# Fast Fourier Analysis for $\mathrm{SL}_2$ over a Finite Field and Related Numerical Experiments

John D. Lafferty and Daniel Rockmore

**CONTENTS**



We study the complexity of performing Fourier analysis for the group $\mathrm{SL}_2(\mathbf{F}_q)$, where $\mathbf{F}_q$ is the finite field of $q$ elements. Direct computation of a complete set of Fourier transforms for a complex-valued function $f$ on $\mathrm{SL}_2(\mathbf{F}_q)$ requires $q^6$ operations. A similar bound holds for performing Fourier inversion. Here we show that for both operations this naive upper bound may be reduced to $O(q^4 \log q)$, where the implied constant is universal, depending only on the complexity of the "classical" fast Fourier transform. The techniques we use depend strongly on explicit constructions of matrix representations of the group.

Additionally, the ability to construct the matrix representations permits certain numerical experiments. By quite general methods, recent work of others has shown that certain families of Cayley graphs on $\mathrm{SL}_2(\mathbf{F}_q)$ are expanders. However, little is known about their exact spectra. Computation of the relevant Fourier transform permits extensive numerical investigations of the spectra of these Cayley graphs. We explain the associated calculation and include illustrative figures.

## 1. INTRODUCTION

### Fast Fourier Analysis

We begin by recalling some definitions. Let $G$ be a finite group and $L^2(G)$ the algebra of complex-valued functions on $G$ with respect to convolution. Fix a complete set $\mathcal{R}$ of inequivalent irreducible representations of $G$. Then

$$\sum_{\rho \in \mathcal{R}} d_\rho^2 = |G|,$$

where $d_\rho$ is the degree of the representation $\rho$.

If $f \in L^2(G)$, the *Fourier transform of $f$ at $\rho$*, denoted $\hat{f}(\rho)$, is the matrix

$$\hat{f}(\rho) = \sum_{s \in G} f(s)\rho(s).$$

AMS Subject Classification: 20-04, 05C25, 20C30

Rockmore was supported in part by a National Science Foundation Mathematical Sciences Postdoctoral Fellowship.


1058-6458/92 \$0.50 per page

The (discrete) *Fourier transform*, or DFT, of $f$ (with respect to $\mathcal{R}$) is the set of matrices $\{\hat{f}(\rho)\}_{\rho\in\mathcal{R}}$.

The Fourier transform of $f$ determines $f$ via the *Fourier inversion formula*

$$f(s) = \frac{1}{|G|}\sum_{\rho\in\mathcal{R}} d_\rho \operatorname{tr}(\hat{f}(\rho)\rho(s^{-1})).$$

Let $T(G)$ denote the minimal number of operations needed to compute a Fourier transform of $f$, with the complete set of representations $\mathcal{R}$ and the function $f$ given as initial data. Similarly, let $I(G)$ denote the minimal number of operations needed to recover the function $f$ from a Fourier transform $\{\hat{f}(\rho)\}_{\rho\in\mathcal{R}}$ via Fourier inversion.

By direct computation, a naive upper bound of $|G|^2$ is obtained for both $T(G)$ and $I(G)$. In this paper we examine this computation for the group $G = \mathrm{SL}_2(K)$, with $K$ a finite field, and derive fast algorithms for Fourier analysis for this situation. These algorithms depend on certain explicit constructions of the matrix representations for this group. The construction of these representations also enables us to obtain a wealth of numerical data for certain interesting Cayley graphs for $\mathrm{SL}_2(K)$.

A remark concerning complexity results is in order. Our complexity estimates are given in the *linear computational model*, which quickly seems to be becoming standard in the analysis of generalized DFT algorithms [Baum and Clausen 1991; Baum et al. 1991]. That is, computation of a DFT over a finite group $G$ may be viewed as the evaluation of a certain $|G|\times|G|$ complex matrix at an arbitrary vector $f$. If $A$ is any $r\times t$ complex matrix and $b\geq 2$, the *b-linear complexity* $L_b(A)$ of $A$ is defined to be the minimal number of linear operations (complex additions, subtractions and scalar multiplications) needed to compute the product $Ax$ for an arbitrary vector $x$, where scalar multiplication is restricted to scalars of absolute value at most $b$. In this model the $b$-linear complexity of a group $G$ is defined to be the minimum $b$-linear complexity over all possible DFTs for $G$. For comparison and adaptation of related results, our $T(G)$ is in fact the 2-linear complexity of $G$. Analogously, $I(G)$ is the minimal 2-linear complexity of a Fourier inversion matrix, under the same formulation.

There have been several recent advances in the development of fast algorithms for performing Fourier analysis on finite groups. Of relevance here are the techniques developed for treating arbitrary finite groups [Clausen 1989a; Diaconis and Rockmore 1990]. In brief, these algorithms rely on deriving a recurrence for the computation with respect to a subgroup. It is useful to briefly review the main idea for speeding the computation of a Fourier transform.

Let $G$ be a group and $H\leq G$ a subgroup. Fix a set of coset representatives $\{s_1,\dots,s_k\}$ for $G/H$. If $\rho$ is a matrix representation of $G$, we can expand $\hat{f}(\rho)$ as

$$\begin{aligned}\hat{f}(\rho) &= \sum_{i=1}^{k}\sum_{t\in H} f(s_it)\rho(s_it)\\ &= \sum_{i=1}^{k}\rho(s_i)\sum_{t\in H} f_i(t)\rho(t),\end{aligned}$$

where $f_i\in L^2(H)$ is defined by $f_i(t) = f(s_it)$. Thus, if $\rho{\downarrow}H$ denotes the restriction of $\rho$ to $H$, we see that the last sum may be rewritten as

$$\hat{f}(\rho) = \sum_{i=1}^{k}\rho(s_i)\hat{f}_i(\rho{\downarrow}H).$$

In general, $\rho{\downarrow}H$ need not remain irreducible. Assume that

$$\rho\sim\eta_1\oplus\cdots\oplus\eta_r,$$

where each $\eta_i$ is an irreducible representation of $H$ and $\sim$ denotes equivalence of representations. In the language of matrices, this direct sum decomposition means that there exists a basis in which the restrictions to $H$ of the representations $\{\rho_i\}$ are block diagonal, with the matrices for the $\{\eta_j\}$ on the diagonal. Such "$H$-adapted bases" can always be found. Consequently, the restricted transforms $\{\hat{f}_i(\rho{\downarrow}H)\}$ can be built from the collection of precomputed transforms $\{\hat{f}_i(\eta_j)\}$. This allows us to write the following recurrence for $T(G)$ [Clausen 1989a, Theorem 1.1; Diaconis and Rockmore 1990, Theorem 1]:

$$T(G)\leq\frac{|G|}{|H|}T(H)+\frac{|G|}{|H|}\sum_{\rho} d_\rho^\alpha \tag{1.1}$$

where $\alpha$ is the exponent of the complexity bound for matrix multiplication.

### Fourier Analysis for $\mathrm{SL}_2(q)$

In attempting to apply (1.1) to $G = \mathrm{SL}_2(K)$ we find an obstruction to heightened efficiency.

Let $K = \mathbf{F}_q$ be the finite field of $q$ elements, where $q$ is a power of the prime $p$. We will write $\mathrm{SL}_2(q)$ for $\mathrm{SL}_2(\mathbf{F}_q)$, and denote the associated complexities as $T(q)$ and $I(q)$.

A natural subgroup for restriction is $B = B(q)$, the subgroup of upper triangular matrices. This is a *metabelian* group; that is, it contains an abelian normal subgroup $U$ such that the quotient $B/U$ is abelian. It is known [Clausen 1989b; Rockmore 1990a] that for such a group we have

$$T(B) \leq O(|B| \log |B|).$$

As will be explained fully in Section 2, the representations of $\mathrm{SL}_2(q)$ occur essentially as $q$ irreducible representations of degree $q$. Thus, (1.1) now specializes to

$$\begin{aligned} T(q) &\leq \frac{|\mathrm{SL}_2(q)|}{|B|} T(B) + \frac{|\mathrm{SL}_2(q)|}{|B|} \sum_{i=1}^{p} d_\rho^\alpha \\ &\leq O((q+1)\cdot q^3 \log q + q \cdot q q^\alpha) \\ &\leq O(q^4 \log q + q^{\alpha+2}). \end{aligned}$$

In most applications $\alpha = 3$, so the term $O(q^5)$, coming from matrix multiplication, dominates.

We are able to get around this by finding certain bases for the representations that allow us to reduce the number of matrix multiplications. In particular, we have:

**Theorem 1.1.** *The number $T(q)$ of operations necessary to compute a Fourier transform of a function $f \in L^2(\mathrm{SL}_2(q))$ is $O(q^4 \log q)$.*

In the proof of this result we will compute an explicit constant for the bound.

By general considerations, Baum and Clausen show that complexity bounds for computation of the DFT of a group $G$ in turn give bounds for the complexity of Fourier inversion. More precisely, in the matrix formulation discussed above, computation of Fourier inverses with respect to a given set of irreducible representations of $G$ is "almost" the same as evaluation of the transpose of the associated DFT matrix at an arbitrary complex vector [Baum and Clausen 1991, Theorem 1]. In fact, if $A$ is a given DFT matrix for $G$, so that $A^{-1}$ is the associated Fourier inversion matrix, it follows from [Baum and Clausen 1991, Theorem 3] that

$$L_b(A^{-1}) \leq L_b(A) + |G|\,.$$

Hence, Theorem 1.1 implies a like upper bound for $I(q)$:

**Theorem 1.2.** *The number $I(q)$ of operations needed to recover a function $f \in L^2(\mathrm{SL}_2(q))$ from its Fourier transform is $O(q^4 \log q)$.*

It is worth pointing out that the relation between the transform and inversion bounds is obtained by recognizing that the algorithm for computing a Fourier transform is a *linear algorithm.* Such an algorithm is realized as a directed acyclic graph with additions and subtractions labeling the nodes, and scalars (for multiplication) labeling the edges. In this setting, the computation of the transposed matrix product is essentially given by a linear algorithm in which the arrows are reversed [Bshouty et al. 1988]. In Section 7 we give a more explicit realization of the Fourier inversion algorithm, which still yields the asserted bound.

### Fourier Analysis, Graphs and Eigenvalues

The ability to compute matrix representations can prove to be a great aid in numerical investigations of Cayley graphs. Let $G$ be a finite group and $S \subset G$ a subset of $G$ such that $S = S^{-1}$. The *Cayley graph* $X = X(G,S)$ of $G$ with respect to $S$ is the undirected graph with vertex set $G$ and having an edge between $a$ and $b$ if and only if $as = b$ for some (necessarily unique) $s \in S$.

The *adjacency matrix* of a graph with $m$ vertices is the $m \times m$ matrix (with rows and columns indexed by vertices of the graph) whose entries are 1 or 0, depending on whether or not there is an edge joining the vertices corresponding to the entry's row and column. The *spectrum* of a graph is the spectrum of its adjacency matrix. Various connectivity and "network" properties of a graph can be judged by studying its spectrum. One such property centers on the notion of expansion, which measures the number of neighbors of a vertex subset of a graph.

**Definition 1.3.** A $k$-regular graph $G = (V, E)$, with $n = |V|$ vertices and edge set $E$, is an $(n, k, c)$-*expander* if

$$|\partial A| \geq c\left(1 - \frac{|A|}{n}\right)|A|$$

for every subset $A \subset V$, where $\partial A = \{y \in V \setminus A : (y, x) \in E \text{ for some } x \in A\}$.

Since every $k$-regular graph is an $(n, k, c)$-expander for some $c > 0$, this definition is intended to be applied to families of graphs, typically with $n \to \infty$ and $k$ and $c$ held fixed. We refer to [Bien 1989; Lubotzky; Sarnak 1990] for complete descriptions and references concerning the mathematics of expanders. Here we limit ourselves to a brief summary of the known relations between the spectrum of a graph and the expansion coefficient $c$. The most striking of these connections stems from discrete analogues of inequalities relating the spectrum of the Laplacian on a finite-volume Riemannian manifold to its Cheeger constant.

Recall that a *combinatorial Laplacian* may be defined on a graph $X = (V, E)$ as follows. The choice of an orientation for each edge of the graph gives rise to a natural complex $d : L^2(V) \to L^2(E)$, which may be thought of as the $|E| \times |V|$ matrix given by

$$(d)_{(e,v)} = \begin{cases} 1 & \text{if } v = (e, f) \text{ for some } f \in V, \\ -1 & \text{if } v = (f, e) \text{ for some } f \in V, \\ 0 & \text{otherwise.} \end{cases}$$

The combinatorial Laplacian $\Delta : L^2(V) \to L^2(V)$ is then realized as the $|V| \times |V|$ matrix $d^*d$, and it is a simple matter to show that

$$\Delta f(v) = \deg(v) f(v) - \sum_{w \in V} A_{(v,w)} f(w),$$

where $A$ is the adjacency matrix. In particular, when $X$ is $k$-regular, which is the only case we shall consider, we have $\Delta = kI - A$, where $I$ is the identity matrix.

The *Cheeger constant* $h(X)$ of the graph $X$ is defined in analogy with the Riemannian case by setting

$$h(X) = \inf_{A,B \subset V} \frac{|E(A, B)|}{\min(|A|, |B|)},$$

where $E(A, B) = \{e = (x, y) \in E : x \in A, y \in B\}$ is the set of edges connecting $A$ and $B$. It is easy to show that every graph $X$ is an $(n, k, h(X)/k)$-expander. Conversely, for an $(n, k, c)$-expander, the inequality $h(X) \geq c/2$ holds.

The connection with the spectrum comes from a discrete version of Cheeger's inequality for Riemannian manifolds [Alon 1983]:

$$\lambda_1(X) \geq \frac{h^2(X)}{2k},$$

where $\lambda_1(X)$ is the smallest nonzero eigenvalue of the Laplacian. A partial converse to this discrete Cheeger inequality has been proved [Alon and Milman 1985]:

$$h(X) \geq \tfrac{1}{2}\lambda_1(X).$$

For $k$-regular graphs, $k$ is the largest eigenvalue of the adjacency matrix. If we order the eigenvalues $\mu_i$ as $k = \mu_n > \mu_{n-1} \geq \cdots \geq \mu_1$, we have $\lambda_1 = k - \mu_{n-1}$ and $\mu_1 \geq -k$, with equality precisely when the graph $X$ is bipartite.

