JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 10, Number 1, January 1997, Pages 169–214
S 0894-0347(97)00219-1

SEPARATION OF VARIABLES AND THE COMPUTATION OF FOURIER TRANSFORMS ON FINITE GROUPS, I

DAVID K. MASLEN AND DANIEL N. ROCKMORE

1. Introduction

Recently, increased attention has been paid to the problem of finding efficient algorithms for the computation of Fourier transforms on nonabelian groups. The abelian case has a long history (cf. [22, 23, 38]), and ever since the publication of the Cooley-Tukey fast Fourier transform (FFT) [24] these algorithms have been at the heart of digital signal processing (see for example [3, 31, 63, 62, 66] and the many references contained therein). The nonabelian cases have also been motivated by applications. They have been found useful in new approaches to data analysis [11, 26, 27], VLSI design [12, 13], the design of filters [43, 65] and efficient group convolution algorithms [20, 56]. In the continuous setting, there are applications to computer vision, geophysics and climate modeling (see, e.g., [30, 37]).

Apart from applications, these algorithms contribute to the understanding of the representation theoretic content of the fast Fourier transform. Although an abelian group has a unique Fourier transform, a nonabelian group has an infinite number of Fourier transforms, each of which corresponds to different choices of bases for the irreducible representations of the group. The complexity of a finite group G is defined as the least upper bound, over all choices of bases, of the complexities of the algorithms computing Fourier transforms on G. A direct approach to the computation shows that the complexity is bounded above by $|G|^2$. We conjecture that all finite groups have complexity $O(|G| \log^c |G|)$ for some universal constant c. This has already been proved for many different classes of nonabelian groups [7, 18, 57, 55].

The first results of this type obtained for nonabelian groups are due to Willsky. In [65] Willsky studies a particular class of finite state Markov processes evolving on metacyclic groups. In so doing he gives an $O(|G| \log |G|)$ FFT for G a metacyclic group, designed to give an efficient nonlinear filtering algorithm for the situation in which noisy observations are taken. At the close he remarks ([65], p. 205) "there are quite likely to be far larger classes of groups for which fast transforms exist". Since then, the general problem of constructing efficient Fourier transform algorithms has been treated by Beth [12, 13], Diaconis and Rockmore [29], and Clausen [18, 19],

Received by the editors March 8, 1996 and, in revised form, May 23, 1996.
1991 *Mathematics Subject Classification.* Primary 20C15; Secondary 65T10.

A preliminary version of some of this work appears as an extended abstract in the *Proceedings of the 1995 ACM-SIAM Symposium on Discrete Algorithms*, pp. 253–262.

Maslen was partially supported as a Shapiro Visitor while at Dartmouth. Rockmore was partially supported by ARPA as administered by the AFOSR under contract DOD F4960-93-1-0567 as well as NSF DMS Award 9404275 and an NSF Presidential Faculty Fellowship.

with others treating certain particular cases. For an overview of some of the recent work in this area, we refer the reader to [49]; see also [54] for a discussion of applications.

We continue the general investigation of efficient Fourier transform algorithms. We present a divide and conquer strategy for computing nonabelian Fourier transforms, which encompasses many known FFTs, and provides new fast algorithms in other cases. It has two main components. First, we use a set of factorizations of elements of G to write the matrix sum of products that defines the Fourier transform in terms of a sequence of sums of products which are easier to compute. We call this technique separation of variables and the corresponding algorithm is the separation of variables algorithm.

The second part of our strategy uses a subgroup chain for the group and the notion of a subgroup-adapted set of representations. When computing with a subgroup-adapted set of representations the matrix multiplications that occur in the separation of variables algorithm have a highly structured and sparse form and may therefore be computed efficiently. We provide a thorough analysis of the structure of these matrices and the operation count of the corresponding matrix multiplications. The main tool used here is a form of Schur's Lemma which determines the structure of the representation matrix of a group element which commutes with a subgroup. The bulk of the new computational savings of this paper come from this use of commutativity. We believe this is a new contribution to the subject, although some aspects of it do appear in the work of Clausen on the symmetric group [18, 19] and that of Rockmore on wreath products [55]. The general idea of factoring representation matrices in terms of sparse matrices was first formulated by Clausen [19].

Our techniques are quite general. We obtain upper bounds for the complexity of the Fourier transform of any group or homogeneous space in terms of representation theoretic data. These bounds are expressed in terms of multiplicities of restrictions of irreducible representations from one subgroup to another. We thereby obtain a general procedure for bounding the complexity of the Fourier transform on a group or homogeneous space, which enables us to find explicit bounds even when the representation matrices are extremely complicated. In this way we derive both previously known and new results as part of a general theory, instead of using ad hoc techniques.

This paper does not use the full strength of the separation of variables approach, but despite this we recover the fastest known algorithms for many abelian groups, the symmetric groups, and their wreath products. Furthermore, we obtain new fast algorithms for matrix groups over finite fields. A more detailed analysis of the computation improves the results; that is the content of part II of this work [50]. By dividing the work in this way we hope to present general results of interest without obscuring them with the technical machinery needed for more refined results.

We start the paper in Section 2 with the definitions of Fourier transform, complexity, and adapted representation. In Section 3, we explain the previously known technique of reducing to subgroups. Section 4 forms the theoretical core of the paper; it contains the definition of the separation of variables algorithm, the analysis of matrix products, and the general complexity results that we use in our examples. Following this, Section 5 develops results on the complexities of specific groups. We start it by deriving the Cooley-Tukey algorithm in the context of finite abelian groups, the results of Clausen and Baum [18] on the symmetric group, results on classical Weyl groups, and the results of Rockmore [55] on wreath products. We then give algorithms for the general linear and unitary groups over a finite field, and finish our examples with some results on classical Chevalley groups over finite fields. In Section 6, we apply the separation of variables algorithm to homogeneous spaces. We analyze the matrix products that occur in this new setting, and give results for homogeneous spaces of the symmetric group, classical Weyl groups, and the general linear and unitary groups. Finally, we summarize the consequences of this work and indicate the contents of part II of this paper [50].

Our bounds depend on some explicit knowledge of the restrictions to a subgroup and often involve the number of conjugacy classes in a group (i.e., the number of irreducible representations). For some of our results we need asymptotics for these quantities. To avoid interrupting the flow of the paper we have postponed this discussion to an appendix at the end.

Acknowledgements

Special thanks to Tom Hagedorn for explaining his interesting recent work on multiplicities for restricted representations, and to Prof. Dennis Stanton for explaining how to work with the power series in the appendix. Thanks also to Prof. Michael Clausen for some very helpful conversations.

2. Background

2.1. **Nonabelian Fourier transforms.** The familiar discrete Fourier transform (DFT) of a finite data sequence and its efficient computation via the Cooley-Tukey fast Fourier transform [24] has a natural formulation in terms of the representation theory of cyclic groups. Phrased in this context, the algorithm generalizes to an algorithm for computing Fourier transforms on arbitrary finite groups. What follows is a brief review of the basic concepts and definitions needed to formulate the problem on arbitrary groups. For a complete introduction to the representation theory of finite groups Serre's book [58] is a good reference.

Recall that a (complex) matrix representation of a finite group G is a map ρ from G into the group of $d \times d$ invertible matrices with complex entries, $GL_d(\mathbf{C})$, such that

$$\rho(st) = \rho(s)\rho(t)$$

for every $s, t \in G$. In this case d is called the degree or dimension of the representation ρ, and is denoted d_ρ and $V = \mathbf{C}^d$ is called the representation space of ρ.

Two representations ρ_1 and ρ_2 are said to be equivalent if they differ only by a change of basis, i.e., if there exists an invertible matrix A such that $\rho_1(s) = A^{-1}\rho_2(s)A$ for all $s \in G$. Notice that 1-dimensional matrix representations are uniquely determined by their equivalence class, whereas multidimensional irreducible representations have an infinite number of equivalent realizations.

A subspace $W \subset V = \mathbf{C}^d$ is said to be G-invariant if for all $s \in G$, $\rho(s)W \subset W$. The representation ρ is said to be irreducible if V has no G-invariant subspaces other than the trivial subspaces $\{0\}$ and V, and is called reducible otherwise. Up to equivalence there are only a finite number of irreducible representations of any finite group — in fact there are as many as there are conjugacy classes in the group.

Irreducible representations are the fundamental building blocks of all representations of a finite group. That is, any representation is equivalent to a direct sum of irreducible representations, where the direct sum of two representations is the matrix direct sum of the representations.

There are several equivalent definitions of the Fourier transform for a finite group [12, 20, 29]. The following is the most convenient for this paper.

Definition 2.1 (Fourier Transform)**.** Let G be a finite group and f be a complex-valued function on G.

(i) Let ρ be a matrix representation of G. Then the **Fourier transform of** f **at** ρ, denoted $\hat{f}(\rho)$, is the matrix sum,

$$\hat{f}(\rho) = \sum_{s \in G} f(s)\rho(s). \tag{1}$$

(ii) Let $\mathcal{R}$ be a set of matrix representations of G. Then the **Fourier transform of** f **on** $\mathcal{R}$ is the set of Fourier transforms of f at the representations in $\mathcal{R}$.

Fast Fourier transforms or **FFTs** are algorithms for computing Fourier transforms efficiently. The Fourier transform at a set $\mathcal{R}$ of representations can always be related to the Fourier transform at the single representation $\Delta = \bigoplus_{\rho \in \mathcal{R}} \rho$. Thus, by taking direct sums, results about single representations may be transferred to results about sets of representations.

The most important case of a Fourier transform occurs when the set $\mathcal{R}$ is a complete set of inequivalent irreducible representations of G. In this situation we shall simply refer to such a calculation as the **computation of a Fourier transform**. A Fourier transform determines f through the Fourier inversion formula.

Theorem 2.1 (Fourier inversion formula; see, e.g., [27], p. 13)**.** *Let G be a finite group, f a complex-valued function on G, and $\mathcal{R}$ a complete set of irreducible matrix representations of G. Then,*

$$f(s) = \frac{1}{|G|} \sum_{\rho \in \mathcal{R}} d_\rho \, trace\Big(\hat{f}(\rho)\rho(s^{-1})\Big) \tag{2}$$

where $d_\rho = \dim(\rho)$.

Example: The "usual" discrete Fourier transform. The irreducible matrix representations of the cyclic group $\mathbf{Z}/n\mathbf{Z} = \{0, 1, \ldots, n-1\}$, are all one-dimensional. For each integer j with $0 \leq j \leq n-1$, define the representation ζ_j, by $\zeta_j(k) = \exp(\frac{2\pi ijk}{n})$ for $k \in \mathbf{Z}/n\mathbf{Z}$. The set of such representations is a complete set of inequivalent irreducible representations for $\mathbf{Z}/n\mathbf{Z}$ and the corresponding Fourier transform is usually known as the discrete Fourier transform. This computation is central to the subject of digital signal processing (cf. [53]).

The arithmetic complexity for computing a Fourier transform conceivably depends on the choice of basis for the irreducible representations. The notion of the complexity of a finite group provides a classification of finite groups according to the complexity of the most efficient algorithm to compute some such transform on the group.

Definition 2.2 (Complexity)**.** Let G be a finite group, and $\mathcal{R}$ any set of matrix representations of G. Let $T_G(\mathcal{R})$ denote the minimum number of operations needed to compute the Fourier transform of f on $\mathcal{R}$ via a straight-line program for an arbitrary complex-valued function f defined on G. $T_G(\mathcal{R})$ is called the **complexity of the Fourier transform for the set** $\mathcal{R}$. Define the **complexity of the group** G to be

$$\mathcal{C}(G) = \min_{\mathcal{R}}\{T_G(\mathcal{R})\}$$

where $\mathcal{R}$ varies over all complete sets of inequivalent irreducible matrix representations of G.

The computational model used here is a common one in which an operation is defined as a single complex multiplication followed by a complex addition.

Elementary representation theory shows that the sum of the squares of the degrees of a complete set of irreducible representations of G is equal to $|G|$ (see, e.g., [58], p. 18). Consequently direct computation of any Fourier transform gives the upper and lower bounds

$$|G| - 1 \leq \mathcal{C}(G) \leq |G|^2.$$

As mentioned in Section 1, the techniques introduced in this paper show how structural properties of the group and a judicious choice of the set of representations $\mathcal{R}$ provide significantly better upper bounds for group complexity. When bounding $T_G(\mathcal{R})$ it is often easier to work with a related quantity, the **reduced complexity**, denoted by $t_G(\mathcal{R})$ and defined by

$$t_G(\mathcal{R}) = T_G(\mathcal{R})/\left|G\right|. \tag{3}$$

This definition simplifies the statements and proofs of many following results.

Remark. Another common interpretation of the Fourier transform is as a change of basis for the group algebra $\mathbf{C}[G]$, from the basis of point masses on G to a basis of matrix coefficients coming from a complete set of inequivalent irreducible representations. When this point of view is adopted, the complexity of the Fourier transform can be measured as the c-linear complexity of the associated change of basis matrix (cf. [8]). The c-linear complexity of a group G is defined to be the minimum c-linear complexity of any such matrix for G. Assuming a choice of unitary representations (which is always possible) the results stated here can all be interpreted as statements about the 2-linear complexity of finite groups.

