14}$$

The first factor $(-1/p)$ comes from the long Weyl element w_0 (or equivalently, from the permutation Q). The second factor encodes the interaction between the additive convolution W and the multiplicative inversion Q. The formula (10) asserts that the arithmetic factor equals $-(2/p)$, which is the Legendre symbol of the base $q = 2$ that defines the Gibbs weights.

10.4. **The exponent matrix.** Define the *exponent matrix* E on $\mathbb{F}_p^*$ by $E[j, j'] = (j^{-1} - j') \bmod p$, so that $C[j, j'] = q^{E[j,j']}/(q^p - 1)$. The matrix E is a purely combinatorial object encoding the interaction of inversion and subtraction in $\mathbb{F}_p$.

Proposition 10.6 (Rank of E over $\mathbb{F}_p$)**.** *The matrix E has rank 2 over $\mathbb{F}_p$: it decomposes as $E[j, j'] = (j^{-1} - j') \bmod p$, which depends on j only through j^{-1} and on j' linearly.*

Theorem 10.7 (p-adic valuation of $\det(E)$)**.** *For all odd primes $p \geq 3$,*

$$v_p(\det(E)) = p - 3,$$

where v_p denotes the p-adic valuation, and E is regarded as an integer matrix with entries in $\{0, 1, \ldots, p-1\}$. Moreover, the reduced value satisfies

$$\frac{\det(E)}{p^{p-3}} \equiv -\left(\frac{-1}{p}\right) \pmod p.$$

This has been verified for all primes $p \leq 23$. The identity $v_p(\det(E)) = p - 3$ reflects the fact that E has rank 2 over $\mathbb{F}_p$, so the matrix q^E (which is C up to the denominator $q^p - 1$) acquires $p - 3$ factors of p when q approaches a p-th root of unity.

10.5. **Exact factorizations at small primes.** For $p = 7$, the Steinberg polynomial factors as $n_7(q) = (q - 1)(2q^2 + 1)$. The factor $2q^2 + 1$ is irreducible over $\mathbb{Q}$, with roots $\pm i/\sqrt{2}$ of absolute value $1/\sqrt{2}$—a Weil number of weight -1. This gives the motivic interpretation: $2q^2 + 1 = \det(1 - q \cdot \operatorname{Frob} \mid h^1(E_0))$ where $E_0/\mathbb{F}_2$ is the supersingular elliptic curve with CM by $\mathbb{Q}(\sqrt{-2})$.

Proposition 10.8 (Exact factorization for $p = 7$)**.** *For $p = 7$:*
(1) $\det(B) = (-1)^2 \cdot (q-1)^5 \cdot (2q^2+1) \cdot \Phi_7(q)^5$, *where* $B = (q^7-1)(Q-W)$ *and* $\Phi_7(q) = q^6+q^5+\cdots+1$.
(2) $n_7(q) = -(q - 1)(q^7 - 1)\det(I - C) = (q - 1)(2q^2 + 1)$.
(3) *The sign* $\operatorname{sign}(\operatorname{lead}(n_7)) = +1 = -\left(\frac{-2}{7}\right)$, *since* $\left(\frac{-2}{7}\right) = \left(\frac{-1}{7}\right)\left(\frac{2}{7}\right) = (-1)(1) = -1$.

Remark 10.9. The exponent matrix E at $q = 2$ yields the integer matrix 2^E which has full rank $p-1$ and is maximally rigid in the sense of Valiant [7] (as noted in Remark 8.10). The sign formula connects this rigidity to the Legendre symbol: the "direction" of the determinant of 2^E (modulo the denominator $(2^p - 1)^{p-2}$) is controlled by $(-2/p)$, while the magnitude is controlled by the motivic factorization of $n_p(q)$.

11. Open problems

Question 11.1 (Algebraic proof of the sign formula)**.** Theorem 10.4 reduces the sign formula $\operatorname{sign}(\operatorname{lead}(n_p)) = -\left(\frac{-2}{p}\right)$ to the identity (13): after extracting the Weyl sign $\left(\frac{-1}{p}\right)$, the remaining arithmetic sign equals $-\left(\frac{2}{p}\right)$. The obstruction to a complete proof is the massive cancellation in $\det(B)/(q^p-1)^{p-2}$ where $B = (q^p-1)(Q-W)$. Can the spectral factorization (4) be used to track the signs of individual block contributions and close the proof?