An additional graph-theoretic invariant is related to the low end of the spectrum [Biggs 1974]. The vertex chromatic number $\nu(X)$ is bounded below as a result of the inequality

$$\nu(x) \geq 1 - \frac{k}{\mu_1}.$$

If we set $\mu = \max_{i \neq n} |\mu_i|$, graphs with small $\mu$ relative to $k$ are not only good expanders, but also have high chromatic number.

Finally, it is worth mentioning that the second eigenvalue of the adjacency matrix also bounds the diameter of the graph, as a result of the inequality [Chung 1989; Sarnak 1990]

$$\operatorname{diam}(X) \leq \left\lceil \frac{\log(n-1)}{\log(k/\mu)} \right\rceil.$$

For more references and a proper discussion of all these results, see [Bien 1989; Lubotzky; Sarnak 1990].

It is known that certain families of Cayley graphs for $\mathrm{SL}_2(q)$ are expanders. In particular, it is shown in [Lubotzky] that the uniform bound $\lambda_1(\Gamma\backslash\mathbf{H}) \geq \frac{3}{16}$ of Selberg's theorem, where $\Gamma$ is a discrete congruence subgroup of $\mathrm{SL}_2(\mathbf{R})$ acting on the hyperbolic plane $\mathbf{H}$, implies that the Cayley graphs

$$X_p = X(\mathrm{SL}_2(p), \mathcal{G}_1)$$

form a family of expanders, where $\mathcal{G}_1$ is the generating set

$$\mathcal{G}_1 = \left\{ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\}.$$

The proof effectively transfers the spectral bound on the manifold to a lower bound on the expansion coefficient of the graphs. In particular, in the absence of further information on this coefficient, the validity of Selberg's conjecture that $\lambda_1(\Gamma\backslash\mathbf{H}) \geq \frac{1}{4}$ would make these graphs more attractive.

The significance of Fourier analysis for the numerical study of the spectrum of a Cayley graph lies in the following link. If $\delta_S$ is the characteristic function of the subset $S$ defined on $G$, the adjacency matrix for $X(G, S)$ is precisely the Fourier transform of $\delta_S$ at the regular representation of $G$. Consequently, the spectrum of $X(G, S)$ is the collection of eigenvalues that occur in the Fourier transforms of $\delta_S$ at a complete set of irreducible representations of $G$. Since the dimension of any given irreducible representation of $G$ cannot exceed $|G|^{1/2}$, the corresponding numerical analysis is much faster.

For example, in the case $G = SL_2(p)$, direct numerical analysis would require that the eigenvalues of a single matrix of size $p^3$ be found. This would require $O(p^9)$ operations. However, by using the Fourier transforms, we instead determine the eigenvalues of $p$ matrices of size $p$, which requires only $O(p \cdot p^3) = O(p^4)$ operations. So, while for primes greater than 10 the full adjacency matrices are already too large to consider, by working with individual Fourier transforms we can consider primes on the order of 500.

When $S$ is not simply an arbitrary subset, but instead a union of conjugacy classes, the analysis simplifies further. Still assuming $S = S^{-1}$, one may show [Diaconis 1988] that the eigenvalues $\mu_i$ of the adjacency matrix are exactly the average values of the irreducible characters:

$$\mu_i = \frac{1}{\dim \rho_i} \sum_{s \in S} \operatorname{tr} \rho_i(s).$$

Using this correspondence, [Lubotzky] uses character tables to tabulate the eigenvalues and expanding properties for $SL_2(q)$ and various unions of conjugacy classes. In considering $S$ to be a union of conjugacy classes, however, one obtains a family of $k(q)$-regular graphs $X_q$, where $k(q)$ increases with the size of the graph.

In this paper we are interested in Cayley graphs for $SL_2(q)$ with respect to sets of generators, such as $\mathcal{G}_1$, which are of fixed size and not a union of conjugacy classes. In this situation the full set of irreducible representations, and not just the characters, is ostensibly required in order to obtain the spectrum.

In Section 2 we review briefly the representation theory of $SL_2(K)$. Section 3 details the algorithm for efficient computation of the Fourier transform over $SL_2(K)$. The explicit constructions of Section 3 are then followed by their application to the investigation of Cayley graphs for $SL_2(K)$ in the next two sections. In Section 4 we discuss implementation aspects of the experiment, while in Section 5the numerical results are presented and explained. Section 6 contains some closing remarks and open questions. We postpone the discussion of Fourier inversion and convolution to an appendix (Section 7), so as to not interrupt the flow from theory to application in Sections 4 and 5.

## 2. REPRESENTATION THEORY FOR $SL_2$

In what follows, $K = \mathbf{F}_q$ will denote the finite field of $q$ elements, where $q = p^n$ for some prime $p \neq 2$. Several important subgroups of $SL_2(q)$ must be distinguished. Let $U \leq SL_2(q)$ be the subgroup of unipotent matrices:

$$U = \left\{ \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} : u \in K \right\}.$$

Let $T \leq SL_2(q)$ denote the subgroup of diagonal matrices:

$$T = \left\{ \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} : \alpha \in K^\times \right\}.$$

Let $B \leq SL_2(q)$ denote the subgroup of upper triangular matrices:

$$B = \left\{ \begin{pmatrix} \alpha & u \\ 0 & \alpha^{-1} \end{pmatrix} : \alpha \in K^\times, u \in K \right\}.$$

Note that $U$ is isomorphic to $K$ considered as a group additively (denoted $K^+$), while $T$ is isomorphic to the cyclic group $K^\times$. Also, $U$ is normal in $B$; in fact, $B$ is the semidirect product of $U$ by $T$.

Lastly, set $w = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Note that $w$ is of order 4 and that

$$w^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The theory of the representations of $\mathrm{SL}_2(q)$ is well known. They fall into two classes, the *principal series* and the *discrete series*. Discrete series representations are also sometimes called *cuspidal*. Essentially, the main difference between the two classes is that for a discrete series representation $\rho : \mathrm{SL}_2(q) \to \mathrm{GL}(V_\rho)$ there is no $U$-fixed vector—that is, no nonzero vector $v \in V_\rho$ such that $\rho(u)v = v$ for all $u \in U$. If $\rho$ is a principal series representation, there is such a vector. Another way to say this is that an irreducible representation $\rho$ of $\mathrm{SL}_2(q)$ is cuspidal if and only if $\rho{\downarrow}U$ does not contain the trivial representation. (If $\rho$ is a representation of a group $G$ and $H \leq G$ is a subgroup, we denote by $\rho{\downarrow}H$ the representation of $H$ given by restriction of $\rho$ to $H$.)

The character table of $\mathrm{SL}_2(q)$ has been known for a long time [Schur 1907; Jordan 1907]. However, the discovery of actual realizations of these representations as group actions on vector spaces is more recent and is generally attributed to Kloosterman [1946] and Tanaka [1967]. Our synopsis follows [Naimark and Stern 1980, 150–160].

### Construction of the Principal Series Representations

The principal series representations of $\mathrm{SL}_2(q)$ are constructed as induced representations. We recall that if $G$ is a group, $H \leq G$ is a subgroup, and $\eta$ is a representation of $H$ in a vector space $V_\eta$, a representation of $G$ may be obtained as follows: Let $\mathrm{Ind}(V_\eta)$ denote the vector space of functions $f : G \to V_\eta$ such that

$$f(st) = \eta(s) f(t) \tag{2.1}$$

for all $s \in H$. There is a representation of $G$ on $\mathrm{Ind}(V_\eta)$ by right translation,

$$(\rho(g)f)(t) = f(tg).$$

This is called the *representation of $G$ induced from $H$ by $\eta$*, and is denoted as $\eta{\uparrow}G$. Note that the dimension of the induced representation is $[G{:}H]d_\eta$, where $d_\eta$ is the dimension of $V_\eta$.

The principal series representations are obtained by inducing characters from $B$ to $\mathrm{SL}_2(q)$. More precisely, the irreducible representations of $T$ are all one-dimensional, given by characters. If $\alpha$ is a generator for $K^\times$, the characters are defined by

$$\psi_j(\alpha^k) = \exp\left(\frac{2\pi i jk}{q-1}\right),$$

where $j$ takes all values between 0 and $q-2$. It is easy to check that any $\psi_j$ extends to a one-dimensional representation (or character) of $B$, denoted $\tilde{\psi}_j$, by

$$\tilde{\psi}_j\left(\begin{pmatrix} k & u \\ 0 & k^{-1} \end{pmatrix}\right) = \psi_j(k).$$

Recall that, if $A$ is any abelian group, the set of characters of $A$ is a group isomorphic to $A$, called the *dual group* to $A$ and denoted by $\hat{A}$. In the case of $K^\times$, two characters are to be singled out: the *trivial* character that maps every element to 1 ($\psi_0$ in the notation above), and the sgn character, the unique nontrivial square root of the trivial character (sgn $= \psi_{(q-1)/2}$). The trivial character is often denoted simply as 1.

Let $\rho_\psi$ denote $\tilde{\psi}{\uparrow}\,\mathrm{SL}_2(q)$, where $\psi$ is any character of $K^\times$.

**Theorem 2.1.** *Let $\psi_1, \psi_2$ be characters of $K^\times$.*

(i) *Suppose that $\psi_i^2 \neq 1$, for $i = 1, 2$. Then $\rho_{\psi_i}$ is irreducible (of dimension $q+1$). Furthermore, $\rho_{\psi_1}$ and $\rho_{\psi_2}$ are equivalent if and only if $\psi_1 = \psi_2$ or $\psi_1^{-1} = \psi_2$.*
(ii) *Let $\psi_1 =$ sgn. Then $\rho_{\psi_1}$ is equivalent to the direct sum of two inequivalent irreducible representations, each of degree $\frac{1}{2}(q+1)$.*
(iii) *$\rho_1$ is equivalent to the direct sum of the trivial representation of $\mathrm{SL}_2(q)$ and an irreducible $q$-dimensional representation of $\mathrm{SL}_2(q)$.*

*These are all the principal series representations.*

The explicit construction of the matrix representations is treated more carefully in Section 3, in considering the computation of Fourier transforms at these representations.

### Construction of the Discrete Series Representations

The discrete series representations may be realized in several ways. The method given here is a combination of ideas due to Silberger and Piatetski-Shapiro.

One way of constructing the discrete series representations for $\mathrm{SL}_2(q)$ is to first construct the discrete-series representations for $\mathrm{GL}_2(q)$, and then take advantage of the fact that the restrictions of these representations to $\mathrm{SL}_2(q)$ are mostly irreducible. It is then only necessary to pick out a subset of these whose restrictions are inequivalent [Silberger 1969].

The following construction of the discrete series for $\mathrm{GL}_2(q)$ follows [Piatetski-Shapiro 1983].

Let $L$ denote the unique quadratic extension of $K$. (Being a finite field, $K$ has a unique quadratic extension, given by adjoining the square root of any nonsquare in $K$. For example, one can adjoin the square root of any generator of the cyclic group $K^\times$.) The Galois group of $L/K$ consists of two elements, the identity map and the Frobenius map, in this case given by raising any given element to the $q$-th power. Recall that the norm map $\mathcal{N} : L \to K$, given by

$$\mathcal{N}(\alpha) = \alpha^{q+1},$$

is surjective onto $K^\times$. The subset $C \subseteq L^\times$ consisting of elements of norm 1 is a cyclic subgroup of $L^\times$ of order $q+1$. Call a character of $L^\times$ *decomposable* if its restriction to $C$ is trivial, that is, if $\psi(c) = 1$ for all $c \in C$. Otherwise, call it *nondecomposable.* To say it another way, consider the group homomorphism $R : \widehat{L^\times} \to \hat{C}$ given by restriction,

$$R(\psi)(c) = \psi(c)$$

for all $c \in C$. Then $R$ is surjective and its kernel equals the set of decomposable characters. Thus, the fiber over each character of $C$ has order $q-1$; in particular, there are $q-1$ decomposable characters, and hence $q^2 - q$ nondecomposable characters. In fact, the decomposable characters may be constructed directly by composing any character of $K^\times$ with the norm map.

There is a natural correspondence between nondecomposable characters of $L^\times$ and discrete series representations of $\mathrm{GL}_2(q)$. If $\nu$ is a nondecomposable character of $L^\times$, let $\rho_\nu$ denote the corresponding discrete series representation of $\mathrm{GL}_2(q)$, which we now construct.

Using the Bruhat decomposition of $\mathrm{GL}_2(q)$,

$$\mathrm{GL}_2(q) = DUwU \amalg DU,$$

where $U$ is as above and

$$D = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} : ab \neq 0 \right\},$$

it is enough to define the representation on $U$, $D$ and the matrix $w$, and then to check certain compatibility conditions.

To define the representation, fix some nontrivial character $\chi$ of $K^+$, as follows: if $q = p$, set $\chi(j) = e^{2\pi i j/p}$; otherwise, set $\chi(j) = e^{2\pi i (\mathrm{tr}\, j)/p}$, where tr denotes the trace map from $K$ to $\mathbf{F}_p$, the finite field of $p$ elements.

Any discrete series representation of $\mathrm{GL}_2(q)$ can be realized as a group action of $\mathrm{GL}_2(q)$ on the vector space $V_{\rho_\nu}$ of complex-valued functions on $K^\times$. Let $f : K^\times \to \mathbf{C}$ be any function in $V_{\rho_\nu}$. Setting $t_u = \left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right)$ and $d_{a,b} = \left(\begin{smallmatrix} a & 0 \\ 0 & b \end{smallmatrix}\right)$, and recalling that $w = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, define

$$(\rho_\nu(t_u)f)(x) = \chi(xu)f(x), \tag{2.2}$$

$$(\rho_\nu(d_{a,b})f)(x) = \nu(b)f(a^2x), \tag{2.3}$$

$$(\rho_\nu(w)f)(x) = \sum_y \nu(y^{-1})j_\nu(xy)f(y), \tag{2.4}$$

where

$$j_\nu(z) = \frac{1}{q} \sum_{\mathcal{N}(t)=z} \chi(t + t^q)\nu(t)$$

and the sum here is over $t \in L^\times$. Note that $t + t^q$ is just the trace of $t$ from $L$ to $K$, and therefore lies in $K$.