2.2. **Adapted sets of representations.** As remarked earlier, there are an infinite number of matrix representations equivalent to any given nontrivial multidimensional matrix representation, all related by a change of basis. Even among equivalent representations the complexity of the associated Fourier transform might vary. For this reason and others, subgroup-adapted sets of representations have been found to be useful for efficiently computing Fourier transforms. Use of these representations permits the computation of a Fourier transform on a finite group G, to be built up from the computation of several Fourier transforms on a chosen subgroup H.

To briefly explain the idea, let H be a subgroup of a group G. An **H-adapted set of representations of** G has the property that when considered as representations of H via restriction, they may be constructed as matrix direct products of representations from a fixed set of inequivalent irreducible matrix representations of H. As shown in [29] and [18] (and explained in the next section), a Fourier transform on G can always be written as a sum over a set of matrix multiplications against Fourier transforms at the restrictions of the representations to the subgroup

H. Requiring that the restrictions are H-adapted allows us to reduce the computation of Fourier transforms at the restricted representations to Fourier transforms at a set of irreducible representations.

Definition 2.3 (Subgroup-adapted representations)**.** Let G be a finite group and $\mathcal{R}$ be a set of matrix representations of G and let H be a subgroup of G. If ρ is a representation of G, let $\rho\!\downarrow\! H$ denote the representation of H obtained by restricting ρ to H. We say that $\mathcal{R}$ is H**-adapted** if there is a set $\mathcal{R}_H$ of inequivalent irreducible matrix representations of H such that the set of restricted representations

$$\mathcal{R}\!\downarrow\! H \;=\; \{\rho\!\downarrow\! H \mid \rho \in \mathcal{R}\}$$

is a matrix direct sum of representations in $\mathcal{R}_H$.

Notice that if $\mathcal{R}$ is H-adapted, then the set $\mathcal{R}_H$ is uniquely determined by $\mathcal{R}$. When $H = G$, the property of being G-adapted allows us to reduce the computation of the Fourier transform of f on $\mathcal{R}$ to a Fourier transform on G at a set of inequivalent irreducible representations.

Lemma 2.2. *If $\mathcal{R}$ is a G-adapted set of matrix representations of G then $T_G(\mathcal{R}) = T_G(\mathcal{R}_G)$.*

Remark. The FFT algorithms presented in the following sections all assume the use of adapted sets of representations. The requirement of adaptability does not limit us, as any set of representations is equivalent to an adapted set of representations. An equivalent concept is that of an adapted basis, also known as a Gel'fand-Tsetlin basis. A basis for a representation space is adapted to a subgroup if the matrix representation obtained by expressing the representation in coordinates for this basis is also adapted. Adaptedness for a set of bases is defined similarly. Adapted bases always exist and, in fact, can always be constructed.

To outline one such construction, we collect several previously known results. Babai and Rónyai [6] have shown that a complete set of irreducible representations of a finite group G can be constructed in polynomial time from the multiplication table of G. Further techniques from [6] or [4] provide efficient algorithms for decomposing representations into their irreducible constituents. By applying these results to the original set of representations restricted to the subgroup H, a complete set of irreducible representations for H is then found. A change of basis to insure that all representations of G are H-adapted is computed by the construction of certain projection operators. This last step is detailed in the book of Fässler and Stiefel [32] which also provides a wealth of examples of uses of adapted bases in a variety of computational problems.

3. Coset decompositions and the Fourier transform

In previous work, adapted representations have already been used to speed the computation of Fourier transforms by factoring the computation through a subgroup [18, 29]. The idea is to use the coset decomposition of elements in the group to relate a Fourier transform on G to Fourier transforms on a subgroup H. This may be thought of as the simplest example of the separation of variables technique (cf. Section 4).

To explain, let H be a subgroup of G and let $Y \subset G$ be a set of coset representatives for G/H. Thus, G can be factored as the disjoint union of subsets

$yH = \{yh \mid h \in H\}$ for all $y \in Y$. For any representation ρ of G we can use the relation $\rho(ab) = \rho(a)\rho(b)$ to produce a factorization of $\hat{f}(\rho)$ by

$$\begin{aligned} \hat{f}(\rho) &= \sum_{s\in G} f(s)\rho(s) \\ &= \sum_{y\in Y} \rho(y) \sum_{t\in H} f_y(t)\rho(t) \end{aligned} \tag{4}$$

where for each $y \in Y$, f_y is the function on H defined by $f_y(t) = f(yt)$ for all $t \in H$. Consequently, with the notation of (4) we can rewrite $\hat{f}(\rho)$ as a sum of Fourier transforms on H,

$$\hat{f}(\rho) = \sum_{y\in Y} \rho(y)\widehat{f_y}(\rho\downarrow H). \tag{5}$$

If we had computed the Fourier transform of f_y on $\mathcal{R}\downarrow H$ for a complete set of irreducible representations $\mathcal{R}$ of G and for all $y \in Y$, then each $\hat{f}(\rho)$ could be built from the Fourier transforms $\widehat{f_y}(\rho\downarrow H)$ using equation (5). The matrices $\rho(y)$ are analogous to the "twiddle factors" that arise in the Cooley-Tukey algorithm when G is an abelian group. In that case all irreducible representations are one-dimensional and the matrices $\rho(y)$ are simply roots of unity.

In general, a restricted representation $\rho\downarrow H$ may be reducible, even when ρ is irreducible. If ρ is H-adapted, then $\rho\downarrow H$ is not only equivalent to, but also equal to, a matrix direct sum of irreducible representations, and all equivalent irreducible representations that occur in this sum are equal. In this case $\widehat{f_y}(\rho\downarrow H)$ can be constructed as a block diagonal matrix from the matrices of the appropriate Fourier transforms of f_y at irreducible representations of H.

The discussion above yields directly an algorithm for computing the Fourier transform of any function f on G using any given H-adapted set $\mathcal{R}$ of representations of G:

1. Choose a set of coset representatives Y for G/H. For each $y \in Y$ compute the Fourier transform of f_y on $\mathcal{R}_H$.
2. For each $\rho \in \mathcal{R}$ build the restricted transforms $\widehat{f_y}(\rho\downarrow H)$. These will be block diagonal matrices with blocks given by the individual Fourier transforms of f_y at the representations of $\mathcal{R}_H$.
3. Compute the products $\rho(y)\widehat{f_y}(\rho\downarrow H)$ and add them together.

To obtain an upper bound for the complexity of this basic algorithm it is useful to introduce some notation. Let $\mathcal{R}$ be a set of matrix representations of G and let Y be any subset of G. Then we define

$$M_G(Y,\mathcal{R}) = \begin{cases} \text{the minimum number of operations required to compute} \\ \text{the collection of sums, } \left\{\sum_{y\in Y}\rho(y)F(y,\rho) \mid \rho \in \mathcal{R}\right\}, \\ \text{where for each } \rho \in \mathcal{R} \text{ and } y \in Y,\ F(y,\rho) \text{ is an arbitrary} \\ d_\rho \times d_\rho \text{ matrix.} \end{cases} \tag{6}$$

Similarly, define a "reduced" version of (6) by

$$m_G(Y,\mathcal{R}) = \frac{M_G(Y,\mathcal{R})}{|G|}. \tag{7}$$

Theorem 3.1 ([29], Proposition 1; [18])**.** *Let H be a subgroup of G and let $\mathcal{R}$ be a complete H-adapted set of inequivalent irreducible matrix representations of G. Let*

$Y \subset G$ be a set of coset representatives for G/H. Then with the notation of (6) *and* (7)

$$T_G(\mathcal{R}) \le |G/H|\, T_H(\mathcal{R}_H) + M_G(Y,\mathcal{R}) \tag{8}$$

or equivalently

$$t_G(\mathcal{R}) \le t_H(\mathcal{R}_H) + m_G(Y,\mathcal{R}). \tag{9}$$

A better bound may be obtained using the block diagonal form of $\widehat{f}_y(\rho\!\downarrow\! H)$. We take this into account in Sections 4.2 and 4.1.

The inequalities (8) and (9) can be viewed as recurrences which bound the complexity of a group in terms of the complexity of a subgroup. The recurrence may be iterated through a chain of subgroups for G. For example consider the chain of subgroups

$$G = K_n > K_{n-1} > \cdots > K_0. \tag{10}$$

We say that $\mathcal{R}$, a set of irreducible representations of G, is **adapted to the chain** (10) provided $\mathcal{R}$ is K_i-adapted for each subgroup K_i in the chain. Using the notation of Definition 2.3, this implies that each $\mathcal{R}_{K_i}$ is K_j-adapted for $j \le i$. Theorem 3.1 now generalizes immediately.

Theorem 3.2. *Let G have the chain of subgroups* (10) *and for $i = 1, \ldots, n$, let Y_i be a set of coset representatives for K_i/K_{i-1}. If $\mathcal{R}$ is a set of matrix representations of G which is adapted to this chain, then*

$$t_G(\mathcal{R}) \le t_{K_0}(\mathcal{R}_{K_0}) + \sum_{i=1}^{n} m_{K_i}(Y_i, \mathcal{R}_{K_i}). \tag{11}$$

Theorems 3.1 and 3.2 suggest that one approach to minimizing an upper bound of t_G, and hence T_G, is to try to efficiently evaluate sums of the form $\sum_{y\in Y} \rho(y)F(y)$, where the $F(y)$ are $d_\rho \times d_\rho$ matrices. Towards this end several possibilities are evident. The subgroup chain can be varied, as can the choice of coset representatives, so as to obtain matrices $\rho(y)$ with useful computational properties. Another idea is to attempt to use the properties of the matrix elements of $\rho(y)$ as special functions on the set Y. In this paper we explore the first approach.

When $G = H \times K$ is a direct product we get a special case of Theorem 3.1. The irreducible representations of G may all be obtained as tensor products of those of H and K, and the product basis constructed by the tensoring of a basis for the irreducible representations of H with those of K yields irreducible representations which are both H-adapted and K-adapted, up to a relabeling of the matrix rows and columns (cf. [13], Satz 5.8). If $\mathcal{R}'$ and $\mathcal{R}''$ are sets of matrix representations of H and K respectively then let $\mathcal{R}' \otimes \mathcal{R}''$ be the set of matrix tensor products of representations in $\mathcal{R}'$ with those in $\mathcal{R}''$.

Theorem 3.3. (i) *If $\mathcal{R}'$ and $\mathcal{R}''$ are sets of matrix representations of representations of H and K respectively, then*

$$t_{H\times K}(\mathcal{R}' \otimes \mathcal{R}'') \le t_H(\mathcal{R}') + t_K(\mathcal{R}'').$$

(ii) *Let ρ be an irreducible K-adapted matrix representation of $H \times K$. Then there are irreducible matrix representations, ρ_H, ρ_K, of H and K respectively such that $\rho = \rho_H \otimes \rho_K$, as matrix representations, and hence ρ is also H-adapted.*

(iii) *Let $\mathcal{R}$ be a complete set of irreducible representations of $H \times K$. If $\mathcal{R}$ is both H-adapted and K-adapted then there are sets, $\mathcal{R}_H$, $\mathcal{R}_K$, of irreducible matrix representations of H and K respectively, such that $\mathcal{R} = \mathcal{R}_H \otimes \mathcal{R}_K$, as sets of matrix representations.*

(iv) *Let $\mathcal{R}$ be a set of irreducible matrix representations of a finite group G with center Z. Then $\mathcal{R}$ is Z-adapted. Therefore if $G = H \times K$ is a product of groups and H is abelian, then $\mathcal{R}$ is H-adapted.*

Proof. Part (i) is a result of Atkinson [2] and Karpovsky [42]. Part (ii) and its corollaries, parts (iii) and (iv), are simple consequences of Schur's lemma (cf. Lemma 4.2). □

Convention. Almost all of the results in remaining sections depend only on the adaptability of the representations and not the particular choice of adapted representation. For this reason explicit reference to a fixed $\mathcal{R}$ is often superfluous and we suppress this in much of the notation, e.g., we will write t_K for $t_K(\mathcal{R}_K)$ and $m_K(Y)$ for $m_K(Y, \mathcal{R}_K)$.

4. The main idea—Separation of variables

In this section we present the main new computational techniques for efficiently computing nonabelian Fourier transforms. We start by generalizing the approach of Section 3 to obtain the separation of variables algorithm. This algorithm reduces the computation of a sum of products to other, potentially smaller, repeated sums of products. Special cases of the separation of variables algorithm also occur in the work of Clausen on the symmetric group, (cf. [19], esp. Section 10). We then give a detailed analysis of the complexity of matrix multiplication when the matrices have a special structure related to a subgroup-adapted representation. These results on matrix multiplication produce the bulk of the new computational savings presented in this paper. The key idea here is that if representations are adapted to a subgroup, then any element in the centralizer of this subgroup is, by Schur's Lemma, guaranteed to have a sparse representation matrix. If coset representatives can be factored as products of such elements, then multiplication by the representation matrices of these coset representatives may be performed efficiently. When these elements are also contained in a proper subgroup of the group for which the representation remains adapted, the representation matrices are even sparser. Finally, we look at the effect of using a subgroup chain in this setting and present some general results on the complexities of our algorithms.