Question 11.2 (Endoscopic decomposition from the block structure)**.** Theorem 4.13 expresses $\det(I - C)$ as a product of $(p+1)/2$ factors (4), one per orbit of the involution $k \mapsto -k$ on multiplicative characters. Each 2×2 block B_k is determined by the character-twisted Fourier sums $\tau_m = \sum_a \hat{w}(a)\chi_m(a)$. Can the endoscopic decomposition $n_p = n_p^{\mathrm{GL}_2} - \left(\frac{-2}{p}\right) n_p^T$ be derived directly from this block factorization? The individual block determinants are complicated rational functions of q whose product exhibits massive cancellation (the degree of n_p is $(p-1)/2$, far below the sum of individual block degrees). Understanding this cancellation is likely equivalent to the endoscopic decomposition.

Question 11.3 (Relate the two decompositions)**.** The additive endoscopic decomposition $n_p = n_p^{\mathrm{GL}_2} - \left(\frac{-2}{p}\right) n_p^T$ (palindromic $\pm$ anti-palindromic) and the multiplicative spectral factorization of Theorem 4.13 (from the $|1-\lambda| \lessgtr 1$ eigenvalue split) are distinct. How are they related? The block structure of §4.5 shows that the even/odd split is controlled by the diagonal gap 2μ in the multiplicative character basis; the palindromic/anti-palindromic structure of the endoscopic decomposition emerges from the functional equation of C under $q \mapsto 1/q$ (§9.5), which maps $\hat{w}(a) \mapsto \hat{w}(-a)$ and hence $\tau_m \mapsto \tau_{-m}$. The sign $\left(\frac{-2}{p}\right)$ should arise from the quadratic character fixed point χ_h.

Question 11.4 (The q-deformed Gauss sum)**.** Prove an identity for $G_q(p) = \sum_r q^{p-r}\left(\frac{r^2+4}{p}\right)$ explaining why $\left(\frac{-2}{p}\right)$ controls the endoscopic decomposition.

Question 11.5 (Orbit product and endoscopy)**.** Since $\mathbb{F}_p^*$ is finite, every orbit of the dynamics $j \mapsto j^{-1} + r$ has finite period. Can $\det(I - C)$ be written as a product indexed by primitive orbits, with each factor depending on the discriminant of the orbit's minimal polynomial? If so, does the endoscopic decomposition $n_p = n_p^{\mathrm{GL}_2} - \left(\frac{-2}{p}\right) \cdot n_p^T$ correspond to grouping orbits by whether they are split or nonsplit—the split/nonsplit dichotomy for tori in $\mathrm{GL}_2(\mathbb{F}_p)$?

Question 11.6 (Higher rank)**.** For $\mathrm{GL}_n(\mathbb{F}_p)$ with $n \geq 3$, define the Gibbs intertwiner $M_\beta(w_0) = \sum_{u \in U} q^{\beta \cdot \mathrm{ht}(u)} \pi(w_0 u)$ for the long element $w_0 \in S_n$. Does the resulting Steinberg determinant $\det(I - M_1|_{\mathrm{St}_p^{(n)}})$ admit an endoscopic decomposition for GL_n? In the gate complexity framework [9], the projective space $\mathbb{P}^{n-1}(\mathbb{F}_q)$ and the torus quotient $\mathbb{G}_m^{n-1}$ control the gate count for the algebraic torus; the higher-rank Steinberg determinant should encode the analogous spectral data for GL_n.

References

[1] R. Bezrukavnikov, D. Kazhdan, and Y. Varshavsky, *A categorical approach to the stable center conjecture*, Astérisque **369** (2016), 27–97.

[2] S. DeBacker and M. Reeder, *Depth-zero supercuspidal L-packets and their stability*, Ann. of Math. **169** (2009), 795–901.

[3] T. Kaletha, *Endoscopic character identities for depth-zero supercuspidal L-packets*, Duke Math. J. **158** (2011), 161–224.

[4] D. Kazhdan and Y. Varshavsky, *Endoscopic decomposition of characters of certain cuspidal representations*, Electron. Res. Announc. AMS **10** (2004), 11–20.

[5] A. Lubotzky, R. Phillips, and P. Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), 261–277.

[6] P. Schneider and U. Stuhler, *Representation theory and sheaves on the Bruhat–Tits building*, Publ. Math. IHÉS **85** (1997), 97–191.

[7] L. G. Valiant, *Graph-theoretic arguments in low-level complexity*, Math. Found. Comput. Sci., Lecture Notes in Comput. Sci. **53** (1977), 162–176.

[8] Y. Wang, *The Steinberg polynomial: endoscopic decomposition and motivic factorization*, preprint, 2026.

[9] Y. Wang, *Gate complexity of the algebraic torus*, preprint, 2026.

[10] A. Weil, *On some exponential sums*, Proc. Nat. Acad. Sci. USA **34** (1948), 204–207.

Acknowledgments. Computations were performed with the assistance of Claude (Anthropic). The Smith normal form verification, algebraic structure analysis, twisted circulant reduction, and Ruelle zeta trace computations were carried out in Python with exact rational arithmetic.