Now extend the map to all of $\mathrm{GL}_2(q)$ by multiplication; that is, define $\rho_\nu(g)$, for $g \in \mathrm{GL}_2(q)$, by expressing $g$ as a product of matrices $t_u$, $d_{a,b}$ and $w$, which is always possible. It can be shown [Piatetski-Shapiro 1983, 38–40] that $\rho_\nu(g)$ does not depend on the choice of decomposition for $g$, so the definition makes sense.

**Theorem 2.2.** *The representation $\rho_\nu$ of $\mathrm{GL}_2(q)$ is irreducible. For nondecomposable characters $\nu$ and $\nu'$, the representations $\rho_\nu$ and $\rho_{\nu'}$ are equivalent if and only if either $\nu = \nu'$ or $\nu$ is equal to the composition of $\nu'$ with the nontrivial element of $\mathrm{Gal}(L/K)$ (that is, if $\nu(\alpha) = \nu'(\alpha^q)$ for all $\alpha \in L$). These are all of the discrete series representations of $\mathrm{GL}_2(q)$.*

The relation to the discrete series representations of $\mathrm{SL}_2(q)$ is as follows.

**Theorem 2.3.** *Let $\rho_\nu$ be the discrete series representation of $\mathrm{GL}_2(q)$ defined above and let the same notation denote its restriction to $\mathrm{SL}_2(q)$.*

(i) *$\nu$ and $\nu'$ have the same restriction to $C$ if and only if $\rho_\nu$ and $\rho_{\nu'}$ are equivalent.*
(ii) *If $\nu^2$ is not the identity on $C$, then $\rho_\nu$ is an irreducible representation of* $\mathrm{SL}_2(K)$.
(iii) *Suppose $\nu$ is nondecomposable and $\nu^2$ is trivial on $C$. Then $\rho_\nu$ is equivalent to the direct sum of two inequivalent irreducible representations of degree $\frac{1}{2}(q-1)$.*

*This constructs all the discrete series representations of* $\mathrm{SL}_2(q)$.

Thus, a complete set of discrete series representations for $\mathrm{SL}_2(q)$ may be given by choosing a set of coset representatives for $\widehat{L^*}/\hat{C}$, constructing the associated discrete series representations for $\mathrm{GL}_2(K)$ (except at the identity coset), and then decomposing the discrete series representation corresponding to the nondecomposable character whose square is trivial on $C$.

## 3. COMPUTATION OF FOURIER ANALYSIS

In this section we give algorithms for performing Fourier analysis on $\mathrm{SL}_2(q)$. The naive upper bound $q^6$ for both $T(q)$ and $I(q)$ is reduced to $O(q^4 \log q)$, where the implied constant is universal and depends only on the complexity of the classical FFT (fast Fourier transform) for abelian groups.

**Theorem 3.1.** [Baum et al. 1991, Theorem 3] *Let $A$ be any finite abelian group. Then*

$$T(A) = I(A) \leq 8\,|A| \log |A|\,.$$

In both cases—Fourier transforms and Fourier inversion—the computation may be split into two parts, one taking place at the principal series and one taking place at the discrete series (compare the two subsections of Section 3).

The algorithms involve finding computationally tractable bases for the representations. Along the way, explicit formulas for matrices representing the elements $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)$ are given. This will permit some explicit calculations to be done for investigation of the spectrum of certain Cayley graphs on $\mathrm{SL}_2(q)$ (see Section 4).

### Fourier Analysis at the Principal Series

As explained in Section 2, the principal series representations are essentially constructed as induced representations from the subgroup $B$. They occur as (i) $\frac{1}{2}(q-1)$ representations of degree $q+1$; (ii) two representations of degree $\frac{1}{2}(q+1)$ and (iii) one representation of degree $q$. Thus, direct computation of the Fourier transforms at all of these representations takes

$$(q^3-q)\left(\tfrac{1}{2}(q-1)(q+1)^2+2\left(\tfrac{1}{2}(q+1)\right)^2+q^2\right) = O(q^6)$$

operations. In this section we show that in fact this may be reduced to $O(q^4 \log q)$.

The key to the savings is the recognition that the "standard" basis for an induced representation proves to be computationally useful in this case. Essentially, the computation may be reduced to a computation of Fourier transforms on the subgroup $T$ (of diagonal matrices) where abelian FFT methods may be used.

In actuality, what we consider is the computation of all Fourier transforms $\hat{f}(\rho_\psi)$, for $\psi \in \hat{K}^\times$. These are reducible only when $\psi = 1$ or $\psi = \mathrm{sgn}$. In both of these cases, $\rho_\psi$ is in fact multiplicity-free, and the change of basis to bring $\hat{f}(\rho)$ into the appropriate block diagonal form requires at most $2(q+1)^3$ operations (two matrix multiplications). This does not change the order of the result.

Using the notation of Section 2, recall that the principal series representations are the induced representations

$$\rho_\psi : \mathrm{SL}_2(q) \to \mathrm{GL}(\mathrm{Ind}(V_\psi)),$$

where $\mathrm{Ind}(V_\psi)$ is the vector space of functions $f : \mathrm{SL}_2(q) \to \mathbf{C}$ satisfying $f(bs) = \psi(b)f(s)$ for all $b \in B$ and $s \in \mathrm{SL}_2(q)$. Furthermore, $\mathrm{SL}_2(q)$ acts on this space by right translation,

$$(\rho_\psi(s)f)(s') = f(s's).$$

To obtain a matrix realization of this representation, a choice of basis must be made for $\mathrm{Ind}(V_\psi)$. By (2.1), any function in $\mathrm{Ind}(V_\psi)$ is determined by its values on a set of coset representatives for $B\backslash \mathrm{SL}_2(q)$. Fix the coset representatives

$$\ldots,\ s_u = \begin{pmatrix} 0 & 1 \\ -1 & -u \end{pmatrix},\ \ldots,\ s_\infty = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

where $u$ varies over $\mathbf{F}_q$. The notation comes from the natural correspondence between $B\backslash \mathrm{SL}_2(q)$ and the projective line over $\mathbf{F}_q$. Thus, let $e_u$, for $u \in$

$\mathbf{F}_q \cup \{\infty\}$, denote the corresponding element of $\mathrm{Ind}(V_\psi)$ defined by $e_u(s_v) = \delta_u(v)$, where

$$\delta_u(v) = \begin{cases} 1 & \text{if } u = v, \\ 0 & \text{otherwise.} \end{cases}$$

We give the basis the order

$$e_{\alpha^2}, e_{\alpha^4}, \ldots e_{\alpha^{q-3}}, e_\alpha, e_{\alpha^3}, \ldots, e_0, e_\infty,$$

where $\alpha$ is a fixed generator of $K^\times$ (in particular, $\alpha$ is not a square in $K^\times$).

Consider first the action of $U$ on the $\{e_u\}$. In general, to simplify the notation, we will sometimes write $se_u$ instead of $\rho_\psi(s)e_u$, where $s \in \mathrm{SL}_2(q)$. A straightforward matrix computation using (2.1) shows that

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} e_\infty = e_\infty$$

and

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} e_u = e_{u-a}$$

for $u \neq \infty$.

At this point it is important to note the following fact:

**Lemma 3.2.** *With respect to the ordered basis $\{e_u\}$ for $\mathrm{Ind}(V_\psi)$, the matrices $\rho_\psi(a)$ for $a \in U$ are of the form*

$$\begin{pmatrix} & & & 0 \\ & A(u) & & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

*where $A(u)$ is a $q \times q$ permutation matrix, that is, a matrix having one 1 in each row and column, and 0's everywhere else. Furthermore, as $u$ varies over $K$, the entries in $A(u)$ occur in distinct positions, that is, $\sum_{u\in K} A(u)$ is the $q \times q$ matrix with 1's everywhere. Lastly, the matrices $\rho_\psi(u)$ are independent of $\psi$.*

Now consider the action of $T$ on this basis. Once more, a straightforward matrix computation shows that

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} e_\infty = \psi(\alpha) e_\infty$$

and

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} e_u = \psi(\alpha^{-1}) e_{\alpha^2 u}$$

for $u \neq \infty$.

**Lemma 3.3.** *With respect to the ordered basis $\{e_u\}$ for $\mathrm{Ind}(V_\psi)$, the matrices $\rho_\psi(\alpha^j)$ are of the form*

$$\psi(\alpha^{-j}) \begin{pmatrix} C\left(\frac{q-1}{2}\right)^j & 0 & 0 & 0 \\ 0 & C\left(\frac{q-1}{2}\right)^j & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \psi(\alpha^{2j}) \end{pmatrix},$$

*where, for $n$ a positive integer, $C(n)$ is the $n \times n$ cyclic matrix*

$$C(n) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

Lastly, one can show that $we_\infty = \psi(-1)e_0$, $we_0 = e_\infty$ and $we_u = \psi(u)e_{-u^{-1}}$ for $u \neq 0, \infty$.

To summarize, we have obtained explicit realizations of the principal series representations for the group elements $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & -1 \\ 0 & 0 \end{smallmatrix}\right)$, $w$ and $w^{-1}$.

**Theorem 3.4.** *For $f \in L^2(\mathrm{SL}_2(q))$, all Fourier transforms $\hat{f}(\rho)$ at a complete set of principal series representations $\rho$ of $\mathrm{SL}_2(q)$ can be computed in at most*

$$8q^4 \log q + 2q^4 + q^3 - q^2 = O(q^4 \log q)$$

*operations.*

*Proof.* The Bruhat decomposition for $\mathrm{SL}_2(q)$ gives a decomposition of the computation of $\hat{f}(\rho_\psi)$ as

$$\begin{aligned} \hat{f}(\rho_\psi) &= \sum_{s\in \mathrm{SL}_2(q)} f(s)\rho_\psi(s) \\ &= \sum_{u\in K} \sum_{\beta \in K^\times} \sum_{u'\in K} f(u'\beta w u)\rho_\psi(u')\rho_\psi(\beta w u) \\ &= \sum_{u\in K} \sum_{\beta \in K^\times} \sum_{u'\in K} \big(f_{\beta,u}(u')\rho_\psi(u')\big)\rho_\psi(\beta w u), \end{aligned}$$

where $f_{\beta,u} \in L(K)$ is defined by

$$f_{\beta,u}(u') = f(u'\beta w u).$$

Here we make the identifications

$$u \in K \leftrightarrow \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}$$

$$\beta \in K^\times \leftrightarrow \begin{pmatrix} \beta & 0 \\ 0 & \beta^{-1} \end{pmatrix}.$$

By Lemma 3.2 the $\rho_\psi(u)$ are permutation matrices whose form is independent of $\psi$. We obtain the following consequence:

**Lemma 3.5.** *Let all notation be as above. To compute the matrices $\{\hat{f}_{\beta,u'}(\rho_\psi{\downarrow}U)\}$ for all $\beta \in K^\times$, $u' \in K$ and all $\psi \in \hat{K}^\times$, at most $q^3 - q^2$ additions and no multiplications are needed.*

*Proof.* Using the notation of Lemma 3.2 we write

$$\hat{f}_{\beta,u}(\rho_\psi{\downarrow}U) = \sum_{u'\in K} f_{\beta,u}(u') \begin{pmatrix} & A(u') & & \begin{matrix}0\\ \vdots\\ 0\end{matrix} \\ 0 & \cdots & 0 & 1 \end{pmatrix} = \begin{pmatrix} & \hat{f}_{\beta,u}(A) & & \begin{matrix}0\\ \vdots\\ 0\end{matrix} \\ 0 & \cdots & 0 & S_{\beta,u} \end{pmatrix},$$

where $S_{\beta,u} = \sum_{u'\in K} f_{\beta,u}(u')$. By Lemma 3.2, the entries of the upper block $\hat{f}_{\beta,u}(A)$ are just the values $f_{\beta,u}(u')$, each value appearing exactly once in each row and column. Hence the only computations done are the $q$ additions to form $S_{\beta,u}$. Repeated for each $\beta \in K^\times$ and $u \in K$, this gives at most $q \cdot q \cdot (q-1)$ additions. □

We return to the computation of $\hat{f}(\rho_\psi)$. Since, by Lemma 3.2, the restricted transforms

$$\hat{f}_{\beta,u}(\rho_\psi{\downarrow}U)$$

(in the basis of choice) are independent of $\psi$, we denote this matrix as $\Theta_{j,u}$. Then we write

$$\hat{f}(\rho_\psi) = \sum_{u\in K}\sum_{j=0}^{q-1} \Theta_{j,u}\rho_\psi(\alpha^{-j})\rho_\psi(w)\rho_\psi(u).$$

Thus, we now consider the computation of the inner sum

$$\sum_{j=0}^{q-1} \Theta_{j,u}\rho_\psi(\alpha^{-j})$$

for all $\psi \in \hat{K}^\times$. By Lemma 3.3 we rewrite this as

$$\sum_{j=0}^{q-1} \psi(\alpha^{-j})\Theta_{j,u} \begin{pmatrix} C\left(\frac{q-1}{2}\right)^j & 0 & 0 & 0 \\ 0 & C\left(\frac{q-1}{2}\right)^j & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \psi(\alpha^{2j}) \end{pmatrix}. \tag{3.1}$$

No multiplications are needed to compute the matrix product $M \cdot C(n)$ for any $r \times n$ matrix $M$. Thus, at most $q^2$ multiplications are needed to rewrite (3.1) as

$$\sum_{j=0}^{q-1} \psi(\alpha^j) \begin{pmatrix} A_{j,u}C\left(\frac{q-1}{2}\right)^j & B_{j,u}C\left(\frac{q-1}{2}\right)^j & * & * \\ C_{j,u}C\left(\frac{q-1}{2}\right)^j & D_{j,u}C\left(\frac{q-1}{2}\right)^j & * & * \\ * & * & * & * \\ * & * & * & * \end{pmatrix}$$

where the asterisks denote arbitrary complex matrices of the appropriate dimension. Note that the submatrix

$$\Gamma_{j,u} = \begin{pmatrix} A_{j,u}C\left(\frac{q-1}{2}\right)^j & B_{j,u}C\left(\frac{q-1}{2}\right)^j \\ C_{j,u}C\left(\frac{q-1}{2}\right)^j & D_{j,u}C\left(\frac{q-1}{2}\right)^j \end{pmatrix}$$

is again independent of $\psi$.