The separation of variables idea also arises naturally in the compact group setting [48]. The use of factorizations into generalized Euler angles in conjunction with adapted representations and Schur's lemma is a standard technique for obtaining matrix coefficients of the classical Lie groups [45]. This is the original purpose for which adapted sets of representations, in the guise of Gel'fand-Tsetlin bases, were invented [33].

4.1. **Sums of products — the separations of variables idea.** Let G be a finite group, Y a subset of G, ρ a matrix representation of G, and for each $y \in Y$, let $F(y)$ be a $d_\rho \times d_\rho$ matrix. In this section we focus on a method for computing sums of the form

$$\sum_{y \in Y} \rho(y) F(y). \tag{12}$$

This is a general setting which encompasses the algorithmic issues which we treat in this paper. For example, if we take $Y = G$ and $F(y) = f(y) \cdot I_{d_\rho}$,[1] for some complex-valued function, f on G, then the sum (12) is $\hat{f}(\rho)$. If we let Y be a set of coset representatives of a subgroup, $H < G$, and $F(y) = \widehat{f_y}(\rho \downarrow H)$, where $f_y(h) = f(y \cdot h)$ for $y \in Y$ and $h \in H$, then we are precisely in the setting of Theorem 3.1. Thus, the results of this section may be applied both directly to the computation of Fourier transforms and indirectly in conjunction with the methods of Section 3.

We shall now define an algorithm for computing (12), which we call the **separation of variables algorithm**. Let S be a subset of G which contains the identity element e, and such that any element of Y may be written as a product of elements of S. For any word w in the elements of S, let $\bar{w}$ denote the element of G obtained by multiplying out the formal product for w. For any set of words W let $\bar{W}$ denote the corresponding set of elements of G.

Choose a set of words X in the elements of S, such that $|X| = |Y|$ and $\bar{X} = Y$. Thus, the words in X may be thought of as a choice of factorization of each element of Y in terms of S. Let γ be the maximum length of any word in X. Let X_0 be the set of words, all of the same length γ, obtained from X by "padding" on the left with the identity if necessary. For each i with $0 \leq i \leq \gamma$, define X_i to be the set of subwords of X_0 obtained by removing the rightmost i symbols from each word of X_0. Note that X_i is a set of words of length $\gamma - i$ in S. For w in X_0 let $F_0(w) = F(\bar{w})$.

The separation of variables algorithm proceeds in γ steps, computing for each i from 1 to γ the recursively defined matrix-valued functions F_i on X_i,

$$F_i(w) = \sum_{s \in S, ws \in X_{i-1}} \rho(s) F_{i-1}(ws) \tag{13}$$

for any w in X_i. The algorithm completes by computing F_γ, which is, by the following lemma, the constant function whose value is the sum (12), with domain X_γ consisting of only the empty word.

Lemma 4.1. *For any set of words X such that $|X| = |Y|$ and $\bar{X} = Y$, the separation of variables algorithm described above computes $\sum_{y \in Y} \rho(y)F(y)$. I.e., with all notation as above*

$$F_\gamma = \sum_{y \in Y} \rho(y) F(y).$$

Proof. We show by induction that for $0 \leq i \leq \gamma$,

$$\sum_{w \in X_i} \rho(\bar{w}) F_i(w) = \sum_{y \in Y} \rho(y) F(y). \tag{14}$$

[1]For any positive integer d, I_d will denote the $d \times d$ identity matrix.

To start, note that (14) holds for $i = 0$ by the definition of X_0 and F_0. Now let $1 \leq i \leq \gamma$, and assume the induction hypothesis for $i-1$. Then by (13)

$$\begin{aligned}\sum_{w\in X_i}\rho(\bar{w})F_i(w) &= \sum_{w\in X_i}\rho(\bar{w})\left[\sum_{s\in S; ws\in X_{i-1}}\rho(s)F_{i-1}(ws)\right]\\ &= \sum_{s\in S; ws\in X_{i-1}}\rho(\bar{w}s)F_{i-1}(ws)\\ &= \sum_{v\in X_{i-1}}\rho(\bar{v})F_{i-1}(v).\end{aligned}$$

When $i = \gamma$ the only word in X_γ is the empty word, and $F_\gamma = \rho(e)F_\gamma$. This proves the lemma. □

The expression (13) shows the recursive nature of the separation of variables approach, as this sum may be rewritten in the same form as the original problem (12): By writing

$$F_i(s_\gamma \cdots s_i) = \sum_{s\in X_{i-1}(s_\gamma\cdots s_i)}\rho(s)F_{i-1}(s_\gamma \ldots s_i \cdot s) \tag{15}$$

where $X_{i-1}(s_\gamma \cdots s_i) = \{s \in S : s_\gamma \cdots s_i s \in X_{i-1}\}$, we reduce the original problem to γ subproblems of the same form. Hence we may apply the separation of variables algorithm to any of these subproblems, provided we first choose a finer factorization of the elements $X_{i-1}(s_\gamma \cdots s_i)$. The separation of variables algorithm is the "divide" portion of a divide and conquer strategy for computing Fourier transforms; it reduces the computation of sums of products to the computation of other sums of products. Its construction only relies on having chosen factorizations for elements of the set Y. On the other hand, the "conquer" part of our strategy, which we treat in Section 4.2, uses subgroup chains and adapted bases.

It is easy to see how the separation of variables algorithm leads to the results of Section 3. The first observation is that the separation of variables algorithm may be applied simultaneously to a whole set $\mathcal{R}$ of representations, simply by considering the direct sum $\Delta = \bigoplus_{\rho\in\mathcal{R}}\rho$. Fix a subgroup $H < G$ and a set Z of coset representatives for G/H. To apply Lemma 4.1 let $Y = G$, and for any $y \in Y$, let $F(y) = f(y) \cdot I_{d_\Delta}$. Let X be the set of all words $z \cdot h$ of length two with $z \in Z$ and $h \in H$. Then for $z \in Z$ we have $X_0(z) = H$, $X_1 = Z$, and

$$F_1(z) = \sum_{h\in H}\Delta(h)f(z\cdot h) = \widehat{f}_z(\Delta\downarrow H). \tag{16}$$

When $i = 2$ we obtain

$$\bigoplus_{\rho\in\mathcal{R}}\hat{f}(\rho) = \hat{f}(\Delta) = F_2 = \sum_{z\in Z}\rho(z)F_1(z) \tag{17}$$

and the separation of variables algorithm for computing $\hat{f}(\Delta)$ is exactly the algorithm considered in Section 3.

Separation of variables may be applied to the computation of both of the sums (16) and (17) by using factorizations of elements of H and of elements of Z respectively. The resulting composite algorithm is precisely the separation of variables algorithm for the set of words obtained by taking pairwise products of the padded words (i.e., the elements of X_0) used in both the algorithms for computing (16) and

(17). This is a general property of the separation of variables technique; using it recursively is equivalent to using a single algorithm for a different set of words.

The applications of Section 5 will always proceed by using coset representatives to obtain a coarse factorization of group elements and then refining this factorization by factoring the coset representatives themselves.

Example. It is instructive to see how the separation of variables algorithm works on a simple example. Let G be the symmetric group S_3, and suppose that $Y = G = S_3$. Let $S = \{e, t_2, t_3\}$, where $t_2 = (1\ 2)$, $t_3 = (2\ 3)$ are transpositions, and let X be the set of words

$$X = X_0 = \{eee, et_3e, t_2t_3e, eet_2, et_3t_2, t_2t_3t_2\}.$$

Then $X_1 = \{ee, et_3, t_2t_3\}$, $X_2 = \{e, t_2\}$, and $X_3 = \{\phi\}$, where ϕ denotes the empty word. The quantities $F_i(w)$ computed at each stage of the algorithm are

$$\begin{aligned}
F_1(ee) &= \rho(e)F_0(eee) + \rho(t_2)F_0(eet_2),\\
F_1(et_3) &= \rho(e)F_0(et_3e) + \rho(t_2)F_0(et_3t_2),\\
F_1(t_2t_3) &= \rho(e)F_0(t_2t_3e) + \rho(t_2)F_0(t_2t_3t_2),\\
\\
F_2(e) &= \rho(e)F_1(ee) + \rho(t_3)F_1(et_3),\\
F_2(t_2) &= \rho(t_3)F_1(t_2t_3),\\
\\
F_3(\phi) &= \rho(e)F_2(e) + \rho(t_2)F_2(t_2).
\end{aligned}$$

4.2. **Products of pairs of matrices.** The results introduced in Sections 3 and 4.1 have focused on rewriting the Fourier transform as a recursively structured summation of matrix products. This is the "divide" component of our divide and conquer strategy. In this section we consider conditions that will ensure that a matrix product involving $\rho(a)$ for a representation ρ and element a of G may be computed efficiently. This is the "conquer" portion of our divide and conquer strategy.

The main tool we use is a form of Schur's Lemma. This simple result pins down the structure of intertwining matrices for a given matrix representation.

Lemma 4.2 (Schur's Lemma; see, e.g., [58], p. 13)**.** *Let K be a subgroup of G and ρ a K-adapted representation of G such that $\rho = \eta_1 \oplus \cdots \oplus \eta_1 \oplus \cdots \oplus \eta_r \oplus \cdots \oplus \eta_r$ where $\eta_1, \ldots, \eta_r$ are inequivalent irreducible matrix representations of K, and η_i occurs with multiplicity m_i. Then the centralizer of the collection of matrices $\rho(K)$ is*

$$\left(\mathrm{Mat}_{m_1}(\mathbf{C}) \otimes I_{d_{\eta_1}}\right) \oplus \cdots \oplus \left(\mathrm{Mat}_{m_r}(\mathbf{C}) \otimes I_{d_{\eta_r}}\right) \tag{18}$$

where I_k denotes the $k \times k$ identity matrix, $\otimes$ the usual tensor product of matrices, $\mathrm{Mat}_n(\mathbf{C})$ is the algebra of $n \times n$ complex matrices, and d_{η_i} is the dimension of η_i.

If $a \in G$ is in the centralizer of a subgroup K, then its representation matrix, $\rho(a)$, is in the centralizer of $\rho(K)$. If ρ is a K-adapted representation, then $\rho(a)$ has the form (18) after some fixed permutation of rows and columns. We interpret this as saying that the matrix $\rho(a)$ is sparse and as such can be multiplied efficiently against an arbitrary $d_\rho \times d_\rho$ matrix.

Corollary 4.3. *Let all notation be as in Lemma 4.2, and let a be a group element lying in the centralizer of K. Then for an arbitrary* $d_\rho \times d_\rho$ *matrix F, the product* $\rho(a)F$ *can be computed in at most* $d_\rho \left(\sum_i d_{\eta_i} m_i^2\right)$ *operations.*

Proof. The bound comes from considering the number of nonzero entries of the matrix $\rho(a)$. There are at most $\sum_i d_{\eta_i} m_i^2$ nonzero entries and each nonzero entry occurs at most d_ρ times in the expressions making up the entries of the matrix product $\rho(a)F$. □

When a is in a proper subgroup of G that contains K, Corollary 4.3 can be improved. To explain, let $H \geq K$ and let ρ and η be representations of H and K respectively. Define

(19) $$\mathcal{M}(\rho, \eta) = \text{the multiplicity of } \eta \text{ in } \rho\!\downarrow\! K.$$

Also define

(20) $$\mathcal{M}(H, K) = \max_{\rho,\eta} \mathcal{M}(\rho, \eta)$$

as ρ and η run over complete sets of irreducible representations of H and K respectively.

Corollary 4.4. *Let* $H \geq K$ *be subgroups of G,* $\mathcal{R}$ *a complete set of irreducible representations of G adapted to the chain* $G \geq H \geq K$*, and suppose that for each* ρ *in* $\mathcal{R}$*,* $F(\rho)$ *is a* $d_\rho \times d_\rho$ *matrix. Let a be in the centralizer of K in H. Then the set of matrix products* $\{\rho(a) \cdot F(\rho) | \rho \in \mathcal{R}\}$ *may be computed in at most* $|G| \cdot \mathcal{M}(H, K)$ *operations.*

Proof. For any ρ in $\mathcal{R}$, $\mathcal{M}(H, K)$ is an upper bound for the number of nonzero entries in any column of $\rho(a)$. Hence the number of operations needed to compute any entry of the matrix $\rho(a) \cdot F(\rho)$ is bounded by $\mathcal{M}(H, K)$. There are d_ρ^2 such entries so the computation of this matrix product takes $\mathcal{M}(H, K) d_\rho^2$ operations. Summing over all representations and using the relation $\sum_{\rho \in \mathcal{R}} d_\rho^2 = |G|$ gives the result. □

For most purposes the upper bounds of Corollaries 4.3 and 4.4 are all we require to get good bounds for group complexity. However, in some situations a more detailed analysis of the matrix multiplications is necessary. We shall now consider the multiplication of two matrices which are block diagonal according to some subgroup restrictions and also have the block scalar form (18), though possibly for different subgroups.