If we form the matrix $\hat{\Gamma}_u(\psi)$ with entries $\hat{\Gamma}_u^{i,k}(\psi)$, where $\Gamma_u^{i,k}(\alpha^j)$ is the $(i,k)$-entry of $\Gamma_{j,u}$, we can now write

$$\hat{f}(\rho_\psi) = \sum_{u\in K} \begin{pmatrix} \hat{\Gamma}_u(\psi) & * \\ * & * \end{pmatrix} \rho_\psi(w)\rho_\psi(u).$$

For each $i$, $k$ and $u$, it takes at most $8q\log q$ operations to compute $\hat{\Gamma}_u^{i,k}(\psi)$ for all $\psi \in \hat{K}^\times$, using the abelian FFT. Thus, a total of at most $8q^4\log q$ operations are needed.

Finally, since $\rho_\psi$ has only one nonzero entry in each row and column, the matrix product $\hat{\Gamma}_u(\psi)$ takes at most $q^2$ operations to compute. Thus, to do this for all $\psi$ and $u$, at most another $q^4 - q^3$ operations are needed. Finally, since the matrices $\rho_\psi(u)$ are permutation matrices, no multiplications are needed in the end. Collecting all terms, we find that at most

$$\begin{aligned} 8q^4\log q + q^4 + q^4 - q^3 + q^3 + q^3 - q^2 &= 8q^4\log q + 2q^4 + q^3 - q^2 \\ &= O(q^4\log q) \end{aligned}$$

operations are needed to compute all Fourier transforms at the principal series representations. This concludes the proof of Theorem 3.4. □

### Fourier Analysis at the Discrete Series

We now turn to the computation at the discrete series representations. As discussed in Section 2, we follow [Silberger 1969] in constructing these representations as restrictions of discrete series representations of $\mathrm{GL}_2(K)$, which can be constructed explicitly [Piatetski-Shapiro 1983].

More precisely (see Section 2), there is a correspondence between discrete series representations for $\mathrm{GL}_2(K)$ and nondecomposable characters of the unique quadratic extension $L$ of $K$. We recall that $C \subset L^\times$ denotes the set of elements of norm 1 in this extension. If $\pi$ is any nondecomposable character of $L^\times$, we denote by $\rho_\pi$ the corresponding discrete series representation on $V_\pi$, the vector space of complex-valued functions on $K^\times$. The action of $\mathrm{SL}_2(q)$ by $\rho_\pi$ is as indicated in (2.2)–(2.4).

There are many "natural" choices of basis for $V_\pi$. From a computational point of view, and particularly from the point of view of investigating expander properties, an especially simple choice of basis is that of the delta functions $e_x$ for $V_\pi$, defined by $e_x(y) = \delta_{x,y}$. We assume some fixed ordering of the $e_x$, say $\dots, e_{\alpha^j}, \dots$, where $\alpha$ generates $K^\times$. Then, using (2.2), we obtain

$$\rho_\pi\left(\left(\begin{smallmatrix}1 & b\\ 0 & 1\end{smallmatrix}\right)\right)(e_x) = \chi(bx)e_x.$$

This gives a matrix realization as

$$\rho_\pi\left(\left(\begin{smallmatrix}1 & b\\ 0 & 1\end{smallmatrix}\right)\right) = \begin{pmatrix}\ddots & & 0\\ & \chi_x(b) & \\ 0 & & \ddots\end{pmatrix},$$

where $\chi_x(b) = \chi(xb)$ for all $x \in K^\times$ and $\left(\begin{smallmatrix}1 & b\\ 0 & 1\end{smallmatrix}\right) \in U$.

In the same way, (2.3) yields

$$\rho_\pi\left(\left(\begin{smallmatrix}\alpha & 0\\ 0 & \alpha^{-1}\end{smallmatrix}\right)\right)e_x = \pi(\alpha^{-1})e_{\alpha^{-2}x},$$

and (2.4) gives

$$\begin{aligned}\rho_\pi\left(\left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)\right)e_x(y) &= \sum_z \pi(z)^{-1}j_\pi(zy)e_x(z)\\ &= \pi(x)^{-1}j_\pi(xy)e_y.\end{aligned}$$

Thus, $\rho_\pi(w)$ is a circulant matrix, with $(x,y)$-entry equal to $j_\pi(xy)$ times a diagonal matrix:

$$\rho_\pi\left(\left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)\right) = (j_\pi(xy))_{x,y}\begin{pmatrix}\ddots & & 0\\ & \pi(y)^{-1} & \\ 0 & & \ddots\end{pmatrix}.$$

**Theorem 3.6.** *For any $f \in L^2(\mathrm{SL}_2(K))$, the Fourier transforms $\hat{f}(\rho)$ at all discrete series representations $\rho$ of $\mathrm{SL}_2(K)$ can be computed in*

$$8\log q(q^4+q^3+2q^2)+3q^4+q^2(q-1)^2 = O(q^4\log q)$$

*operations.*

*Proof.* As before, we identify $u \in K$ with $\left(\begin{smallmatrix}1 & u\\ 0 & 1\end{smallmatrix}\right)$ and $\beta \in K^\times$ with $\left(\begin{smallmatrix}\beta & 0\\ 0 & \beta^{-1}\end{smallmatrix}\right)$. As in the previous subsection, we fix a generator $\alpha$ of $K^\times$ and a nontrivial additive character $\chi$ of $K^+$, and use the Bruhat decomposition to write

$$\begin{aligned}\hat{f}(\rho) &= \sum_{g\in\mathrm{SL}_2(q)} f(g)\rho(g)\\ &= \sum_{u\in U}\sum_{t\in T}\sum_{u'\in U} f(u'twu)\rho(u'twu) + \sum_{b\in B} f(b)\rho(b).\end{aligned} \tag{3.2}$$

Thus the sum naturally breaks into two parts, one over $B$ and one over $BwU$. Consider first the sum over $BwU$.

For any $u \in U$ and $t \in T$, let $f_{u,t}$ be the function in $L^2(U)$ defined by $f_{u,t}(u') = f(u'twu)$. Then the sum over $BwU$ equals $\sum_{u\in U}\sum_{t\in T} Q(u,t)$, where

$$\begin{aligned}Q(u,t) &= \Big(\sum_{u'\in U} f_{u,t}(u')\rho(u')\Big)\rho(t)\rho(w)\rho(u)\\ &= \sum_{u'\in U}\begin{pmatrix}\ddots & & 0\\ & f_{u,t}(u')\chi_x(u') & \\ 0 & & \ddots\end{pmatrix}\rho(t)\rho(w)\rho(u)\\ &= \begin{pmatrix}\ddots & & 0\\ & \hat{f}_{u,t}(\chi_x) & \\ 0 & & \ddots\end{pmatrix}\rho(t)\rho(w)\rho(u).\end{aligned}$$

For any fixed $u \in U$ and $t \in T$, the Fourier transforms $\hat{f}_{u,t}(\chi_x)$ for all $x \in K^\times$ can be computed in at most $8q\log q$ operations using the abelian FFT. Thus, at most $8q^3\log q$ operations are needed to compute all the inner diagonal matrices, which are independent of the discrete series representation $\rho$.

Let $F(u,t)$ denote the diagonal matrix

$$\begin{pmatrix}\ddots & & 0\\ & \hat{f}_{u,t}(\chi_x) & \\ 0 & & \ddots\end{pmatrix}.$$

Proceeding directly, note that the matrices $\rho(t)$ are generalized permutation matrices, that is, each row and column contains exactly one nonzero entry, so that for any fixed $\rho$, the matrix products $F(u,t)\rho(t)$ can be computed in $q-1$ operations for any $t \in T$, adding up to $q(q-1)^2$ operations for all $t \in T$ and $u \in U$. Thus, the inner sums

$$M_u(\rho) = \sum_{t \in T} F(u,t)\rho(t)$$

for all discrete series representations $\rho$ can be computed in at most $q^2(q-1)^2 + 8q^3 \log q$ operations.

Next we must compute $M'_u(\rho) = M_u(\rho)\rho(w)$. Because $\rho(w)$ is a circulant matrix, any $(q-1) \times (q-1)$ matrix can be multiplied by $\rho(w)$ in at most $8q^2 \log q$ operations (again by abelian FFT methods). Letting $\rho$ vary over all discrete series representations and $u$ vary over $U$, we conclude that all the products $M'_u(\rho)$ can be computed in at most $8q^4 \log q$ operations.

Finally, we must sum $\sum_{u \in U} M'_u(\rho)\rho(u)$. Again, the matrices $\rho(u)$ are diagonal. Thus, for any fixed $\rho$, the sum requires at most $q^3$ operations. For all $\rho$, we need at most $q^4$ operations.

We now turn to the second term in (3.2),

$$\sum_{b \in B} f(b)\rho(b).$$

Since $B$ is a metabelian group, all Fourier transforms of any function in $L^2(B)$ may be computed in at most $16q^2 \log q$ operations [Baum et al. 1991, Theorem 4]. Any necessary change of basis requires at most $2q^3$ operations. Thus, over all $\rho$, at most $2q^4$ additional operations need be performed. Collecting terms yields at most

$$8q^3 \log q + 8q^4 \log q + q^2(q-1)^2 + q^4 + 16q^2 \log q + 2q^4$$

operations, and we obtain the upper bound in the statement of the theorem. $\square$

Adding together the bounds in Theorems 3.4 and 3.6 and using simple inequalities to eliminate terms of lower order in $q$, we get the following result:

**Theorem 3.7.** *If $q$ is a power of an odd prime, the number $T(q)$ of operations needed to compute a Fourier transform of a function $f \in L^2(\mathrm{SL}(q))$ is at most $25q^4 \log q$.*

## 4. IMPLEMENTING THE COMPUTATION

As explained in Section 1, explicit matrix representations can be used to investigate the spectra of Cayley graphs. In this section we detail two types of experiments that we carried out. We gratefully acknowledge the help and suggestions of A. Lubotzky and P. Sarnak regarding these investigations.

**Asymptotics of spectra of families of Cayley graphs on $\mathrm{SL}_2(p)$.** In these experiments we consider Cayley graphs on fixed sets of generators (whose elements may vary with $p$, but whose size does not) and then consider the spectra as $p$ gets large. We compute spectra for the following three sets of generators.

$$\begin{aligned}
\mathcal{G}_1 &= \left\{ \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right), w, w^{-1} \right\},\\
\mathcal{G}_2 &= \left\{ \left(\begin{smallmatrix}1&(p+1)/2\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&(p-1)/2\\0&1\end{smallmatrix}\right), w, w^{-1} \right\},\\
\mathcal{G}_3 &= \left\{ w\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\-1&1\end{smallmatrix}\right)w^{-1}, w, w^{-1} \right\}.
\end{aligned}$$

Of interest is the behavior of the second-largest eigenvalue, multiplicities of the eigenvalues and the range of eigenvalues.

**Expanding properties of randomly chosen pairs of generators.** Here the goal is to gain insight into the "expanding behavior" (that is, the second-largest eigenvalue) of a generic pair of generators of $\mathrm{SL}_2(p)$. It is known [Kantor and Lubotzky 1990] that almost every pair of elements generates $\mathrm{SL}_2(p)$, but little is known of the expanding behavior for different pairs of generators. The idea, then, is to compute Fourier transforms for all pairs of generators for some small range of primes, and to compare second-largest eigenvalues for the associated Cayley graphs.

In practice, computing full spectra for all generating pairs is too large a computational task (see the discussion following Lemma 4.7). Thus, we limit ourselves to computing spectra for some sizable set of random generating pairs over a larger range of primes.

Both experiments have associated implementation issues. We treat general issues first and then explain the two computations separately. Note that, while we are only interested in the case of $\mathrm{SL}_2(K)$ for $K$ a prime field, it is straightforward to extend these methods to arbitrary finite fields, using algorithms such as those described in [Lenstra and Lenstra 1990].

In this section, $p$ will denote an odd prime and $K = \mathbf{F}_p$ the field with $p$ elements.

### Working in the Base Field

Section 2 shows that the representations of $\mathrm{SL}_2(p)$ essentially occur as $p$ representations of size $p$. As a matter of practicality, then, we must limit ourselves to primes on the order of 500 if the eigenvalues are to be computed at the complete set of irreducible representations. For such small primes efficiency is not of the essence in working with the underlying fields.

The two basic operations necessary for our purposes are inversion in $\mathbf{F}_p^\times$ and modular exponentiation in $\mathbf{F}_p$. The former is carried out with the help of the Euclidean algorithm, and the latter by the method of repeated squares. For completeness, we record here the easily derived complexity of these well-known algorithms.

**Proposition 4.1.** *Let $1 \le a, b \le p-1$ and $n \in \mathbf{Z}$. The Euclidean algorithm computes $\gcd(a,b)$ in time*

$$O(\log^3 \max(a,b)).$$

*The method of repeated squares computes $a^n \in \mathbf{F}_p$ in time $O(\log n \log^2 p)$.*

Given these basic operations, one may efficiently calculate Legendre symbols and find generators of the cyclic group $\mathbf{F}_p^\times$. In particular, since we are working with relatively small primes, we may afford ourselves the luxury of finding the smallest generator of the field. Alternatively, a randomized algorithm may be used.

### Working in the Quadratic Extension

As described in Section 2, the discrete series representations require calculations in the quadratic extension $L$ of the base field. To prepare for computations here, the first task is to find a generator of the cyclic group $L^\times$. For concreteness, we construct $L = K(\sqrt{\varepsilon})$, where $\varepsilon$ is the smallest generator of $K^\times$ (as a positive integer). With $L$ realized as a two-dimensional vector space over $K$ in the canonical way, exponentiation in $L$ is again carried out by the method of repeated squares, with the same complexity estimate as for the base field.

For $j \in K^\times$, let $C(j)$ be the circle of radius $j$ in $L$, that is, the set of elements of $L$ of norm $j$. The unit circle $C(1)$ is easily constructed, as it is parametrized by $K \cup \{\infty\}$. The elements $z_t = (x_t, y_t) \in C(1)$ are given by $z_\infty = (-1, 0)$ and

$$x_t = (1+\varepsilon t^2)(1-\varepsilon t^2)^{-1},$$
$$y_t = 2t\,(1-\varepsilon t^2)^{-1}$$

for $t \in K$. The circle $C(\varepsilon)$ of radius $\varepsilon$ is then obtained by simply multiplying by the vector $(0,1)$, and a generator of $L^\times$ may be obtained by a randomized algorithm on this circle. Furthermore, since $L^\times$ is of order $p^2-1$ and each $z \in C(\varepsilon)$ satisfies $N(z) = \varepsilon$, checking whether or not $z$ is a generator requires calculating at most the first $p-1$ powers of $z$ in $L^\times$. We thus have the following estimate:

**Proposition 4.2.** *A generator for $L^\times$ can be found in randomized time $O(p)$. More precisely, a generator can be obtained with probability $1-\alpha^n$ in time $O(np)$, where $\alpha$ is the distribution of nongenerators on the circle $C(\varepsilon) \subset L^\times$.*

However, note that since $C(\varepsilon)$ is of order $p+1$, the efficiency of this procedure is not at all crucial for our purposes, and we may in fact allow ourselves to find the generator of smallest "Euclidean norm"

$$|z_t|^2 = |\,(x_t, y_t)|^2 = x_t^2 + y_t^2.$$

In any case, it is desirable to be able to compute a canonical choice of generator, since the multiplicative characters used in the representations depend on this choice.

### Checking the Computations

In any implementation, it is desirable to use representation-theoretic identities to check the correctness of the representations obtained. Simple identities that can be used for this purpose are $\rho(w)\rho(w^{-1}) = I$, $\rho^p\big(\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)\big) = \rho^p\big(\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)\big) = I$ and

$$\rho\big(\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)\big)\rho\big(\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)\big) = I.$$

The most important check, however, is a direct consequence of Schur's lemma:

**Proposition 4.3.** *The matrix identity*

$$\sum_{t\in T}\sum_{u\in U}\sum_{u'\in U}\rho(t)\rho(u)\rho(w)\rho(u') + \sum_{t\in T}\sum_{u\in U}\rho(t)\rho(u) = 0$$

*holds, where $\rho(t)$ may be calculated in terms of $\mathcal{G}_1$ as*

$$\rho(w)\rho^{t^{-1}}\big(\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)\big)\rho(w)\rho^{t}\big(\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)\big)\rho(w^{-1})\rho^{t^{-1}}\big(\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)\big)$$

*and in terms of* $\mathcal{G}_2$ *as*

$$\rho(w)\rho^{2t^{-1}}\left(\left(\begin{smallmatrix}1 & (p-1)/2\\0 & 1\end{smallmatrix}\right)\right) \\ \times \rho(w)\rho^{2t}\left(\left(\begin{smallmatrix}1 & (p-1)/2\\0 & 1\end{smallmatrix}\right)\right)\rho(w^{-1})\rho^{2t^{-1}}\left(\left(\begin{smallmatrix}1 & (p-1)/2\\0 & 1\end{smallmatrix}\right)\right).$$

*Furthermore, the complexity of this calculation is of order* $O(p^{3+\alpha})$, *where* $\alpha$ *is the exponent of the complexity bound for matrix multiplication.*

*Proof.* The identity follows directly from the Bruhat decomposition

$$\mathrm{SL}_2(p) = TUwU \amalg TU,$$

together with Schur's lemma, which implies that

$$\sum_{g\in \mathrm{SL}_2(p)} \rho(g) = 0$$

whenever $\rho$ is irreducible and unitary. To compute the torus $T$ in terms of $\mathcal{G}_i$, simply observe the matrix identity

$$\left(\begin{smallmatrix}\alpha & 0\\0 & \alpha^{-1}\end{smallmatrix}\right) = \left(\begin{smallmatrix}1 & -\alpha^{-1}\\0 & 1\end{smallmatrix}\right)w\left(\begin{smallmatrix}1 & -\alpha\\0 & 1\end{smallmatrix}\right)w^{-1}\left(\begin{smallmatrix}1 & -\alpha^{-1}\\0 & 1\end{smallmatrix}\right). \qquad \square$$

## Computing Random Generating Pairs

We now present an efficient method for deciding whether two given elements $x, y \in \mathrm{SL}_2(p)$ form a *generating pair*, that is, whether $\{x, y\}$ generates $\mathrm{SL}_2(p)$. The main idea is to use the classification of subgroups of $\mathrm{PSL}_2(p) = \mathrm{SL}_2(p)/\{\pm I\}$, and the fact that (for $p > 3$) $H \le \mathrm{SL}_2(p)$ is a proper subgroup if and only if $\pi(H) \le \mathrm{PSL}_2(p)$ is a proper subgroup, where $\pi : \mathrm{SL}_2(p) \to \mathrm{PSL}_2(p)$ is the usual projection map.

Given $x, y \in \mathrm{SL}_2(p)$, then, the strategy is to test whether their images $\pi(x)$ and $\pi(y)$ generate one of the possible types of proper subgroups; as we will see, doing this is relatively straightforward. If $\pi(x)$ and $\pi(y)$ do not generate a subgroup, they generate all of $\mathrm{PSL}_2(p)$, so $x$ and $y$ comprise a generating pair.

The following theorem is found in [Suzuki 1982] and usually attributed to [Dickson 1958].

**Theorem 4.4.** *The following are all possible proper subgroups of* $\mathrm{PSL}_2(p)$:

(a) *Abelian subgroups.*
(b) *Dihedral groups of order* $2n$, *where* $n$ *divides* $\frac{1}{2}(p+1)$ *or* $\frac{1}{2}(p-1)$.
(c) *The alternating group* $A_4$.
(d) *Noncommutative subgroups of the image of the upper triangular subgroup, and its conjugates.*
(e) *The symmetric group* $S_4$, *if* $p^2 = 1 \pmod{16}$.
(f) *The alternating group* $A_5$, *if either* $p = 5$—*in which case* $\mathrm{PSL}_2(5) = A_5$—*or* $p^2 = 1 \pmod 5$.

Our algorithm tests for cases (a)–(f) in that order. The specific tests are given below. In this analysis the symbol $=$ means equality in $\mathrm{PSL}_2(p)$ (that is, equality between images under $\pi$).

**Test for (a).** Check if $xy = yx$.

**Test for (b).** If $x$ and $y$ generate a dihedral group of order $2n$, there are only two possibilities: either $x^2 = y^n = 1$ and $xyx = y^{-1}$, or $x^2 = y^2 = 1$ and $(xy)^n = 1$. (Also, the roles of $x$ and $y$ can be reversed.) Since every element has finite order, the test can be reduced to checking if

1. $x^2 = 1$ and $xyx = y^{-1}$, or
2. $x^2 = y^2 = 1$.

Note that inversion in $\mathrm{SL}_2(p)$ is very quick: the inverse of $\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ is simply $\left(\begin{smallmatrix}\delta & -\beta\\ -\gamma & \alpha\end{smallmatrix}\right)$. Therefore, testing for the conditions above is very efficient.

**Test for (c).** Testing whether $x$ and $y$ generate $A_4$ is much like the previous case, in the sense that it is a question of trying out all possible two-generator presentations for the group. Here the presentation must be either $x^2 = y^3 = (xy)^3 = 1$ or $x^3 = y^3 = (xy)^2 = 1$. The first case corresponds to the generators (12)(34) and (123), and the second to (123) and (234).

**Test for (d).** A subgroup of the upper triangular group fixes the point at infinity of the projective line $P^1(\mathbf{F}_p) = \mathbf{F}_p \cup \{\infty\}$, where the action of $\mathrm{PSL}_2(p)$ on $P^1(\mathbf{F}_p)$ is by fractional linear transformations:

$$\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}(\omega) = \frac{\alpha\omega+\beta}{\gamma\omega+\delta}.$$

Any conjugate of the upper triangular group fixes some other point of $P^1(\mathbf{F}_p)$. Consequently, this test can be reduced to checking whether $x$ and $y$ have any fixed points in common.

Clearly, $\infty$ is a fixed point of $x$ and $y$ if and only if both matrices are upper triangular. Checking for other points requires a bit more work. A point $\omega \ne \infty$ is fixed by $\left(\begin{smallmatrix}\alpha & \beta\\ \gamma & \delta\end{smallmatrix}\right)$ if and only if

$$\gamma\omega^2 + \omega(\delta-\alpha) - \beta = 0.$$

Therefore the check amounts to determining if two quadratic polynomials over $\mathbf{F}_p$ share any roots over $\mathbf{F}_p$.

**Tests for (e) and (f).** These tests have to be performed only when the corresponding divisibility conditions (see Theorem 4.4) are satisfied. One way to go about them is to list all possible two-generator presentations of $S_4$ or $A_5$ and to check for each one of them explicitly.

We chose a different approach, since our goal was to compute (for each prime) the spectrum for each of 25 pairs of generators chosen at random (Figure 8). The results of [Kantor and Lubotzky 1990] imply that the occurrence of $S_4$ or $A_5$ as the subgroup generated by two random elements of $\mathrm{SL}_2(p)$ is rare. For this reason, our implementation altogether ignored cases (e) and (f) of Theorem 4.4, and relied instead on the fact that, if $x$ and $y$ generate a proper subgroup, the eigenvalue 1 will appear in the spectrum with multiplicity equal to the index of the subgroup in $\mathrm{SL}_2(p)$, and this index will be much greater than one. Therefore, it is more efficient to first compute the spectrum whenever a pair $\{x, y\}$ fails tests (a)–(d), and then weed out those pairs that do not generate the whole group, as detected by the high multiplicity of the eigenvalue 1.

### Representations at a Random Generating Pair

Using the Bruhat decomposition

$$\mathrm{SL}_2(p) = TUwU \amalg TU,$$

we may distinguish the upper-triangular component, parametrized as $g_{\alpha,u} \in TU$, for $\alpha \in \mathbf{F}_p^\times$ and $u \in \mathbf{F}_p$, and its complement, parametrized as $g_{\alpha,u,v} \in TUwU$, for $\alpha \in \mathbf{F}_p^\times$, $u \in \mathbf{F}_p$ and $v \in \mathbf{F}_p$.

Thus, to generate a random element of $\mathrm{SL}_2(p)$, we may first determine which component of the group it should belong to, by generating a random integer $1 \le r \le p(p^2-1)$ and checking whether $r \le p(p-1)$ (in which case the element will be upper triangular) or not. In either case, we need a random $\alpha \in \mathbf{F}_p^\times$ and a random $u \in \mathbf{F}_p$. In the more probable event that we are constructing an element of $TUwU$, we also need a random $v \in \mathbf{F}_p$. We then form

$$g_{\alpha,u} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \in TU$$

or

$$g_{\alpha,u,v} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & v \\ 0 & 1 \end{pmatrix} \in TUwU,$$

as appropriate.

Rather than multiply the matrices for the representations evaluated at each of the factors in the above expressions, it is more efficient to calculate the representations directly.

For the principal series representations, we find

$$\begin{aligned} \rho_\psi(g_{\alpha,u})e_\infty &= \psi(\alpha)e_\infty, \\ \rho_\psi(g_{\alpha,u})e_x &= \psi(\alpha^{-1})e_{\alpha^2(x-u)}, \\ \rho_\psi(g_{\alpha,u}^{-1})e_\infty &= \psi(\alpha^{-1})e_\infty, \\ \rho_\psi(g_{\alpha,u}^{-1})e_x &= \psi(\alpha)e_{\alpha^{-2}x+u}, \end{aligned}$$

and

$$\begin{aligned} \rho_\psi(g_{\alpha,u,v})e_\infty &= \psi(-\alpha^{-1})e_{-\alpha^2 u}, \\ \rho_\psi(g_{\alpha,u,v})e_x &= \begin{cases} \psi(\alpha)e_\infty & \text{if } x-v=0, \\ \psi(\alpha^{-1}(x-v))e_{-\alpha^2((x-v)^{-1}+u)} & \\ & \text{if } x-v\neq 0, \end{cases} \\ \rho_\psi(g_{\alpha,u,v}^{-1})e_\infty &= \psi(\alpha^{-1})e_v, \\ \rho_\psi(g_{\alpha,u,v}^{-1})e_x &= \begin{cases} \psi(-\alpha)e_\infty & \text{if } \alpha^{-2}x+u=0, \\ \psi(-\alpha(\alpha^{-2}x+u))e_{v-(\alpha^{-2}x+u)^{-1}} & \\ & \text{if } \alpha^{-2}x+u\neq 0. \end{cases} \end{aligned}$$

Similarly, we find for the discrete series

$$\begin{aligned} \rho_\pi(g_{\alpha,u})e_x &= \pi(\alpha^{-1})\chi(ux)e_{\alpha^{-2}x}, \\ \rho_\pi(g_{\alpha,u}^{-1})e_x &= \pi(\alpha)\chi(-u\alpha^2 x)e_{\alpha^2 x}, \end{aligned}$$

and

$$\begin{aligned} \rho_\pi(g_{\alpha,u,v})e_x(y) &= \pi(\alpha x)^{-1}\chi(vx+u\alpha^2 y)j_\pi(\alpha^2 xy), \\ \rho_\pi(g_{\alpha,u,v}^{-1})e_x(y) &= \pi(-\alpha x)^{-1}\chi(-vy-u\alpha^2 x)j_\pi(\alpha^2 xy). \end{aligned}$$

## 5. DISCUSSION OF NUMERICAL RESULTS

In this section we present the results of several numerical investigations of the spectra of Cayley graphs for $\mathrm{SL}_2(p)$. In particular, we present data on the spectra of Cayley graphs associated with generator sets $\mathcal{G}_1$ and $\mathcal{G}_2$, exhibiting the behavior of the second-largest eigenvalue of the principal series representations, and of the largest eigenvalue of the discrete series representations. As discussed in Section 1, these eigenvalues are related to the expansion coefficient of the graph. We also present figures exhibiting the full spectrum of these and other Cayley graphs. Finally, we discuss the expanding properties of randomly chosen generating pairs.

### Second-Largest Eigenvalues

As mentioned in Section 4, we work with the generating sets

$$\begin{aligned}\mathcal{G}_1 &= \{\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right), w, w^{-1}\},\\ \mathcal{G}_2 &= \{\left(\begin{smallmatrix}1&(p+1)/2\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&(p-1)/2\\0&1\end{smallmatrix}\right), w, w^{-1}\},\\ \mathcal{G}_3 &= \{w\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right), \left(\begin{smallmatrix}1&0\\-1&1\end{smallmatrix}\right)w^{-1}, w, w^{-1}\}.\end{aligned}$$

To obtain a sense of the structure of the associated Cayley graphs, notice that $w$ has order 4, while $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ has order $p$ and $w\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ has order 3. Thus, the Cayley graph of $\mathrm{SL}_2(\mathbf{F}_5)$ with respect to $\mathcal{G}_1$ has cycles of order 4, 5 and 6. A fragment of this graph is shown in Figure 1.