To state these results, let $G \geq H \geq K$, and let ρ be a representation of G adapted to this chain. We introduce the notation

$$\mathrm{End}_K(\rho\!\downarrow\! H) = \mathrm{span}_{\mathbf{C}} \left(\rho(H)\right) \cap \mathrm{Centralizer}\left(\rho(K)\right),$$

so $\mathrm{End}_K(\rho \downarrow H)$ is the algebra of matrices with block diagonal form according to $\rho\!\downarrow\! H$ that also have the form (18) up to a fixed permutation of rows and columns. In particular, if $a \in H$ is in the centralizer of K, then $\rho(a)$ is in $\mathrm{End}_K(\rho\!\downarrow\! H)$.

Suppose $F_1 \in \mathrm{End}_{K_1}(\rho \downarrow H_1)$ and $F_2 \in \mathrm{End}_{K_2}(\rho \downarrow H_2)$, where the subgroup chains $H_1 \geq K_1$ and $H_2 \geq K_2$ both occur as subchains of some fixed subgroup chain of G for which ρ is adapted. We wish to examine the complexity of the matrix multiplication $F_1 \cdot F_2$. There are a number of special cases to consider corresponding to the different possible orderings of the subgroups H_1, K_1, H_2, K_2, in the subgroup chain. By exchanging F_1 and F_2 the number of cases under consideration is reduced from six to three. We shall consider one of these cases in detail and then indicate the adaptations needed to treat the other two.

Theorem 4.5. *Let $H_1 \geq H_2 \geq K_1 \geq K_2$ be a chain of subgroups of G, and let ρ be a representation of G adapted to this chain. Suppose that for $i = 1, 2$, $F_i \in \mathrm{End}_{K_i}(\rho \downarrow H_i)$. Then the matrix multiplication $F_1 \cdot F_2$ can be computed in no more than*

$$\sum_{\rho_{H_1}, \rho_{H_2}, \rho_{K_1}, \rho_{K_2}} \mathcal{M}(\rho_{H_1}, \rho_{K_1}) \mathcal{M}(\rho_{H_2}, \rho_{K_2}) \mathcal{M}(\rho_{H_1}, \rho_{H_2}) \mathcal{M}(\rho_{H_2}, \rho_{K_1}) \mathcal{M}(\rho_{K_1}, \rho_{K_2}) \tag{21}$$

scalar operations, where for $L \in \{H_1, H_2, K_1, K_2\}$, the index ρ_L ranges over all irreducible representations of the subgroup L (up to equivalence) having nonzero multiplicity in $\rho \downarrow L$.

Proof. Both matrices F_1 and F_2 belong to $\mathrm{End}_1(\rho \downarrow H_1)$ and are therefore block diagonal with blocks corresponding to the restriction of ρ to H_1. By considering the matrix multiplication one block at a time we may restrict ourselves to the case where $H_1 = G$ and ρ is an irreducible representation of G. From now on we let $G_1 = G = H_1$, $G_2 = H_2$, $G_3 = K_1$ and $G_4 = K_2$.

In this situation it is useful to index the rows or columns of the chain-adapted representation ρ by a 7-tuple, $\Lambda = (\lambda_2, \rho_2, \lambda_3, \rho_3, \lambda_4, \rho_4, \lambda_5)$, where for $i = 2, 3$, or 4, ρ_i is an irreducible representation of G_i occurring as a matrix direct summand of $\rho_{i-1} \downarrow G_i$, where $\rho_1 = \rho$ and λ_i is a variable indexing the particular occurrences of ρ_i as a matrix direct summand of ρ_{i-1}. When $i = 5$, λ_5 simply indexes a basis of ρ_4 of dimension $\mathcal{M}(\rho_4, 1)$, where 1 is the trivial representation of the trivial group. Said differently, λ_2 indexes the blocks of $\rho \downarrow G_2$ which contain copies of ρ_2, λ_3 indexes the blocks of $\rho_2 \downarrow G_3$ which contain copies of ρ_3, and λ_4 indexes the blocks of $\rho_3 \downarrow G_4$ which contain copies of ρ_4. By Lemma 4.2 the entries of a matrix $F_1 \in \mathrm{End}_{G_3}(\rho \downarrow G)$ have the form

$$[F_1]_{\Lambda, \Lambda'} = f_1(\lambda_2, \lambda_2', \rho_2, \rho_2', \lambda_3, \lambda_3', \rho_3) \cdot \delta_{\rho_3, \rho_3'} \delta_{\lambda_4, \lambda_4'} \delta_{\rho_4, \rho_4'} \delta_{\lambda_5, \lambda_5'}$$

for some complex-valued function f_1, and the entries of a matrix F_2 in $\mathrm{End}_{G_4}(\rho \downarrow G_2)$ have the form

$$[F_2]_{\Lambda, \Lambda'} = \delta_{\lambda_2, \lambda_2'} \delta_{\rho_2, \rho_2'} \cdot f_2(\rho_2, \lambda_3, \lambda_3', \rho_3, \rho_3', \lambda_4, \lambda_4', \rho_4) \cdot \delta_{\rho_4, \rho_4'} \delta_{\lambda_5, \lambda_5'}$$

for some complex-valued function f_2. Therefore, the expression for the matrix product entry $[F_1 \cdot F_2]_{\Lambda, \Lambda'}$ is

$$\delta_{\rho_4, \rho_4'} \delta_{\lambda_5, \lambda_5'} \cdot \sum_{\lambda_3''} f_1(\lambda_2, \lambda_2', \rho_2, \rho_2', \lambda_3, \lambda_3'', \rho_3) \cdot f_2(\rho_2', \lambda_3'', \lambda_3', \rho_3, \rho_3', \lambda_4, \lambda_4'). \tag{22}$$

The variables appearing in the expression (22) range over values according to Diagram 1.

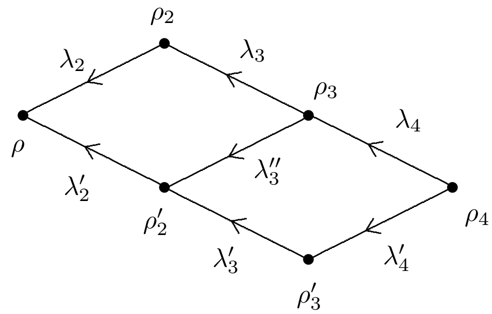

Diagram 1

In Diagram 1 a directed edge $\beta \overset{\lambda}{\longleftarrow} \alpha$ indicates that β is an irreducible representation which occurs as a matrix direct summand of the restriction of α, and that λ is a variable indexing the copy of β in this restriction. The number of operations required to compute the matrix product $F_1 \cdot F_2$ is then bounded by the number of distinct ways of assigning values to $\lambda_2, \lambda_2', \rho_2, \rho_2', \lambda_3, \lambda_3', \lambda_3'', \rho_3, \rho_3', \lambda_4, \lambda_4', \rho_4$ consistent with the conditions represented by Diagram 1.

To count this number, fix the three representations ρ_2', ρ_3, and ρ_4, and count the number of ways, if any, that the remaining variables may be assigned values in a manner consistent with the diagram. These variables may be collected into five sets corresponding to the five edges in Diagram 2; each set consists of the variables that label the path in Diagram 1 corresponding to the edge in Diagram 2.

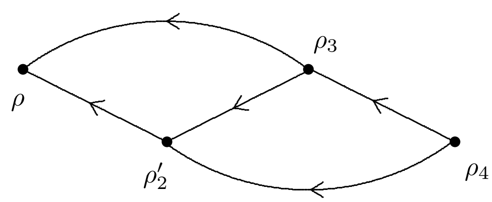

Diagram 2

These sets of variables are $\{\lambda_2, \rho_2, \lambda_3\}$, $\{\lambda_2'\}$, $\{\lambda_3''\}$, $\{\lambda_4\}$, and $\{\lambda_3', \rho_3', \lambda_4'\}$. For a given choice of ρ_2', ρ_3, and ρ_4, the choices of values for variables in different sets are completely independent. For example, the choice of λ_2' is independent of the choice of λ_3''. Now consider the set of variables, $\{\lambda_2, \rho_2, \lambda_3\}$ which corresponds to the edge from ρ_3 to ρ in Diagram 2. Each different way of choosing values of these three variables, consistent with Diagram 1, corresponds to a choice of a copy of ρ_3 appearing as a matrix direct summand of the restriction of ρ to G_3, and hence there are $\mathcal{M}(\rho, \rho_3)$ possible choices. Similarly, the number of ways of choosing values for variables in the set corresponding to any directed edge from a vertex β to a vertex α in Diagram 2 is $\mathcal{M}(\alpha, \beta)$. Therefore the total number of ways of assigning values to all variables in Diagram 1 is

$$\sum_{\rho_2', \rho_3, \rho_4} \prod_{\alpha \leftarrow \beta} \mathcal{M}(\alpha, \beta) \tag{23}$$

where the product in (23) is over all directed edges in Diagram 2, and α, β denote the finishing and starting points of the edges, respectively. This is precisely

$$\sum_{\rho_2', \rho_3, \rho_4} \mathcal{M}(\rho, \rho_3)\mathcal{M}(\rho_2', \rho_4)\mathcal{M}(\rho, \rho_2')\mathcal{M}(\rho_2', \rho_3)\mathcal{M}(\rho_3, \rho_4).$$

Substituting $\rho_{H_1} = \rho$, $\rho_{H_2} = \rho_2'$, $\rho_{K_1} = \rho_3$ and $\rho_{K_2} = \rho_4$ proves the theorem. □

The two other cases we need to consider are

$$H_2 \geq H_1 \geq K_1 \geq K_2$$

and

$$H_1 \geq K_1 \geq H_2 \geq K_2.$$

Extending the proof of Theorem 4.5 to these two other cases is routine; the important difference is that other diagrams must be considered. As before, we let $\{G_i\}$ be the chain of subgroups listed in decreasing order, and let ρ_i index irreducible representations of G_i. In the case $H_2 \geq H_1 \geq K_1 \geq K_2$, Diagram 2 must be replaced by Diagram 3,

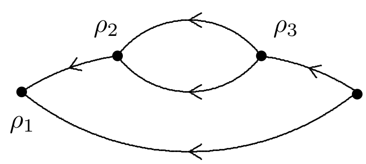

Diagram 3

but the procedure for obtaining the complexity bound from the diagram is the same: Take the product of $\mathcal{M}(\alpha, \beta)$ over all the directed edges in the diagram, and then sum over all choices of representations labeling the vertices consistent with the restriction relations represented by the diagram and with ρ_1 occurring in $\rho \downarrow G_1$. In the case $H_1 \geq K_1 \geq H_2 \geq K_2$, Diagram 2 should be replaced by Diagram 4.

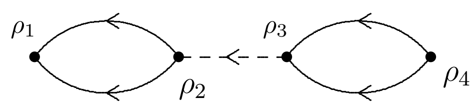

Diagram 4

The dashed edge in Diagram 4 indicates that the factor $\mathcal{M}(\rho_2, \rho_3)$ should be omitted from the product in the expression for the complexity. However, the presence of this edge indicates that ρ_3 still must occur in the restriction of ρ_2 to $H_2 (= G_3)$.

Theorem 4.6. *Let $H_1 \geq K_1$ and $H_2 \geq K_2$ be subgroups of G occurring in some chain of subgroups to which the representation ρ is adapted. Suppose that for $i = 1, 2$, F_i is a matrix in $\operatorname{End}_{K_i}(\rho \downarrow H_i)$.*

(i) *When $H_2 \geq H_1 \geq K_1 \geq K_2$, the matrix multiplication $F_1 \cdot F_2$ can be computed in no more than*

$$\sum_{\substack{\rho_{H_1}, \rho_{H_2}, \\ \rho_{K_1}, \rho_{K_2}}} \mathcal{M}(\rho_{H_2}, \rho_{K_2})\mathcal{M}(\rho_{H_2}, \rho_{H_1})\mathcal{M}(\rho_{H_1}, \rho_{K_1})^2\mathcal{M}(\rho_{K_1}, \rho_{K_2}) \tag{24}$$

scalar operations.

(ii) *When $H_1 \geq K_1 \geq H_2 \geq K_2$, the matrix multiplication $F_1 \cdot F_2$ can be computed in no more than*

$$\sum_{\rho_{H_1},\rho_{K_1},\rho_{H_2},\rho_{K_2}} \mathcal{M}(\rho_{H_1},\rho_{K_1})^2 \mathcal{M}(\rho_{H_2},\rho_{K_2})^2 \tag{25}$$

scalar operations.

For $L \in \{H_1, H_2, K_1, K_2\}$, the index ρ_L in the above sums ranges over the irreducible representations of the subgroup L, up to equivalence. In (24) *we also require that ρ_{H_2} has nonzero multiplicity in $\rho \downarrow H_2$. In* (25) *we require that ρ_{H_1} has nonzero multiplicity in $\rho \downarrow H_1$, and that ρ_{H_2} has nonzero multiplicity in $\rho_{K_1} \downarrow H_2$.*

The diagrammatic techniques introduced in the proof of Theorem 4.5 may be formalized and used to prove complexity results that are even better than those given in this paper. This approach to Fourier transforms on finite groups is explained in the sequel [50]. In particular, the appropriate setting for discussing multiplication of block scalar matrices is a tower of multi-matrix algebras (cf. [34]).