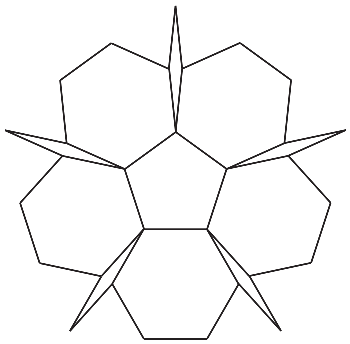

**FIGURE 1.** Fragment of Cayley graph for $\mathrm{SL}_2(\mathbf{F}_5)$ with generators $\mathcal{G}_1$.

If we project the 4-cycles onto lines, we obtain precisely the Cayley graph for $\mathrm{PSL}_2(\mathbf{F}_5)$ with respect to the projected set of generators, since $w$ is its own inverse in this quotient group. Unlike its covering graph, the convex hull of this graph, embedded in Euclidean three-space, can be seen as a regular polytope, as shown in Figure 2.

A fragment of the universal covering graph of $\mathrm{PSL}_2(\mathbf{F}_p)$ with generators $\mathcal{G}_3$ is shown in Figure 3.

Applying the theory presented in Sections 1, 2 and 3, we computed the spectra of these Cayley graphs by constructing the principal and discrete series representations and by computing the eigenvalues of the resulting matrices. More specifically, for a generating set $\mathcal{G} = \{g_1, g_2, g_1^{-1}, g_2^{-1}\}$, we computed the eigenvalues of the matrices

$$\hat{\delta}_{\mathcal{G}}(\rho) = \rho(g_1) + \rho(g_2) + \rho(g_1)^{-1} + \rho(g_2)^{-1}$$

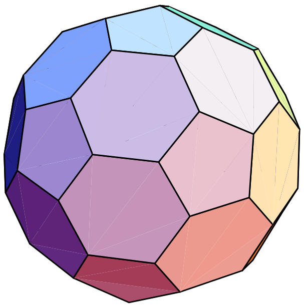

**FIGURE 2.** Cayley graph of $\mathrm{PSL}_2(\mathbf{F}_5)$ with respect to the generator set $\mathcal{G}_1$.

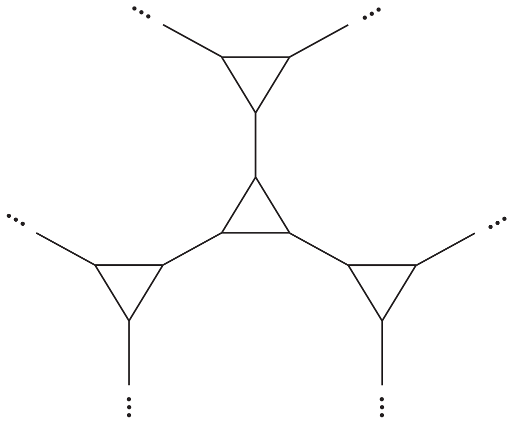

**FIGURE 3.** Fragment of covering Cayley graph for $\mathrm{PSL}_2$ with generators $\mathcal{G}_3$.

as $\rho$ varied over the complete set of discrete and principal series representations.

In Figure 4 we plot, as a function of the prime $p$, the second-largest eigenvalue among the principal series representations (recall that the largest eigenvalue is 4, coming from the identity character). We also plot the largest discrete series eigenvalue. The computations were carried out for all 93 primes between 5 and 500. It is notable that for primes larger than 100 the eigenvalues stabilize quickly to a value around 0.982, where the eigenvalues have been normalized by the degree of the graph.

Figure 5 shows the corresponding eigenvalues for the generating set $\mathcal{G}_2$. Here the eigenvalues give the appearance of stabilizing slightly more slowly, around a somewhat smaller value of approximately 0.972.

Finally, Figure 6 shows the second-largest eigenvalue overall for each of the generating sets $\mathcal{G}_1$ and $\mathcal{G}_2$. It is this eigenvalue that is related to the expansion coefficient through the isoperimetric inequalities referred to in the Introduction.

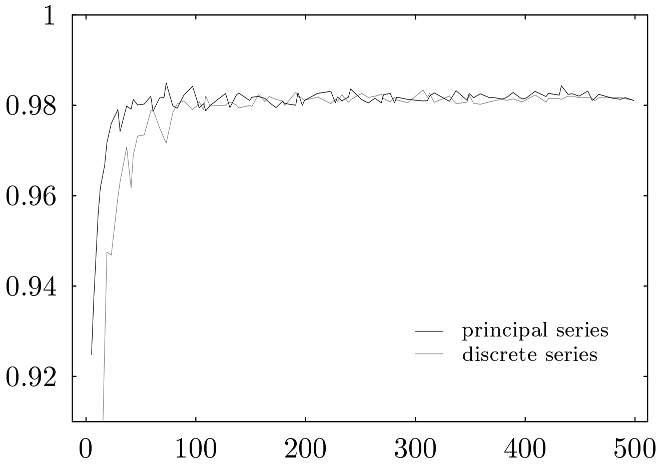


**FIGURE 4.** Principal and discrete series eigenvalues for generators $\mathcal{G}_1$.

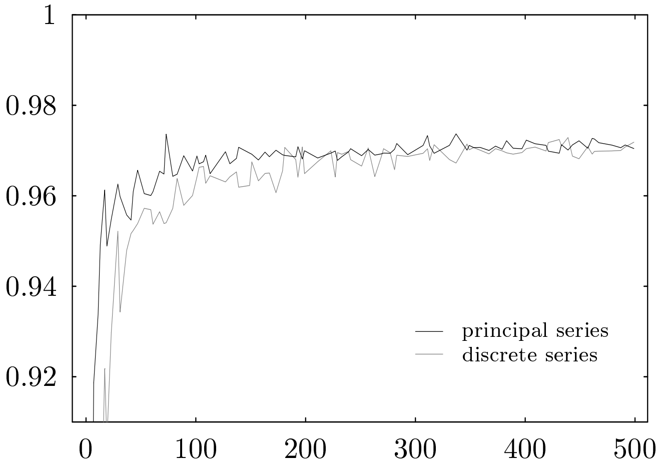


**FIGURE 5.** Principal and discrete series eigenvalues for generators $\mathcal{G}_2$.

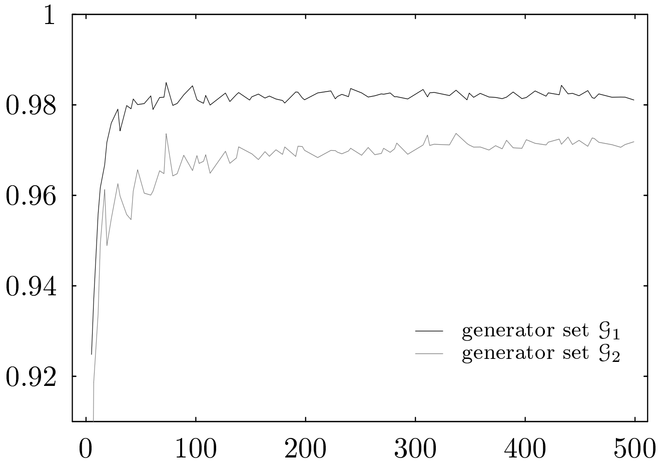


**FIGURE 6.** Second-highest eigenvalue for generators $\mathcal{G}_1$ and $\mathcal{G}_2$.

**The Full Spectrum**

The next series of figures displays the full spectrum for the generator sets $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$. The top left panel in Figure 7 shows the principal series spectrum for $\mathcal{G}_1$ for each of the 28 primes between 5 and 113. In fact, the computations were carried out for primes up to 251; however, at the resolution of these graphs, the spectrum becomes "continuous" outside of the exceptional neighborhood of zero that contains the isolated eigenvalues. In short, the spectra all resemble that for prime 113, the largest shown on this graph. Again the eigenvalues are normalized by the degree of the Cayley graph. The "exceptional eigenvalues" that fall, approximately, into the interval $(-0.30, 0.30)$ are associated with the principal series representations induced from characters $\psi$ satisfying $\psi(-1) = -1$. The Fourier transforms of the characteristic function of the generating set evaluated at these representations do not depend on the group element $w$, since here we have

$$\begin{aligned}\rho\left(\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)\right) + \rho\left(\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)\right) + \rho\left(\left(\begin{smallmatrix}0&1\\-1&0\end{smallmatrix}\right)\right) + \rho\left(\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)\right)\\ = \rho\left(\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)\right) + \rho\left(\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)\right).\end{aligned}$$

Since $\rho(w)$ depends on $\psi$ but $\rho\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ does not, the eigenvalues in this interval appear with multiplicity order $p$. Since the total mass of the spectrum is of order $O(p^3)$, taken with respect to the counting measure, there is no asymptotic contribution from these eigenvalues. In other words, the spectral measure of the universal covering graph will contain a spectral gap in the approximate interval $(-0.30, 0.30)$.

The top right panel in Figure 7 shows the spectra for the discrete series representations associated with the generating set $\mathcal{G}_1$. Here the isolated eigenvalues appearing in a neighborhood of 0 are associated with discrete series representations built from nondecomposable characters $\nu$ such that $\nu(-1) = -1$. It is notable that the spectra resemble their principal series counterparts very closely, excepting the isolated eigenvalue at 1.

The middle row in Figure 7 shows the corresponding spectra for the generating set $\mathcal{G}_2$. Here again the exceptional eigenvalues in the approximate interval $(-0.10, 0.10)$ are due to representations associated with characters that take the value $-1$ at $-1$.


**FIGURE 7.** Principal series (left) and discrete series (right) spectra for the Cayley graph of $\mathrm{SL}_2(\mathbf{F}_p)$, with respect to the generator sets $\mathcal{G}_1$ (top), $\mathcal{G}_2$ (middle) and $\mathcal{G}_3$ (bottom).

The bottom row displays the spectra for the generating set $\mathcal{G}_3$. Note that these spectra, unlike the ones shown in the top and middle rows, have only two isolated eigenvalues, at 1 and $-3$ (unnormalized), excepting the common eigenvalue of 4, which results from the principal series representation induced from the identity.

### Random Generators

Figure 8 is a scatter-plot of the second-highest eigenvalue of Cayley graphs associated with random generating pairs, as described in the previous section. The data is shown here for the 36 primes between 30 and 200, with 25 random pairs generated for each prime.

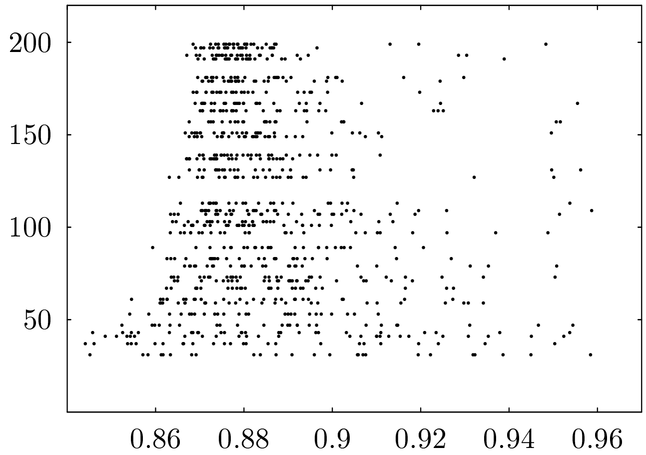


**FIGURE 8.** Second-highest eigenvalue for random generating pairs.

There is a clear accumulation of eigenvalues in the approximate interval from 0.868 to 0.888. This indicates that a random Cayley graph for $SL_2(p)$ is a significantly better expander than those Cayley graphs associated with the "natural" generators considered before. It also suggests that a random 4-regular Cayley graph for $SL_2(p)$, when $p$ is sufficiently large, is not a Ramanujan graph. A graph has the *Ramanujan property* [Bien 1989] if the inequality

$$\lambda_1 \leq 2\sqrt{k-1}$$

is satisfied, where $\lambda_1$ is the second-largest eigenvalue and $k$ is the degree of the graph. Since our graphs are 4-regular, the inequality becomes $\lambda_1 \leq 3.46410$, that is (taking into account our normalizations), the second-largest eigenvalue must be no larger than 0.86602. Figure 8 suggests that, asymptotically as $p \to \infty$, a random 4-regular Cayley graph over $SL_2(p)$ fails to meet this criterion.

### Comparison with Work of Buck

In [Buck 1986], certain computations are carried out that are closely related to ours. In particular, Buck considers the generating pair

$$\left\{ \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right) \right\}$$

over $PSL_2$, giving a Cayley graph of degree 3 comprised of triangles bridged together by a single edge at each vertex. This is the same graph as we have considered for generators $\mathcal{G}_3$, when taken over the projective group $PSL_2$, as shown in Figure 3. Over the cover $SL_2$, we obtain a graph where the triangles become hexagons, and where the lines bridging the triangles become squares. However, by a theorem of Kesten [1959], if $G$ is a countably generated group with normal subgroup $N$, we have

$$\lambda_1(X(G,S)) = \lambda_1(X(G/N,S))$$

for any generating set $S$, so long as the diffusion coefficient of the symmetric random walk on $N$ with respect to any set of generators is 1. ([Buck 1986] discusses an extension of this theorem to amenable groups.) In particular, this situation applies to the quotient of $SL_2$ by its center, and Buck's analysis of the generating function for the symmetric random walk on the graph of Figure 3 thus determines the second-largest eigenvalue for our set of generators $\mathcal{G}_3$. Intuitively, what this result implies for the generating set $\mathcal{G}_3$ is that the amount by which the expansion coefficient increases when we pass from triangles to hexagons is exactly cancelled by the decrease effected by the addition of more cycles (the squares that result from $w$ having order 4 over $SL_2$). For primes larger than 43, our computations agree closely with the asymptotic limit of

$$\frac{1+\sqrt{8\sqrt{2}+13}}{6} \approx 0.988482$$

established by the random-walk analysis.