Theorems 4.5 and 4.6 give exact operation counts for the appropriate matrix multiplications. It is useful to provide some notation for these counts.

Definition 4.1. Let $H_1 \geq K_1$ and $H_2 \geq K_2$ be a chain of subgroups of G and let ρ be a representation of G. Define $C(\rho; H_1, K_1; H_2, K_2)$ to be

1. the sum (21) when $H_1 \geq H_2 \geq K_1 \geq K_2$,
2. the sum (24) when $H_2 \geq H_1 \geq K_1 \geq K_2$, and
3. the sum (25) when $H_1 \geq K_1 \geq H_2 \geq K_2$.

We extend this definition to include the three other possible arrangements of H_1, H_2, K_1, K_2 in the subgroup chain, by the symmetry condition

$$C(\rho; H_1, K_1; H_2, K_2) = C(\rho; H_2, K_2; H_1, K_1).$$

It is clear that $C(\rho; H_1, K_1; H_2, K_2)$ is an upper bound for the complexity of the matrix multiplication of a matrix in $\mathrm{End}_{K_1}(\rho \downarrow H_1)$ with a matrix in $\mathrm{End}_{K_2}(\rho \downarrow H_2)$, whatever the positions of H_1, K_1, H_2, K_2 in the subgroup chain. The next theorem gives another useful bound.

Theorem 4.7. *Let $H_1 \geq K_1$ and $H_2 \geq K_2$ be subgroups of G occurring as subchains of some chain of subgroups to which the representation ρ is adapted. Let $G_1 \geq G_2 \geq G_3 \geq G_4$ be the rearrangement of the H_i and K_i into a single chain. Then*

$$C(\rho; H_1, K_1; H_2, K_2) \leq \mathcal{M}(G_2, G_3). \sum_{\rho_{G_1},\rho_{G_4}} \mathcal{M}(\rho_{G_1}, \rho_{G_4})^2$$

where ρ_{G_1} ranges over inequivalent irreducible representations of G_1 having nonzero multiplicity in ρ, and ρ_{G_4} ranges over inequivalent irreducible representations of G_4 having nonzero multiplicity in ρ_{G_1}.

Proof. For simplicity we only consider the case when $H_1 \geq H_2 \geq K_1 \geq K_2$; all the other cases use a similar line of proof. First note that if ρ is a representation of G, ρ_K is a representation of a subgroup of G, and H is a subgroup of G containing K, then

$$\mathcal{M}(\rho, \rho_K) = \sum_{\rho_H} \mathcal{M}(\rho, \rho_H)\mathcal{M}(\rho_H, \rho_K).$$

By Theorem 4.6 we may bound $C(\rho; H_1, K_1; H_2, K_2)$ as follows:

$$\begin{aligned}
C(\rho; H_1,&K_1; H_2, K_2) \\
&= \sum_{\rho_{H_1},\rho_{H_2},\rho_{K_1},\rho_{K_2}} \mathcal{M}(\rho_{H_1},\rho_{K_1})\mathcal{M}(\rho_{H_2},\rho_{K_2}) \times \\
&\qquad\qquad \times \mathcal{M}(\rho_{H_1},\rho_{H_2})\mathcal{M}(\rho_{H_2},\rho_{K_1})\mathcal{M}(\rho_{K_1},\rho_{K_2}) \\
&\le \mathcal{M}(H_2,K_1) \sum_{\rho_{H_1},\rho_{K_2}} \left(\sum_{\rho_{H_2}} \mathcal{M}(\rho_{H_1},\rho_{H_2})\mathcal{M}(\rho_{H_2},\rho_{K_2})\right) \times \\
&\qquad\qquad \times \left(\sum_{\rho_{K_1}} \mathcal{M}(\rho_{H_1},\rho_{K_1})\mathcal{M}(\rho_{K_1},\rho_{K_2})\right) \\
&= \mathcal{M}(H_2,K_1) \sum_{\rho_{H_1},\rho_{K_2}} \mathcal{M}(\rho_{H_1},\rho_{K_2})^2
\end{aligned}$$

□

4.3. **Complexity of the algorithm.** We now combine the ideas of Sections 4.1 and 4.2 to obtain some general upper bounds for the complexity of a Fourier transform. Assume all notation is as in Section 4.1, so that for a fixed subset $Y \subset G$, X is a set of words from a subset $S \subset G$, whose products equal Y, X_0 is obtained by padding the words of X with copies of the identity element on the left until they all have the same length γ, and X_k is obtained from X_0 by deleting k symbols from the right of each word. Furthermore, let X^i denote the set of words obtained from words of X_0 by deleting the $\gamma - i$ leftmost symbols.

Let $K_n \ge \cdots \ge K_0 = 1$ be a chain of subgroups of G, and assume that ρ is adapted to this chain. Given any $g \in G$, define the indices $c^+(g)$ and $c^-(g)$ by

(26)
$$\begin{aligned}
K_{c^+(g)} &= \text{the smallest subgroup in the chain containing } g \text{, and} \\
K_{c^-(g)} &= \text{the largest subgroup of } K_{c^+(g)} \text{ in the chain and centralizing } g.
\end{aligned}$$

So $\rho(g) \in \mathrm{End}_{K_{c^+(g)}}(\rho{\downarrow}K_{c^-(g)})$. Let b_0^+ and b_0^- be such that

$$F(y) \in \mathrm{End}_{K_{b_0^+}}(\rho{\downarrow}K_{b_0^-})$$

for each $y \in Y$. Then for any i between 0 and γ we let

$$\begin{aligned}
b_i^+ &= \max\{b_0^+, c^+(g) : g \in X^i\}, \\
b_i^- &= \min\{b_0^-, c^-(g) : g \in X^i\}.
\end{aligned}$$

By Definition 4.1 the number of scalar operations needed to perform the matrix product $\rho(s) \cdot F_{i-1}(w)$ appearing in the expression (13) is no greater than $C(\rho; K_{c^+(s)}, K_{c^-(s)}; K_{b_{i-1}^+}, K_{b_{i-1}^-})$.

Theorem 4.8. *Let ρ be a matrix representation of G which is adapted to a chain of subgroups, $K_n \ge \cdots \ge K_0$. Let $Y \subset G$ and for each $y \in Y$ let $F(y)$ be in*

$\mathrm{End}_{K_{b_0^+}}(\rho\!\downarrow\! K_{b_0^-})$. *Then the sum* (12) *may be calculated in no more than*

$$\sum_{k=0}^{\gamma-1} \sum_{\substack{ws\in X_k \\ s\neq e}} C(\rho; K_{c^+(s)}, K_{c^-(s)}; K_{b_k^+}, K_{b_k^-}) \tag{27}$$

scalar operations.

Proof. By the definition of C (Definition 4.1) and Theorem 4.5, the sum (27) is an upper bound for the number of scalar operations needed for all the matrix multiplications occurring in the separation of variables algorithm. We now have to include the matrix additions as well. The proof of Theorem 4.5 shows that for each nonzero entry of the matrix products, the number of scalar multiplications used in the computation of that entry is one more than the number of scalar additions. When we include the scalar additions used to compute the matrix additions occurring in our algorithm we see that the total number of additions used is still no greater than the total number of multiplications. Hence the complexity of the algorithm is the same as the total number of multiplications and is bounded by the sum (27). □

We now give several simpler bounds that are direct corollaries of Theorem 4.8 and the results of Section 4.2. For this, it is useful to introduce the **multiplicity function** $\mathcal{M}$, defined on G. For a fixed chain of subgroups $G \geq K_n \geq \cdots \geq K_0$ define

$$\mathcal{M}(g) = \begin{cases} \mathcal{M}(K_{c^+(g)}, K_{c^-(g)}) & \text{if } g \neq 1, \\ 0 & \text{if } g = 1. \end{cases} \tag{28}$$

For any subset $S \subset G$ define

$$\mathcal{M}(S) = \max_{s\in S} \mathcal{M}(s). \tag{29}$$

Corollary 4.9. *Let $\mathcal{R}$ be a complete set of inequivalent irreducible matrix representations for G, adapted to the subgroup chain $K_n \geq \cdots \geq K_0$. Let Y be any subset of G, and let X be a set of factorizations of elements of Y in terms of elements from a subset $S \subset G$. Let γ be the maximum length of any word in X. Then*

$$m_G(\mathcal{R}, Y) \leq \sum_{k=0}^{\gamma-1} \sum_{ws\in \tilde{X}_k} \mathcal{M}(s) \tag{30}$$

$$\leq \mathcal{M}(S)\left[\sum_{k=0}^{\gamma-1} \left|\tilde{X}_k\right|\right] \tag{31}$$

where $\tilde{X}_k$ is obtained from X_0 by deleting k elements from the right of each word and then deleting all words with an identity element at the far right.

Proof. This is an immediate consequence of Theorem 4.8, Theorem 4.7 and the definition of $\mathcal{M}$. □

Remarks. 1. *Applications.* Corollary 4.9 is the primary result for the applications of Section 5. It has the virtue of simplicity, but when $\mathcal{R}$ is H-adapted, it does not use the block diagonal form of the Fourier transforms of the restricted representations on H. To take this into account, Theorem 4.8 must be used directly.

2. *General results.* Corollary 4.9 might be useful in the search for general results on the complexity of Fourier transforms on any finite group, possibly improving on the general bounds of Clausen [18], or those of Diaconis and Rockmore [29]. As a first step in this direction, let $l_{H,S}(y)$ be the minimum non-negative integer, l, such that y is in the same coset of G/H as some product of l elements of S. If Y is a set of coset representatives for G/H, then the generating function $\mathcal{P}_{G/H,S}(t) = \sum_{y \in Y} t^{l_{H,}(y)}$ is independent of the choice of coset representatives and is sometimes called the **Poincaré polynomial** of G/H with respect to S. Note that in this case $\mathcal{P}'_{G/H,S}(1) = \sum_{y \in Y} l_{H,S}(y)$ is the sum of the lengths of minimal coset representatives for G/H.

Corollary 4.10. *Let $\mathcal{R}$ be a complete set of inequivalent irreducible matrix representations for G, adapted to the subgroup chain $K_n \geq \cdots \geq K_0$. Let H be a subgroup of G, and Y a set of minimal coset representatives for G/H, relative to the subset, S of G. Then*

$$\begin{aligned} m_G(\mathcal{R}, Y) &\leq \mathcal{M}(S) \cdot \mathcal{P}'_{G/H,S}(1) \\ &\leq \mathcal{M}(S) \cdot \gamma \cdot |G/H| \end{aligned}$$

where γ is the maximum length of any element of Y in S.

Notice that this is a general upper bound, depending only on a set of generators for a finite group, and a subgroup chain.

3. *Adapted diameters.* In order to use Corollary 4.10 in conjunction with Theorem 3.2, we must assume that given a chain of subgroups

$$G = K_m \geq \cdots \geq K_0, \tag{32}$$

then for each i, a set of coset representatives for K_i/K_{i-1} can be expressed in terms of $S \cap K_i$. In this case we say that S is a **generating set for the chain of subgroups** (32). When the subgroup chain contains both the whole group, G, and the trivial subgroup, 1, a generating set for the chain is called a **strong generating set** for G with respect to the chain of subgroups (32). Strong generating sets arise naturally in the context of many algorithmic issues in computational group theory [59]. In particular, fast algorithms for their construction for stabilizer subgroup chains in permutation groups are a cornerstone for many important techniques [5].

Using the bounds of Corollary 4.10 in Theorem 3.2, we obtain an upper bound on the complexity of G in terms of the quotient sizes $|K_i/K_{i-1}|$, multiplicity data $\mathcal{M}(S)$ and combinatorial data in the form of the maximum lengths needed to construct the coset representatives at each level. This last aspect is nicely encapsulated in the notion of the **adapted diameter** of a group with respect to a generating set for a given chain of subgroups (cf. [51] for details).

4. *Choosing the generating set or subgroup chain.* The complexity bounds of Theorem 4.8, Corollary 4.9, and Theorem 3.1 only depend on the choice of subgroup chain and on the choice of factorization for group elements. Thus, they do not depend on the choice of a particular adapted basis. We now discuss some ideas which guide these choices with the aim of minimizing the complexity (27) of the separation of variables algorithm. These issues are examined a bit more fully in [49] (esp. Section 3.2).

For a fixed factorization, refining the subgroup chain always decreases the bound (27). This is because the complexity for the matrix product, $C(\rho; H_1, L_1; H_2, L_2)$, is decreased if we increase L_1 or L_2 or if we decrease H_1 or H_2. Refining the subgroup chain therefore decreases $C(\rho; K_{c^+(s)}, K_{c^-(s)}; K_{b^+_{i-1}}, K_{b^-_{i-1}})$. Of course, this is also changing the original problem, as we must assume our representations are adapted to the new subgroup chain, so that Theorem 4.9 applies; this is an additional hypothesis.

It is conceivable that for a given group, a natural chain of subgroups may be given; in this case we are faced with the problem of finding a factorization of group elements that makes the separation of variables algorithm efficient. If we plan to apply separation of variables recursively through the chain,

$$G = K_n \geq K_{n-1} \geq \cdots \geq K_0 = 1, \tag{33}$$

then the factorization we use must be a refinement of a factorization using coset representatives, and the set of generators S, for the factorization is necessarily a strong generating set (cf. Remark 3).

We now construct a strong generating set with minimal $\mathcal{M}(S)$. For any subgroups $H \geq L$ in the subgroup chain, this set will also minimize the quantity

$$\max_{s \in S} C(\rho; K_{c^+(s)}, K_{c^-(s)}; H, L) \tag{34}$$

over all strong generating sets for the chain (33). We start by defining $S(0) = K_0 = 1$, which clearly solves this problem for the trivial group. Then we define $S(i) = S \cap K_i$ inductively by

$$S(i) = S(i-1) \cup (K_i \cap \text{Centralizer}(K_j))$$

where j is chosen to be maximal with respect to the property that $S(i)$ generates K_i. By induction, $S(i)$ is a strong generating set for the chain $K_i \geq \cdots \geq K_0$ and $S = S(n)$ minimizes both $\mathcal{M}(S)$ and (34) amongst strong generating sets. Note that we do not need to calculate any restriction multiplicities to find this generating set.

Minimizing $\mathcal{M}(S)$ places a restriction on the generators which increases the lengths of factorizations. In practice it seems that the advantage of smaller multiplicities outweighs the disadvantage of long factorizations. The possibility of minimizing $\mathcal{M}(S)$ is one of the most important features of our approach to the computation of Fourier transforms; it gives us a place to start the construction of an algorithm for a specific group we might be interested in.

The converse problem is to construct a subgroup chain from a generating set so the complexity of the separation of variables algorithm is small. Suppose now, that we are given a minimal generating set, S. Then an arbitrary ordering of elements of S as $s_1, \ldots, s_n$, defines a subgroup chain via $K_i = \langle s_1, \ldots s_i \rangle$. It is clear that $c^+(s_i) = i$ for this subgroup chain. If we draw a graph with vertices corresponding to elements of S and edges between elements that do not commute then $c^-(s_i)$ can be read straight from the graph as the largest j such that s_i is not connected to any of $s_1, \ldots, s_j$ by an edge. Ordering S corresponds to labeling the vertices of this graph with numbers from 1 to n. Finding an ordering of S such that the numbers $c^+(s_i) - c^-(s_i)$ are minimized is related to the problem of drawing the graph in a form which is "close" to a chain.

5. Applications

The results of Section 4 may be immediately applied to derive useful upper bounds for the complexities of many families of finite groups. We first show how our general machinery reobtains the best known FFTs for some abelian groups, the symmetric groups and their wreath products. We then move on to derive new results for some of the families of classical groups over finite fields as well as their various generalizations.

Our usual approach is via Corollary 4.9. Thus in each situation we require a chain of subgroups with the accompanying sequence of coset representatives. For families of groups which nest naturally (e.g., symmetric groups, general linear groups) the subgroup chains contain the nesting and we get a recursive description of the algorithm. To take full advantage of Corollary 4.9 the coset representatives should admit a factorization in terms of a generating set such that the value of $\mathcal{M}$, defined by (28), on the generators is small. In what follows, the statement of the theorems will be given in terms of the complexity, T_G, defined in Definition 2.2. The proofs are most most easily presented using the reduced complexity t_G.

5.1. **Finite abelian groups.** Applications in digital signal processing and data analysis motivated the need for a fast cyclic discrete Fourier transform (cf. the example of Section 2.1 and the references [22, 23, 38]) and more generally a fast Fourier transform on any abelian group [31, 53]. Application of Corollary 4.9 immediately gives us some well-known results bounding the complexity of the Fourier transform on any finite abelian group.

Theorem 5.1. *Let A be a finite abelian group whose order has the prime factorization $|A| = p_1^{r_1} \dots p_m^{r_m}$. Then for any complete set of irreducible representations $\mathcal{R}$ of A,*

$$\mathcal{C}_A \leq T_A(\mathcal{R}) \leq |A| \sum_{i=1}^{m} r_i p_i.$$

Proof. Since A is abelian, all irreducible representations of A are one-dimensional. Thus, the unique complete set of irreducible representations is adapted with respect to any chain of subgroups of G. Let $S = A$ be the generating set for A. As all representations of A are one-dimensional, $\mathcal{M}(S) = 1$ with respect to any chain of subgroups. Let $A = K_n > \cdots > K_0 = \{1\}$ be any chain of subgroups of A. For a fixed i let Y_i be any complete set of coset representatives for K_i/K_{i-1} and let $X = Y_i$ be the set of trivial factorizations of elements of Y_i (i.e., each element in Y_i is represented by the one element word consisting of itself). Clearly, $X_1 = \{\phi\}$ (where ϕ denotes the empty word) so that $m_{K_i}(\mathcal{R}_{K_i}, Y_i) \leq |Y_i|$, by Corollary 4.9. Applying Theorem 3.2 then yields

$$t_A \leq \sum_{i=1}^{n} \frac{|K_i|}{|K_{i-1}|}. \tag{35}$$

The right-hand side of (35) is a sum of divisors of $|A|$ whose product is equal to $|A|$. Such a sum is minimized precisely when each term $|K_i| \,/\, |K_{i-1}|$ is prime. This type of chain can always be found in an abelian group and any chain of subgroups of A may be refined to such a chain. Hence the theorem is proved. □

The proof of Theorem 5.1 is essentially the derivation of the well-known Cooley-Tukey FFT [24]. Note that when $|A| = 2^n$ we find that $\mathcal{C}(A) \leq 2n \cdot 2^n = 2\,|A| \log_2 |A|$. For primes greater than 2 other techniques have been discovered for further optimizing the discrete Fourier transform (see, e.g., [31]). For any abelian group A, $\mathcal{C}(A) \leq 8\,|A| \log_2 |A|$ (cf. [9]).

5.2. **FFTs for S_n and other Weyl groups.** Applications in data analysis as well as the analysis of certain random walks related to card shuffling (cf. [27]) have motivated recent work related to FFTs for the symmetric group. For a survey of some approaches to these algorithms see [21]. In this section we show how the most efficient known algorithm due to Clausen (cf. [18]) can be rederived by our general approach and then show how our techniques extend directly to the other Weyl groups.

For the symmetric group we use the natural chain of subgroups

$$S_n > S_{n-1} > \cdots > S_1 = \{1\} \tag{36}$$

where S_k is identified with the subgroup of S_n of elements fixing each of the points $k+1, \ldots, n$. This chain has a natural generalization in the other Weyl groups.

Theorem 5.2 (Clausen [18], Theorem 1.4)**.** *Let S_n denote the symmetric group on n elements. If $\mathcal{R}$ is any complete set of irreducible representations of S_n adapted to the chain of subgroups* (36)*, then*

$$\mathcal{C}(S_n) \leq T_{S_n}(\mathcal{R}) \leq \frac{(n+1)n(n-1)}{3} \cdot n!. \tag{37}$$

Proof. Take as generating set the collection of pairwise-adjacent transpositions, $S = \{t_2, \ldots, t_n\}$, where t_j denotes the transposition $(j-1, j)$. Note that

$$\begin{cases} t_j \in S_j \text{ and} \\ t_j \text{ commutes with } S_k \text{ for } k < j-1. \end{cases}$$

Thus, in the notation of Section 4.3

$$\begin{aligned} K_{c^+(t_j)} &= S_j \quad \text{and} \\ K_{c^-(t_j)} &= S_{j-2}. \end{aligned}$$

Furthermore, it is easily derived from the combinatorics of Young tableaux and the "Branching Theorem" (cf. [40], p. 34) that the maximum multiplicity occurring in the restriction of any irreducible representation from S_k to S_{k-2} is two, i.e., $\mathcal{M}(S_k, S_{k-2}) = 2$, so that $\mathcal{M}(t_j) = 2$. Lastly, note that coset representatives for S_n/S_{n-1} of minimal length in the generating set S are given by the elements

$$\begin{aligned} Y &= \{1, t_n, t_{n-1}t_n, \ldots, t_2 \cdots t_n\} \\ &= \{1, (n\; n-1), (n-2\; n-1\; n), \ldots, (1\; \cdots\; n)\}. \end{aligned}$$

If we let X be the corresponding set of words, then the longest product in X has length $\gamma = n-1$, and in the notation of Corollary 4.9,

$$\tilde{X}_k = \{e \cdot e \cdots e \cdot t_{n-k}, \ldots, e \cdot t_3 \cdots t_{n-k}, t_2 \cdot t_3 \cdots t_{n-k}\}.$$

Therefore

$$\sum_{k=0}^{\gamma-1} \left|\tilde{X}_k\right| = \frac{n(n-1)}{2}.$$

Plugging this data into Corollary 4.9 and using Theorem 3.1 gives the recurrence $t_{S_n} \leq t_{S_{n-1}} + n(n-1)$ which is easily iterated to finish the proof. □

Remark. The bound of Theorem 5.2 is on the order of $n!(\log_2 n!)^3$. In this case the representations given by Young's orthogonal form or Young's seminormal form (cf. [40], p. 114) are adapted for the chain of subgroups (36) for S_n. The resulting algorithm is the fastest algorithm currently known for computing a Fourier transform on S_n [21].

The above discussion for S_n generalizes naturally to all Weyl groups. The pairwise adjacent transpositions are simple reflections, and the chain of subgroups (36) is the corresponding chain of parabolic subgroups. From this point of view the coset representatives we use are quite natural; they are the minimal coset representatives for the chain of parabolic subgroups, and their factorization comes from the Bruhat order by taking subwords of the unique minimal coset representative of maximal length. In this language (the book [39] is a good reference for the basic material) the generalization of Theorem 5.2 to the Weyl groups B_n and D_n is straightforward.

We shall consider the chains of parabolic subgroups

$$\begin{aligned} &B_n > B_{n-1} > \cdots, \\ &D_n > D_{n-1} > \cdots, \end{aligned} \tag{38}$$

and the generating sets consisting of the simple reflections. The minimal coset representatives with respect to (38) are well-known as are explicit expressions for the corresponding Poincaré polynomial. There are explicit formulae for the multiplicities of the restrictions of the classical Weyl groups to any parabolic subgroup in terms of the Littlewood-Richardson coefficients .

The results we obtain for the groups B_n and D_n are superseded by the results on wreath products in the next section (cf. Theorem 5.6), as $B_n \cong (\mathbf{Z}/2\mathbf{Z})[S_n]$ is a wreath product of $\mathbf{Z}/2\mathbf{Z}$ by S_n, and D_n is a subgroup of index 2 in B_n. However, the techniques used here illustrate the combinatorial methods used in our construction of FFTs on Chevalley groups (cf. Section 5.6).

Theorem 5.3. *Assume $\mathcal{R}_B$ and $\mathcal{R}_D$ are complete sets of irreducible representations of B_n and D_n respectively, each set adapted to the appropriate chain of subgroups* (38)*. Then*

$$\mathcal{C}(B_n) \leq T_{B_n}(\mathcal{R}_B) \leq \frac{(n+1)n(4n-1)}{3} \cdot |B_n| \tag{39}$$

and

$$\mathcal{C}(D_n) \leq T_{D_n}(\mathcal{R}_D) \leq \frac{4(n+1)n(n-1)}{3} \cdot |D_n|\,. \tag{40}$$

Before we prove Theorem 5.3 we state some lemmas which provide the data needed to apply Corollary 4.9 to this situation.

Lemma 5.4. (i) *The maximum multiplicity occurring in a restriction of any irreducible representation of S_n to S_{n-1}, B_n to B_{n-1}, or D_n to D_{n-1} is* 2*, i.e., $\mathcal{M}(S_n, S_{n-1}), \mathcal{M}(B_n, B_{n-1}), \mathcal{M}(D_n, D_{n-1}) \leq 2$.*

(ii) *The maximum dimension of a representation of $D_3 \cong S_4$ is* 3*.*

Proof. (i) It is well-known that the restriction of an irreducible representation of S_n to S_{n-1} is multiplicity-free (see, e.g., [40]) as is that of B_n to B_{n-1} (see, e.g., [68]). The result for D_n follows easily from that of B_n, and the fact that D_n is of index 2 in B_n.

(ii) This follows from the hooklength formula (see [40], p. 77). □

The minimal coset representatives and the sums of their lengths may be found using the following lemma.