In contrast, an exact asymptotic analysis of the endpoints of the spectra for the generator sets $\mathcal{G}_1$ and $\mathcal{G}_2$ seems more difficult to obtain. Over the projective group $PSL_2$, the covering Cayley graph for the generator set $\mathcal{G}_1$ is made up of adjacent hexagons, as shown in Figure 9.

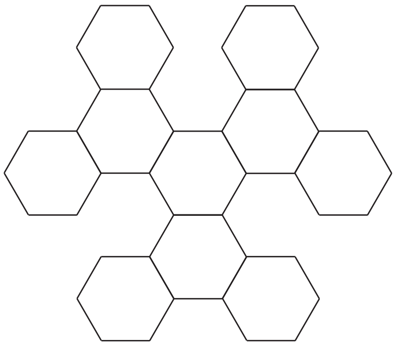

**FIGURE 9.** Covering Cayley graph for $\mathrm{PSL}_2(\mathbf{Z})$ with generators $\mathcal{G}_1$.

For this graph, the generating function analysis is more complicated, and while we can write down a set of six equations in six unknowns that the generating function must satisfy, we are unable to solve this system or obtain the radius of convergence of the return function. Similarly, the graph for generators $\mathcal{G}_2$ is made up of 9-gons, which provides us with the intuition that the spectral gap must be larger than for generators $\mathcal{G}_1$. This intuition is borne out in Figure 6. However, here again the probabilistic analysis appears difficult, though we can write down a system of equations characterizing the generating function.

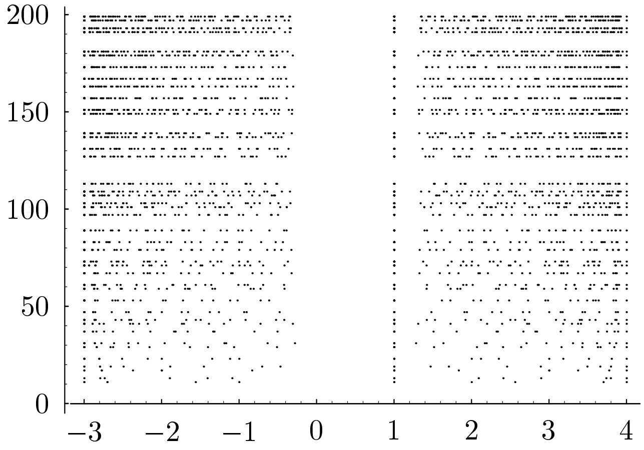


**FIGURE 10.** Action of $\mathcal{G}_3$ on $P^1(\mathbf{F}_p)$.

[Buck 1986] also gives a numerical analysis of the action of certain generating sets of $\mathrm{SL}_2(\mathbf{Z})$ on the projective line $P^1(\mathbf{F}_p)$, together with a conjecture that the action on this finite set "approximates" the action on the infinite group $\mathrm{SL}_2(\mathbf{Z})$. Our computations may be seen as providing further evidence for this phenomenon. In particular, we have observed that the spectrum obtained by evaluating the Fourier transform at a single representation closely approximates the full spectrum as $p$ gets large. Figure 10 plots the spectrum of the Cayley graph for generators $\mathcal{G}_3$ evaluated at the principal series representation induced from the identity, and should be compared to the graphs in the bottom row of Figure 7. This is precisely the graph corresponding to the action of $\mathcal{G}_3$ on the projective line $P^1(\mathbf{F}_p)$ that was considered in [Buck 1986].

## 6. SPECULATIONS AND OPEN PROBLEMS

We conclude this paper by presenting several speculations suggested by the data explained in Section 5.

Figures 4 and 5 suggest that for the generating sets $\mathcal{G}_1$ and $\mathcal{G}_2$, the second-largest eigenvalues are approximately 0.9821 and 0.9716, respectively.

The same figures indicate that from the point of view of the second-largest eigenvalue, the discrete and principal series are very similar. It would be interesting to obtain an analytic proof of a close upper bound or limit. Some recent work of Brooks [1991], building on [Buck 1986], gives techniques for obtaining this. The data also suggest that the convergence of the second-largest eigenvalue may very well be uniform in the following sense. Let $a, b$ be generators of $\mathrm{SL}_2(\mathbf{Z})$, and let $a_p, b_p$ be their images in $\mathrm{SL}_2(p)$. If $\{a_p, b_p\}$ generates $\mathrm{SL}_2(p)$ for all but a finite number of primes $p$, let $X_p(a, b)$ be the associated family of Cayley graphs. The data suggests that for $p$ sufficiently large there is an $\varepsilon_p$, independent of $\{a, b\}$, such that all fluctuations in the second-largest eigenvalue are within $\varepsilon_p$ of the limiting value.

**Open Question 6.1.** For the generating sets $\mathcal{G}_1$ and $\mathcal{G}_2$, do the second-largest eigenvalue occurring over all principal series representations and the second-largest eigenvalue occurring over all discrete series representations converge to the same limit as $p \to \infty$?

More generally, the pairs of graphs in Figure 7 suggest that the spectra of the principal series and discrete series are effectively "the same". Again, it might be of some interest to quantify this similarity in the form of a theorem. Such similarity could perhaps be quantified by comparing the associated spectral measures for operators corresponding to the direct sum of the discrete series representations and the principal-series representations. So, generalizing Open Question 6.1, we ask:

**Open Question 6.2.** For any generating pair, do the spectral measure associated with the direct sum of principal series representations and the spectral measure of the direct sum of the discrete series representations converge to the same limit as $p \to \infty$?

This would certainly be of interest from a computational point of view. As the discussion of Section 3 shows, spectral computations for the discrete series are computationally more intensive by a factor of $p$. A positive answer to Open Question 6.2 would permit any further numerical investigations to be carried out exclusively in the principal series, and consequently for a wider range of primes.

In this direction we would also like to remark on some numerical data not included here. Comparison of Figure 10 with the bottom row of Figure 7 seems to indicate that it may be the case that to understand the spectrum it is sufficient to study the Fourier transform evaluated at a single representation. Preliminary investigation appears to show that the spectra of $\hat{f}(\rho)$ for $\rho \neq \rho_\psi$, where $\psi(-1) = -1$, and $\rho \neq \rho_\nu$, where $\nu(-1) = -1$, in the notation of Theorems 2.1 and 2.3, are "the same", so that in fact perhaps only a single, arbitrary principal series Fourier transform need be computed.

As remarked in Section 5, Figure 7 reflects the convergence of the spectra to the spectrum of the infinite cover for these Cayley graphs by the natural Cayley graph on $\mathrm{SL}_2(\mathbf{Z})$. Again, the methods of [Brooks 1991] could possibly be used to compute precisely the support of the spectral measure for the infinite cover, so as to give the limiting distribution. This would also give the endpoints for the "intervals" seen in these graphs.

Figure 8 suggests many possible questions. The most striking property of this figure is that the majority of second-largest eigenvalues seems to be clustered in a small interval, roughly between .868 and .888. Note that the "Ramanujan number" for these graphs is $\sqrt{3}/2 \approx .86602$, so that none of the graphs generated for $p > 127$ were found to be Ramanujan. On the other hand, eigenvalues in the interval $(0.868, 0.888)$ are significantly lower than those for either of the generating sets $\mathcal{G}_1$ or $\mathcal{G}_2$. This suggests that a random Cayley graph of degree 4 on $\mathrm{SL}_2(p)$ has better expanding properties than those with "naturally" chosen generators.

**Open Question 6.3.** Is there a bound for the second-largest eigenvalue that holds for most generating pairs of $\mathrm{SL}_2(p)$, where "most" is to be interpreted in a sense similar to that of [Kantor and Lubotsky 1990]?

**Open Question 6.4.** Can one find a family of 4-regular Cayley graphs (indexed by $p$) whose second-largest eigenvalue is within these bounds? This would provide a family of Cayley graphs with better expanding properties.

Lastly, we would like to comment on the complexity results of Section 3. Recent work in the area of DFTs for finite groups [Baum 1991; Clausen 1989a,b; Diaconis and Rockmore 1990; Rockmore 1990a,b] has shown that the DFT can be computed in $O(|G| \log |G|)$ operations for several classes of groups. It would be of great interest if for $G = \mathrm{SL}_2(q)$ the results of Section 3 could be improved.

**Open Question 6.5.** Can one prove that

$$T(q) = O(q^3 \log q)?$$

## 7. APPENDIX: FOURIER INVERSION AND CONVOLUTION FOR $SL_2$

We now turn to the problem of efficient Fourier inversion and convolution for $\mathrm{SL}_2(K)$. As we noted in the Introduction, the existence of a fast Fourier inversion algorithm follows from general results [Baum and Clausen 1991] and from the upper bounds of Theorems 3.4 and 3.6. Here we provide a constructive and "implementable" proof of a fast inversion algorithm.

**Theorem 7.1.** *Let all notation be as in Section 1. Then $I(q) \leq O(q^4 \log q)$.*

Efficient algorithms for Fourier inversion and calculation of the Fourier transform for a given group together yield an efficient algorithm for computing group convolutions [Clausen 1989b, Corollary 1.8; Rockmore 1990b, Theorem 2]. In particular, by combining Theorems 7.1 and 3.7, we obtain

**Theorem 7.2.** *Let $f, g \in L^2(\mathrm{SL}_2(K))$. Assuming the additional initial data of the representations of $\mathrm{SL}_2(K)$, the convolution $f * g$ may be computed in at most $O(q^4 \log q)$ operations.*

Given Theorems 7.1 and 3.7, Theorem 7.2 is easy to prove. To compute $f * g$, simply compute the Fourier transforms $\{\hat{f}(\rho), \hat{g}(\rho)\}$, then the products $\{\hat{f}(\rho)\hat{g}(\rho) = \widehat{f * g}(\rho)\}$ and then perform Fourier inversion on this last Fourier transform. As the additional matrix multiplications will take at most $(q+1)q^3$ operations when done directly, the asserted bound is achieved.

To prove Theorem 7.1, the main idea is to try to apply the methods of [Rockmore 1990b] directly to the problem of Fourier inversion on $\mathrm{SL}_2(K)$. Close investigation of the computation involved will yield the asserted bound.

We summarize quickly the algorithm in [Rockmore 1990b]. Let $G$ be a group and $H \leq G$ a subgroup, with $s_1, \dots, s_k$ a complete set of coset representatives for $H \backslash G$. Let $\hat{G}$ and $\hat{H}$ be complete sets of irreducible representations of $G$ and $H$, respectively. Given the initial data of the Fourier transform of a function $f \in L^2(G)$ as a collection of matrices $\{\hat{f}(\rho)\}_{\rho \in \hat{G}}$, we wish to recover the values $\{f(s)\}_{s \in G}$. As in the efficient computation of the Fourier transform, the idea is to reduce this to a problem on $H$—in particular, to efficiently recover the restricted transforms $\{\hat{f}_i(\eta)\}_{\eta \in \hat{H}}$ for $1 \leq i \leq k$ (where the notation is as in Section 3), and then to perform Fourier inversion for the $k$ functions $f_i \in L^2(H)$.

To do this, let $\eta \in \hat{H}$, and suppose that

$$\eta{\uparrow}G \sim \rho_1 \oplus \cdots \oplus \rho_r,$$

where $\sim$ denotes equivalence of representations. Then, in one basis,

$$\hat{f}(\eta{\uparrow}G) = \begin{pmatrix} \hat{f}(\rho_1) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{f}(\rho_r) \end{pmatrix},$$

which can be built directly. However, by [Rockmore 1990b, Theorem 3], there exists a change of basis, and thus an invertible matrix $A_\eta$ (depending on only the representations $\rho_i$ and $\eta$), such that

$$A_\eta \begin{pmatrix} \hat{f}(\rho_1) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{f}(\rho_r) \end{pmatrix} A_\eta^{-1} = \begin{pmatrix} \hat{f}_1(\eta) & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ \hat{f}_k(\eta) & * & \cdots & * \end{pmatrix}, \tag{7.1}$$

where the asterisks denote block matrices of the appropriate dimensions.

We wish to apply this idea for $G = \mathrm{SL}_2(K)$ and $H = B$. If we can recover the restricted transforms on $B$ in $O(q^4 \log q)$ operations, we will have proved Theorem 7.1, since we have [Baum et al. 1991]

$$I(B) \leq 16(|B| \log |B|) \leq 32(q^2 \log q).$$

To proceed, we must first briefly explain the representation theory of $B$. This is a straightforward use of "Mackey theory", which takes advantage of the fact that $B = T \ltimes U$ (for details of such constructions see [Serre 1977, 62–63]). Thus, $U$ is a normal subgroup of $B$, so the irreducible representations of $B$ are built by first considering the action of $T$ on $\hat{U}$: we fix once and for all a nontrivial character $\chi$ of $U$. Every character in $\hat{U}$ can then be written as $\chi_\beta$, for some unique $\beta \in K$ given by

$$\chi_\beta(u) = \chi(\beta u)$$

for $u \in K$ (where we identify $U$ with $K$ under the natural isomorphism). The action of $T$ on $\hat{U}$ is by conjugation,

$$\begin{aligned} \left(\left(\begin{smallmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{smallmatrix}\right)\chi\right)\left(\left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right)\right) &= \chi\left(\left(\begin{smallmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{smallmatrix}\right)\right) \\ &= \chi\left(\left(\begin{smallmatrix} 1 & \alpha^{-2}u \\ 0 & 1 \end{smallmatrix}\right)\right) = \chi_{\alpha^{-2}}\left(\left(\begin{smallmatrix} 1 & u \\ 0 & 1 \end{smallmatrix}\right)\right). \end{aligned}$$

Thus, under the action of $T$, $\hat{U}$ splits into three orbits

$$\hat{U} = \{\chi_0 = 1\} \amalg \{\chi_\alpha : \alpha \in K^\times_{\mathrm{sq}}\} \amalg \{\chi_\alpha : \alpha \notin K^\times_{\mathrm{sq}}\}$$

where $K^\times_{\mathrm{sq}}$ denotes the set of nonzero squares in $K$.