Lemma 5.5 (cf. [39])**.** *Let $\mathfrak{W}$ be a Weyl group with S its set of simple reflections. For any subset $J \subset S$ let $\mathfrak{W}_J$ denote the corresponding parabolic subgroup. Let*

$\mathcal{P}_{\mathfrak{W}/\mathfrak{W}_J,S}(t)$ denote the Poincaré polynomial of $\mathfrak{W}/\mathfrak{W}_J$ in the variable t. Then the sum of the lengths of the minimal coset representatives of $\mathfrak{W}/\mathfrak{W}_J$ is given by

$$\mathcal{P}'_{\mathfrak{W}/\mathfrak{W}_J,S}(1) = \frac{1}{2}\,|\mathfrak{W}/\mathfrak{W}_J|\,[N_S - N_J]$$

where P' denotes the derivative with respect to t and where N_S and N_J are the numbers of reflections in $\mathfrak{W}$, $\mathfrak{W}_J$ and hence the lengths of the longest elements in $\mathfrak{W}$ and $\mathfrak{W}_J$ respectively. In addition the minimal coset representatives for $\mathfrak{W}/\mathfrak{W}_J$ and their minimal factorizations all occur as subwords of a minimal factorization for $w_S w_J$, where w_S is the longest element in $\mathfrak{W}$ and w_J is the longest word in $\mathfrak{W}_J$.

In Table 1 we summarize the data required to bound the complexities for the Weyl groups.

Table 1. Combinatorial data for the Weyl groups.

$\mathfrak{W}$	$\mathfrak{W}_J$	$\mathcal{M}(S)$	$\lvert\mathfrak{W}\rvert$	N_S	$\mathcal{P}'_{\mathfrak{W}/\mathfrak{W}_J,S}(1)$
S_n	S_{n-1}	2	$n!$	$\frac{1}{2}n(n-1)$	$\frac{1}{2}n(n-1)$
B_n	B_{n-1}	2	$2^n n!$	n^2	$n(2n-1)$
D_n	D_{n-1}	3	$2^{n-1}n!$	$n(n-1)$	$2n(n-1)$

It is now straightforward to use Table 1 to obtain recursive bounds for the reduced complexities of these chains of groups.

Proof of Theorem 5.3. From the data in Table 1, Corollary 4.10, and Theorem 3.1, we obtain the recurrences $t_{B_n} \le t_{B_{n-1}} + 2n(2n-1)$ and $t_{D_n} \le t_{D_{n-1}} + 6n(n-1)$. Iterating the recurrence for t_{B_n} gives the result for that series of groups, but for t_{D_n} we need a more careful count.

Let $s_1, \dots, s_n$ denote the simple reflections for D_n, in the order indicated in Diagram 5. Then $\mathcal{M}(s_i) = 2$ for $i \ge 4$, $\mathcal{M}(s_3) = 3$ and $\mathcal{M}(s_i) = 1$ for $i = 1$ or $i = 2$. The maximal minimal coset representative for D_n/D_{n-1} is $s_n \cdots s_3 s_2 s_1 s_3 \cdots s_n$ and the minimal coset representatives have the following minimal factorizations:

$$\begin{aligned} &1,\ s_n,\ s_{n-1}s_n,\ \dots,\ s_3\cdots s_n,\ s_2s_3\cdots s_n,\ s_1s_3\cdots s_n,\ s_1s_2s_3\cdots s_n,\\ &s_3s_1s_2s_3\cdots s_n,\ \dots,\ s_n\cdots s_3s_2s_1s_3\cdots s_n. \end{aligned} \tag{41}$$

The number of times s_3 occurs in these words is exactly equal to the number of times s_1 and s_2 occur in total, so the average value of $\mathcal{M}$ over all occurrences of symbols in the set of minimal factorizations is 2. The sum of the lengths of the minimal coset representatives of D_n/D_{n-1} is $2n(n-1)$. Therefore if we let X be equal to the set of words (41), then we have

$$\sum_{k=0}^{2n-2} \sum_{\substack{ws\in X_k\\ s\neq e}} \mathcal{M}(s) = 4n(n-1).$$

Applying Corollary 4.9 and Theorem 3.1 gives us $t_{D_n} \le t_{D_{n-1}} + 4n(n-1)$. Solving this recurrence completes the proof. □

We have already given the minimal coset representatives for both S_n/S_{n-1} and D_n/D_{n-1}. For B_n/B_{n-1} they are

$$1,\ s_n,\ s_{n-1}s_n,\ \dots,\ s_1\cdots s_n,\ s_2s_1\cdots s_n,\ \dots,\ s_n\cdots s_1\cdots s_n$$

where $s_1, \dots, s_n$ are the simple reflections of B_n labelled according to Diagram 5.

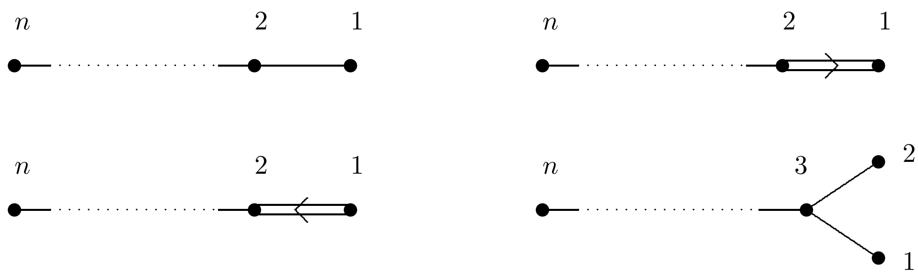

Diagram 5. Labelling the simple roots.

5.3. **Wreath products of the symmetric group.** For wreath products of the form $G[S_n]$, a decomposition similar to that used for Weyl groups is used. Wreath products are of interest in data analysis as the symmetry groups of nested designs [52] and in structural chemistry as the automorphism groups of non-rigid molecules [67]. They are also studied as the automorphism groups of graphs obtained by "composition" (cf. [36]).

Abstractly, $G[S_n]$ has the structure of a semidirect product $G^n \rtimes S_n$ in the following way. Elements of this group may be described by pairs $(f;\pi)$ where $f : \{1,\ldots,n\} \longrightarrow G$, and S_n acts on G^n by

$$f^\pi(j) = f(\pi^{-1}(j))$$

for $\pi \in S_n$ and $f \in G^n$. Multiplication is defined by

$$(f;\pi)\cdot(g;\sigma) = (f\cdot g^\pi;\pi\sigma)$$

where $f\cdot g^\pi(j) = f(j)g^\pi(j)$. In this notation it is clear that both S_n and G^n are naturally identified with subgroups of $G[S_n]$ and that under such an identification G^n is a normal subgroup and therefore $G[S_n]$ is a semidirect product of these subgroups. It is not too difficult to see that such a construction makes sense for any permutation group $H < S_n$. A thorough but accessible treatment of wreath products may be found in [44].

A slight modification of the techniques used in Section 5.2 for the symmetric group yields comparable results for their wreath products. In this case we will use the chain of subgroups

$$G[S_n] > G \times G[S_{n-1}] > G[S_{n-1}] > \cdots \tag{42}$$

where $G[S_{n-1}] < G[S_n]$ denotes the subgroup of elements $(f;\sigma)$ for which σ lies in S_{n-1} and $f(n)$ is the identity element of G.

Theorem 5.6 ([55], Theorem 3)**.** *Let $G[S_n]$ denote the wreath product of S_n by the finite group G and let d_G denote the maximum dimension of an irreducible representation of G. Let $\mathcal{R}$ be any complete set of irreducible representations of $G[S_n]$ adapted to the chain of subgroups* (42)*. Then,*

$$\mathcal{C}(G[S_n]) \le T_{G[S_n]}(\mathcal{R}) \le |G[S_n]| \left[\frac{(n+1)n(n-1)}{3}(d_G)^2 + nt_G\right].$$

Proof. Note that coset representatives for $G[S_n]/(G\times G[S_{n-1}])$ can be chosen to be the same as for S_n/S_{n-1}, so that these coset representatives can be written as words in S, the set of pairwise-adjacent transpositions in S_n. The transposition t_j lies in $G[S_j]$ and commutes with $G[S_{j-2}]$. So if we use the chain of subgroups

$$G[S_j] > G \times G[S_{j-1}] > G[S_{j-1}] > G \times G[S_{j-2}] > G[S_{j-2}]$$

and the fact that the restriction of representations from $G[S_j]$ to $G \times G[S_{j-1}]$ is multiplicity-free (see, e.g., [44]) we find that $\mathcal{M}(S)$ is $2d_G^2$, for d_G the maximum dimension of an irreducible representation of G. Using the minimal coset representatives for S_n/S_{n-1} as coset representatives for $G[S_n]/(G \times G[S_{n-1}])$ as well as Theorem 3.1, we obtain the relation

$$\begin{aligned} t_{G[S_n]} &\leq t_{G\times G[S_{n-1}]} + n(n-1)(d_G)^2 \\ &\leq t_{G[S_{n-1}]} + t_G + n(n-1)(d_G)^2. \end{aligned}$$

Applying this inequality recursively proves the theorem. □

Remark. Notice that a subgroup chain for G will give a chain of subgroups of $G[S_n]$ refining the chain (42). In [55] bases adapted to the subgroup chain (42) are constructed.

5.4. **A new FFT for the general linear group over a finite field.** Let $GL_n(q)$ denote the group of invertible $n \times n$ matrices with entries in the field of q elements where q is a prime power. For data analysis, these groups and their generalizations are of interest as the automorphism groups of the many designs based on finite geometries and codes (see, e.g., [1]). Throughout this section all matrix groups are assumed to be over $\mathbf{F}_q$, the finite field of q elements. Thus, $GL_n \equiv GL_n(q)$, etc.

To apply the results of Section 5 to these groups, we will consider the chain of subgroups

$$GL_n > P_n > GL_{n-1} \times GL_1 > GL_{n-1} > \cdots > GL_1 \tag{43}$$

where P_n is the subgroup of GL_n of all block matrices of the form

$$\left(\begin{array}{c|c} A & v \\ \hline 0 \ \dots \ 0 & c \end{array}\right) \tag{44}$$

with $A \in GL_{n-1}$, $v \in \mathbf{F}_q^{n-1}$ and $c \in \mathbf{F}_q^{\times}$; and $GL_k \times GL_1$ is identified with the subgroup of block diagonal matrices of the form $A \oplus b \oplus I_{n-k-1}$ with A in GL_k and b in GL_1 and I_r denoting the $r \times r$ identity matrix.

Theorem 5.7. *Let $\mathcal{R}$ be any complete set of irreducible representations of $GL_n(q)$ adapted to the chain of subgroups* (43)*. Then for any $n \geq 2$, $q \geq 2$,*

$$\mathcal{C}(GL_n(q)) \leq T_{GL_n(q)}(\mathcal{R}) < \frac{1}{2} 2^{2n} q^{2n-2} \left| GL_n(q) \right|. \tag{45}$$

We postpone the proof of Theorem 5.7 in order to first collect the preliminary results necessary for applying Corollary 4.9. As before, we seek generators for the successive sets of coset representatives for which the values of $\mathcal{M}$ are low.

Let $E_{i,j}$ be the matrix that is zero everywhere except for a 1 in the i,j entry. For $i \neq j$ and any x in $\mathbf{F}_q$, define $X_{i,j}(x) = I + xE_{i,j}$. Let $X^*_{i,i}(x) = I + (x-1)E_{i,i}$. Also let t_i denote the transposition matrix $E_{i-1,i} + E_{i,i-1}$. These elements generate GL_n [41], and they will serve as our generating set.

Factorizations of coset representatives of GL_n/P_n in terms of this generating set are easily derived from the Bruhat decomposition for GL_n (see [41, 15]). Those for $P_n/(GL_n \times GL_1)$ may be derived using some simple matrix algebra.

Lemma 5.8. *With the above notation,*

(i) $GL_n = \coprod_{k=1}^{n} (X_{k,k+1}t_{k+1}) \cdots (X_{n-1,n}t_n) \cdot P_n$.

(ii) $P_n = X_{n-1,n}(X_{n-2,n-1}t_{n-1}) \cdots (X_{1,2}t_2) \cdot t_3 \cdots t_{n-1} \cdot X_{n-1,n}(1) \cdot (GL_{n-1} \times GL_1)$.

We now need to calculate the the value of $\mathcal{M}$ on the elements of $X_{i,i-1}$ and $X_{i-1,i}$, as well as t_i. As a first step note that all these elements are in GL_i and commute with GL_{i-2}. Hence we must bound $\mathcal{M}(GL_n, GL_{n-2})$.

Lemma 5.9. (i) *The maximum multiplicity occurring in the restriction of any irreducible representation of GL_n to GL_{n-1} is at most 2^{n-1}.*

(ii) *For any $n \geq 1$ and $q \geq 2$, the number of conjugacy classes of $GL_n(q)$ is less than q^n.*

(iii) *The maximum multiplicity occurring in the restriction of any irreducible representation of GL_n to GL_{n-2}, $\mathcal{M}(GL_n, GL_{n-2})$, is less than $2^{2n-3}q^{n-1}$.*

Proof. Part (i) follows straight from the paper of Thoma [61]. Part (ii) is proved in Appendix A at the end of the paper. Part (iii) then follows from (i) and (ii) by noting that $\mathcal{M}(GL_n, GL_{n-2})$ is bounded by the product of the number of representations of GL_{n-1} with $\mathcal{M}(GL_n, GL_{n-1})$ and $\mathcal{M}(GL_{n-1}, GL_{n-2})$. □

The following corollary is an immediate consequence of Lemma 5.9.