It is clear that the stabilizer of $\chi_0$ in $T$ is $T$ itself, and that the stabilizers of $\chi_1$ and $\chi_\varepsilon$ both equal $\{\pm I\}$, where $I$ is the identity in $\mathrm{SL}_2(K)$ and $\varepsilon$ is any nonsquare in $K^\times$ (for example, a generator of $K^\times$). Let $-1$ denote the character of the subgroup $\{\pm I\}$ equal to $-1$ on $-I$. Then $\chi_1$ may be extended to the subgroup $\{\pm I\} \times U$ in two ways, as $1 \otimes \chi_1$ and $-1 \otimes \chi_1$, and similarly for $\chi_\varepsilon$. Set

$$\begin{aligned} \pi^+_+ &= (1 \otimes \chi_1){\uparrow}B, \\ \pi^+_- &= (-1 \otimes \chi_1){\uparrow}B, \\ \pi^-_+ &= (1 \otimes \chi_\varepsilon){\uparrow}B, \\ \pi^-_- &= (-1 \otimes \chi_\varepsilon){\uparrow}B. \end{aligned}$$

Each $\pi^t_s$ is then of degree $\frac{1}{2}(q-1)$, and the above four representations of $B$ are inequivalent. Finally, the remaining irreducible representations of $B$ are all one-dimensional, and are obtained by extending the trivial character on $U$ by any character of $T$.

The following lemma will eventually give a quick and simple construction of the intertwining matrices $A_{\pi_s^t}$.

**Lemma 7.3.** *With notation as above, each representation $\pi_s^t\uparrow G$, for $s,t \in \{+,-\}$, is multiplicity-free.*

*Proof.* By the preceding discussion,

$$\pi_s^t\downarrow U \sim \bigoplus_{a\in K_{\mathrm{sq}}^\times} \chi_a.$$

In particular, the restriction is a representation equivalent to the direct sum of distinct characters. Consequently, suppose that $\rho$ is an irreducible representation of $\mathrm{SL}_2(K)$ such that

$$\langle \rho, \pi_s^t\uparrow \mathrm{SL}_2(K)\rangle > 1.$$

(As usual, $\langle \rho_1, \rho_2\rangle$ represents the intertwining number of two representations $\rho_1$ and $\rho_2$ of a group $G$.) By Frobenius reciprocity, $\langle \rho\downarrow B, \pi_s^t\rangle > 1$. But this implies that $\rho\downarrow U$ is equivalent to the direct sum of a set of characters of $U$ (of size greater than 1) with multiplicity greater than 1. However, this contradicts the constructions of Section 3, where we see that $\rho\downarrow U$ contains at most one character with multiplicity greater than 1 if $\dim(\rho) \geq q$. □

To apply the constructions of [Rockmore 1990b], we require a basis for the representations of $\mathrm{SL}_2(K)$ that is "$B$-adapted". More precisely, let $\rho$ be a matrix representation of $\mathrm{SL}_2(K)$ such that $\rho\downarrow B \sim \eta_1 \oplus \cdots \oplus \eta_r$. Then we demand that

$$\rho(b) = \begin{pmatrix} \eta_1(b) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \eta_r(b) \end{pmatrix} \tag{7.2}$$

for all $b \in B$. (Note that the irreducible representations $\eta_i$ are fixed independently of the particular representation of $\mathrm{SL}_2(K)$ that is being decomposed.) In general, such $B$-adapted representations can always be constructed, and in fact we now discuss the necessary explicit construction. The discrete series representations are already $B$-adapted. For the principal series we require a new basis: in Section 3 we used the basis of $\delta$-functions on the set $K^+ \cup \{\infty\}$, and here instead we take $e_\infty \cup \hat{K}^+$ as a basis, which has the required property. The change-of-basis matrix is circulant, so multiplication by it requires at most $8q^2 \log q$ operations to perform, using standard abelian FFT techniques.

Thus, we now assume that we have a $B$-adapted set of irreducible representations and we now wish to construct $A_{\pi_s^t}$. Let

$$\pi_s^t\uparrow \mathrm{SL}_2(K) \sim \rho_1 \oplus \cdots \oplus \rho_r,$$

with $\dim(\rho_i) = d_i$. Fix coset representatives for $\mathrm{SL}_2(K)/B$ with $s_1 = 1$ and

$$s_j = \begin{pmatrix} 1 & j-2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

for $2 \leq j \leq q+1$. Using Lemma 7.3, we see that we are in a situation in which [Rockmore 1990b, Theorem 3] may be applied. We state the construction in the form of a lemma.

**Lemma 7.4.** *Let all notation be as above, the basis for the representations of $\mathrm{SL}_2(K)$ having been chosen in such a way that (7.2) holds. Then $A_{\pi_s^t}$ is a block matrix with blocks $B_{i,j}$, where $1 \leq i \leq r$ and $1 \leq j \leq q+1$ and each $B_{i,j}$ is a $d_i \times \frac{1}{2}(q-1)$ matrix. In particular, if $\pi_s^t$ comprises the first or second $\frac{1}{2}(q-1) \times \frac{1}{2}(q-1)$ diagonal block of $\rho_i\downarrow B$ (we may assume that one of these two instances occurs), $B_{i,j}$ will comprise exactly the first or second $\frac{1}{2}(q-1)$ columns of $\rho_i(s_j)$, respectively.*

Thus, consider now the computation of the matrix multiplication

$$A_{\pi_s^t}\begin{pmatrix} \hat{f}(\rho_1) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{f}(\rho_r) \end{pmatrix} A_{\pi_s^t}^{-1}.$$

We are only interested in the first $\frac{1}{2}(q-1)$ columns. So we first compute these columns for the rightmost pair of factors,

$$\begin{pmatrix} \hat{f}(\rho_1) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{f}(\rho_r) \end{pmatrix} A_{\pi_s^t}^{-1}.$$

Using the block diagonal structure of the right factor, it is easy to see that there are at most $q$ blocks of size $q+1$, so that by direct multiplication we get at most $q(q+1)^3$ operations. We are thus left with the problem of computing

$$\begin{pmatrix} B_{1,1} & B_{1,2} & \cdots & B_{1,q+1} \\ B_{2,1} & B_{2,2} & \cdots & B_{2,q+1} \\ \vdots & \vdots & \ddots & \vdots \\ B_{r,1} & B_{r,2} & \cdots & B_{r,q+1} \end{pmatrix} \begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_{q+1} \end{pmatrix} = \begin{pmatrix} \hat{f}_1(\pi_s^t) \\ \hat{f}_2(\pi_s^t) \\ \vdots \\ \hat{f}_1(\pi_s^t) \end{pmatrix}, \tag{7.3}$$

where each $M_j$ is a $\frac{1}{2}(q-1) \times \frac{1}{2}(q-1)$ matrix.

We wish to show that in fact computation of (7.3) may be viewed again as the computation of a Fourier transform of a suitably defined function on $\mathrm{SL}_2(K)$.

Consider the function $g \in L^2(G)$ defined by

$$\hat{g}_i(\pi_s^t) = M_i$$

and $\hat{g}_i(\eta) = 0$ for all other irreducible representations $\eta$ of $B$ (thus, we have defined $g$ by describing the Fourier transforms of the derived functions $g_i \in L^2(B)$). If $\pi_s^t$ makes up the first block of $\rho_i{\downarrow}B$, say, we see that

$$\begin{aligned} \hat{g}(\rho_i) &= \sum_{k=1}^{q+1} \rho_i(s_j) \begin{pmatrix} \hat{g}_k(\pi_s^t) & 0 \\ 0 & 0 \end{pmatrix} \\ &= \sum_k \begin{pmatrix} B_{i,k} M_k & \bar{0} \end{pmatrix}, \end{aligned}$$

where $\bar{0}$ denotes the $(d_i - \frac{1}{2}(q-1)) \times d_i$ matrix of zeros.

The results of Section 3 show that these computations may all be performed in at most $25q^4 \log q$ operations. Doing this for each of the four $\pi_s^t$, we see that the matrices $\{\hat{f}_k(\pi_s^t)\}_{k,s,t}$ can be recovered in at most

$$4(25q^4 \log q + q(q+1)^3) + 8q^3 \log q \le 108 q^4 \log q$$

operations.

Finally, we need to obtain the restricted transforms at the one-dimensional representations $\tilde{\psi}$ (in the notation of Section 2). Writing down the appropriate matrices for (7.1), we see that in this case we need only recover the first column of a $(q+1) \times (q+1)$ matrix. This requires at most $(q+1)^2$ operations. Repeating for each character gives at most $(q-1)(q+1)^2$ operations. Thus, in total we require at most

$$\begin{aligned} 108q^4 \log q + (q-1)(q+1)^2 + 16q^3 \log q &\le 110 q^4 \log q \\ &= O(q^4 \log q) \end{aligned}$$

operations for Fourier inversion. This completes the proof of Theorem 7.1.

## ACKNOWLEDGEMENTS

We thank Alex Lubotzky and Peter Sarnak for their interest and suggestions, especially regarding the computations. We have also benefited from discussions with Laci Babai, Bob Brooks and Michael Clausen.

## REFERENCES

[Alon 1983] N. Alon, "Eigenvalues and expanders", *Combinatorica* **6** (1983), 83–96.

[Alon and Milman 1985] N. Alon and V. D. Milman, "$\lambda_1$, isoperimetric inequalities for graphs and superconcentrators", *J. Combin. Theory* B**38** (1985), 73–88.

[Baum 1991] U. Baum, "Existence and efficient construction of fast Fourier transforms on supersolvable groups", Doctoral dissertation, Institut für Informatik, Universität Bonn, 1991.

[Baum and Clausen 1991] U. Baum and M. Clausen, "Some lower and upper complexity bounds for generalized Fourier transforms and their inverses", *SIAM J. Comput.* **20** (1991), 451–459.

[Baum et al. 1991] U. Baum, M. Clausen and B. Tietz, "Improved upper complexity bounds for the discrete Fourier transform" *Applicable Algebra in Engineering, Communication and Computing* **2** (1991), 35–43.

[Bien 1989] F. Bien, "Constructions of telephone networks by group representations" *Notices of the AMS* **36**(1) (1989), 5–22.

[Biggs 1974] N. Biggs, *Algebraic Graph Theory*, Cambridge University Press, Cambridge (UK), 1974.

[Brooks 1991] R. Brooks. "Some relations between spectral geometry and number theory", technical report, Department of Mathematics, UCLA, 1991.

[Bshouty et al. 1988] N. Bshouty, M. Kaminski and D. Kirkpatrick, "Addition requirements for matrix and transposed matrix products", *J. of Algorithms* **9** (1988), 354–364.

[Buck 1986] M. W. Buck, "Expanders and diffusers", *SIAM J. Algebraic and Discrete Methods* **7** (1986), 282–304.

[Chung 1989] F. Chung, "Diameters and eigenvalues", *J. Amer. Math. Soc.* **2** (1989) 187–200.

[Clausen 1989a] M. Clausen, "Fast generalized Fourier transforms", *J. Theor. Comp. Sci.* **67** (1989), 55–63.

[Clausen 1989b] M. Clausen, "Fast Fourier transforms for metabelian groups", *SIAM J. Comput.* **18** (1989), 584–593.

[Diaconis 1988] P. Diaconis, *Group Representations in Probability and Statistics,* Institute of Mathematical Statistics, Hayward, CA, 1988.

[Diaconis and Rockmore 1990] P. Diaconis and D. Rockmore, "Efficient computation of the Fourier transform on finite groups", *J. Amer. Math. Soc.* **3** (1990), 297–332.

[Dickson 1958] L. Dickson, *Linear Groups with an Exposition of the Galois Field Theory,* Dover, New York, 1958.

[Jordan 1907] H. Jordan, "Group characters of various types of linear groups", *Amer. J. Math.* **29** (1907), 387–405.

[Kesten 1959] H. Kesten, "Symmetric random walks on groups", *Trans. Amer. Math. Soc.* **92** (1959), 336–354.

[Kantor and Lubotzky 1990] W. Kantor and A. Lubotzky, "The probability of generating a finite classical group", *Geom. Ded.* **36** (1990), 67–87.

[Kloosterman 1946] H. D. Kloosterman, "The behavior of general theta functions under the modular group and the characters of binary modular congruence groups", *Ann. of Math.* **47** (1946), 317–375, 376–447.

[Lenstra and Lenstra 1990] A. K. Lenstra and H. Lenstra, "Algorithms in number theory", pp. 673–717 in *Handbook of Theoretical Computer Science, A: Algorithms and Complexity* (edited by J. Van Leeuwen), MIT Press, Cambridge, MA, 1990.

[Lubotzky] A. Lubotzky, "Discrete groups, expanding graphs, and invariant measures" (to appear).

[Naimark and Stern 1990] M. A. Naimark and A. I. Stern, *Theory of Group Representations*, Springer-Verlag, New York, 1980.

[Piatetski-Shapiro 1983] I. Piatetski-Shapiro, *Complex Representations of* $\mathrm{GL}[2, K]$ *for Finite Fields* $K$*,* Cont. Math. **16**, American Mathematical Society, Providence, RI, 1983.

[Rockmore 1990a] D. Rockmore, "Fast Fourier analysis for abelian group extensions", *Adv. in Appl. Math.*, **11** (1990), 164–204.

[Rockmore 1990b] D. Rockmore, "Efficient computation of Fourier inversion for finite groups", technical report, Department of Mathematics, Harvard, 1990.

[Sarnak 1990] P. Sarnak, *Some Applications of Modular Forms*, Cambridge Univ. Press, Cambridge (UK), 1990.

[Schur 1907] I. Schur, "Untersuchungen über die Darstellungen der endlichen Gruppen durch gebrochene lineare Substitutionen", *J. Reine Angew. Math.* **132** (1907), 85–117.

[Serre 1977] J. P. Serre, *Linear Representations of Finite Groups,* Springer-Verlag, New York, 1977.

[Silberger 1969] A. Silberger, "An elementary construction of the representations of $\mathrm{SL}(2, GF(q))$", *Osaka J. Math.* **6** (1969), 329–338.

[Suzuki 1982] M. Suzuki, *Group Theory, volume I*, Springer-Verlag, New York, 1982.

[Tanaka 1967] S. Tanaka, "Construction and classification of irreducible representations of the special linear group of the second order over a finite field", *Osaka J. Math.* **4** (1967), 65–84.

John D. Lafferty, IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, NY 10598 (jlaff@watson.ibm.com)

Daniel Rockmore, Department of Mathematics, Columbia University, New York, NY 10027
Current address: Department of Mathematics, Dartmouth College, Hanover, NH 03755 (rockmore@aruba.dartmouth.edu)

Received December 19, 1991; revised August 17, 1992