Corollary 5.10. *Let $g \in GL_i$ commute with GL_{i-2}. Then $\mathcal{M}(g) < 2^{2i-3}q^{i-1}$.*

We are now ready to prove Theorem 5.7.

Proof of Theorem 5.7. Applying Corollaries 4.9 and 5.10 to the factorization of the first part of Lemma 5.8 yields

$$\begin{aligned} t_{GL_n} &< t_{P_n} + \sum_{k=2}^{n} 2^{2k-3}q^{k-1} \cdot \left(\sum_{l=1}^{k-1} q^l \right) \\ &\leq t_{P_n} + \frac{4}{15} 2^{2n} q^{2n-2}. \end{aligned} \tag{46}$$

Applying Corollary 4.9 to part (ii) of Lemma 5.8 gives

$$t_{P_n} \leq t_{GL_{n-1} \times GL_1} + q^{n-1}\mathcal{M}(GL_n, GL_{n-2}) + \sum_{k=2}^{n-1} q^k \mathcal{M}(GL_k, GL_{k-2}) + \\ + \sum_{k=3}^{n} q \cdot \mathcal{M}(GL_k, GL_{k-2}).$$

By Theorem 3.3

$$t_{GL_{n-1} \times GL_1} \leq t_{GL_{n-1}} + t_{GL_1},$$

so for $n \geq 2$ we obtain

$$t_{P_n} < t_{GL_{n-1}} + t_{GL_1} + \frac{1}{5} 2^{2n} q^{2n-2}. \tag{47}$$

Now we use these inequalities recursively. In the case of GL_1 we use the naive bound of $q-1$ for t_{GL_1}. An examination of the derivation of inequalities (46) and (47) shows that we have dropped several negative terms along the way, and that these terms dominate all the t_{GL_1} terms that appear. Thus we may ignore the

t_{GL_1} terms that appear during the recursion and at the bottom of the recursion. Summing all the other terms that appear gives the final result

$$t_{GL_n} < \frac{1}{2} 2^{2n} q^{2n-2}.$$

□

Remarks. 1. *Further improvements.* By improving the bound for t_{GL_2} we can improve on Theorem 5.7. Application of the results of [46] show that $t_{GL_2} \leq 200q \log q$. In fact, a generalization of our methods, applied to the appropriate subgroup chain of GL_2, shows that t_{GL_2} may be bounded by $5q - 3$; for details see [50].

2. *Variations of the algorithm.* There is of course nothing canonical about either the generators chosen here for GL_n or the subgroup chain. It seems highly likely that better choices for either are possible. Always, commutativity will need to be exploited and here it may be necessary to effectively compute the centralizers of various subsets of elements. Towards this end, recent advances in computational group theory for matrix groups [10] may prove useful.

3. *Other work.* The problem of finding an efficient algorithm for computing a Fourier transform for $GL_n(q)$ was first considered in [47]. There an algorithm is proposed which uses "models" (direct sums of induced one-dimensional representations which contain each irreducible of the group exactly once) to compute a Fourier transform for GL_n. In so doing the algorithm proceeds in two parts: (1) Computing the Fourier transform at reducible representations which are given by monomial matrices and then (2) applying projection operators to these reducible matrices in order to obtain a collection of unique irreducible Fourier transforms. Some simple asymptotics for the bounds they obtain yield an estimate for the complexity of their algorithm of

$$O(q^{\frac{n^2-2n}{4}} |GL_n(q)|).$$

4. *Direct approach using Theorem 3.2.* It is also necessary to compare our algorithm with the algorithm which uses the subgroup chain but does not factor the coset representatives and thus performs direct matrix multiplication of the twiddle factors. Straightforward analysis then shows that such an algorithm yields an upper bound which depends on the maximum degree of an irreducible representation of GL_n, which is of the order of $q^{\frac{1}{2}(n^2-n)}$. This direct algorithm gives an upper bound of

$$O(nq^{\frac{1}{2}(n^2-3n)} |GL_n(q)|).$$

5.5. **The unitary group over a finite field.** Let $U_n(q^2)$ denote the group of unitary $n \times n$ matrices with entries in the field of q^2 elements, relative to the field automorphism of order 2, where q is some prime power. We shall often abbreviate this to U_n. To simplify our calculations we shall always assume that q is odd. We consider the chain of subgroups

$$U_n > U_{n-1} > \cdots > U_1 \tag{48}$$

where U_k is identified with the subgroup $\mathrm{Diag}(U_k, I_{n-k})$ of U_n.

Theorem 5.11. *Let $\mathcal{R}$ be any complete set of irreducible representations of $U_n(q^2)$ adapted to the chain of subgroups* (48)*. Then for any $n \geq 2$, $q \geq 2$,*

$$\mathcal{C}(U_n(q^2)) \leq T_{U_n(q^2)}(\mathcal{R}) < 3(1+\frac{18}{7q})B_1(q^{-1})q^{3n-3}\left|U_n(q^2)\right| \tag{49}$$

where $B_1(t) = \prod_{k=1}^{\infty} \frac{1+t^k}{1-t^k}$.

Remark. In Appendix A, we show that, for $q \geq 2$,

$$B_1(q^{-1}) < 1 + 2q^{-1} + 4q^{-2} + (42.05)q^{-3}$$

so in particular, $B_1(q^{-1}) < 8.26$, and $B_1(q^{-1})$ tends to 1 as q gets large. We also give a series of related bounds for B_1, and tabulate some values of $B_1(q^{-1})$. Also, the factor $3(1+\frac{18}{7q})$ can be replaced by $3q^4\left[(q^3-1)(q-1)\right]^{-1}$. When $q = 2$ this is $\frac{48}{7}$, but it tends to 3 for large q.

Before proving Theorem 5.11, we shall first prove the following weaker but simpler result:

Claim. Assume $\mathcal{R}$ and B_1 are as in Theorem 5.11. Then

$$t_{U_n}(\mathcal{R}) < \frac{32}{7}B_1(q^{-1})q^{3n-2}. \tag{50}$$

To prove the Claim we proceed as in the case of GL_n and obtain a factorization of any element of U_n as a product of matrices which are either diagonal or have a single 2×2 block with ones elsewhere on the diagonal. The multiplicity results we will need are given in the following lemma.

Lemma 5.12. (i) *The maximum multiplicity occurring in the restriction of any irreducible representation of U_n to U_{n-1} is* 1*.*

(ii) *For any $n \geq 1$ and $q \geq 2$, the number of conjugacy classes of $U_n(q^2)$ is less than $q^n B_1(q^{-1})$.*

(iii) *The maximum multiplicity occurring in the restriction of any irreducible representation of U_n to U_{n-2} is less than $q^{n-1}B_1(q^{-1})$.*

Proof. (i) is a result of Hagedorn [35]. (ii) is proved in Appendix A. (iii) is a direct consequence of (i) and (ii). □

So as to not unduly interrupt the flow of the section the necessary factorization of coset representatives of U_n/U_{n-1} is obtained in Appendix A, by using some simple geometry. To state the result succinctly, we let $u_i(x_1, x_2)$ be the block diagonal matrix with ones on the diagonal except for a 2×2 block of the form

$$\begin{pmatrix} -x_2^q & x_1 \\ x_1^q & x_2 \end{pmatrix}.$$

This matrix is in $U_n(q^2)$ provided that $x_1^{1+q} + x_2^{1+q} = 1$.

Lemma 5.13. *Let N be the group homomorphism on $\mathbf{F}_{q^2}^{\times}$ given by $N(\alpha) = \alpha^{1+q}$ and let R be a complete set of coset representatives for $\mathbf{F}_{q^2}^{\times}/\ker N$. Then every coset of U_n/U_{n-1} has at least one coset representative of the form $\varepsilon \cdot a_2 \cdots a_n$, where ε is an element of $\mathbf{F}_{q^2}$ satisfying $\varepsilon^{1+q} = 1$ and for $2 \leq i \leq n-1$, the matrix a_i has one of the following forms:*

(A) *$a_i = u_i(r, x)$ for some $r \in R$, $x \in \mathbf{F}_{q^2}$ such that $r^{1+q} + x^{1+q} = 1$.*

(B) *$a_i = t_{i+1}u_i(r, x)$ for some $r \in R$, $x \in \mathbf{F}_{q^2}$ such that $r^{1+q} + x^{1+q} = 1$.*

(C) $a_i = t_i t_{i+1} u_i(r, r\delta)$, *where* r *is the unique element of* R *with* $r^{1+q} = \frac{1}{2}$, *and* $\delta \in \mathbf{F}_{q^2}$ *satisfies* $\delta^{1+q} = 1$.

The factor a_n *has the form* (A).

We can now prove the claim.

Proof of the Claim. Applying Corollary 4.9 to the factorization of Lemma 5.13 shows that

$$\begin{aligned} t_{U_n} &\leq t_{U_{n-1}} + q^{n-1}(q^n - (-1)^n)\left[1 + g_{n-1}(q^2) + 3\sum_{k=2}^{n-1} g_{k-1}(q^2)\right] \\ &< t_{U_{n-1}} + 4q^{3n-2}B_1(q^{-1}) \end{aligned}$$

where $g_k(q^2)$ denotes the number of conjugacy classes of $U_n(q^2)$. Using this inequality recursively and noting that $t_{U_1} \leq q+1$ gives the result (50). □

With these preliminary results in hand we now easily prove Theorem 5.11.

Proof of Theorem 5.11. The improvement on the Claim comes from looking at the matrix multiplications in the separation of variables algorithm more carefully. Suppose we are computing the Fourier transform at the adapted irreducible representation, ρ. At some point in the algorithm we will calculate matrix products of the form $\rho(a_n) \cdot \hat{h}(\rho \downarrow U_{n-1})$, where $a_n \in U_n$ commutes with U_{n-2} and $\hat{h}(\rho \downarrow U_{n-1})$ is in $\mathrm{End}_1(\rho \downarrow U_{n-1})$. To obtain the complexity result (50) we used the bound of $\mathcal{M}(a_n)d_\rho^2$ for the complexity of such a matrix multiplication—a bound which comes without assuming any special form of the matrix $\hat{h}(\rho \downarrow U_{n-1})$. However, we could get a better result by using part of Theorem 4.7 to bound the complexity of that matrix multiplication:

$$C(\rho; U_n, U_{n-2}; U_{n-1}, 1) \leq d_\rho^2.$$

Using this new complexity gives us

$$\begin{aligned} t_{U_n} &\leq t_{U_{n-1}} + q^{n-1}(q^n - (-1)^n)\left[1 + 1 + 3\sum_{k=2}^{n-1} g_{k-1}(q^2)\right] \\ &< t_{U_{n-1}} + 3(1 + \frac{2}{q})q^{3n-3}B_1(q^{-1}). \end{aligned}$$

Using this bound recursively proves the theorem. □

5.6. **Chevalley groups.** The techniques used to compute a Fourier transform in GL_n may be extended in a relatively straightforward manner to Chevalley groups and other finite groups of Lie type. We refer the reader to the book of Carter [16] for definitions. We limit the current discussion to the classical Chevalley groups although the techniques generalize in a natural way to other finite groups of Lie type.

As usual, let $A_n(q)$, $B_n(q)$, $C_n(q)$, $D_n(q)$ denote the simply connected forms of the Chevalley groups over a finite field with q elements. Any Chevalley group G has a subgroup chain analogous to (43), where P_{n-1} is replaced by a maximal parabolic subgroup and $GL_{n-1} \times GL_1$ is replaced by the reductive part, and GL_{n-1} by the semisimple part of the parabolic subgroup. More specifically, we shall label the simple roots of a rank n group from 1 to n in the order shown in Diagram 5. Then P_k will denote the parabolic subgroup corresponding to the set of simple roots labeled from 1 to k with reductive part L_k and semisimple part G_k (not to be confused with the exceptional group G_2 when $k = 2$). For any Chevalley group G the chain of subgroups we shall use in the construction of a fast Fourier transform on G will always be

$$G_n \geq P_{n-1} \geq L_{n-1} \geq G_{n-1} \geq \cdots \geq G_1. \tag{51}$$

Theorem 5.14. *For any $n \geq 2$ there is a positive constant K_n, such that for any $q \geq 2$,*

(i) $T_{A_n(q)} \leq K_n q^{2n+1}. |A_n(q)|$,
(ii) $T_{B_n(q)} \leq K_n q^{5n-3}. |B_n(q)|$,
(iii) $T_{C_n(q)} \leq K_n q^{5n-3}. |C_n(q)|$,
(iv) $T_{D_n(q)} \leq K_n q^{5n-6}. |D_n(q)|$, *for* $n \geq 4$, *and* $T_{D_3(q)} \leq K_3 q^{10}. |D_3(q)|$, *for* $n = 4$

where the complexities are taken with respect to a complete set of representations adapted to the chain (51).

We shall give the proof of Theorem 5.14 after we have collected some lemmas on multiplicities and factoring elements in these groups.

We refer the reader to [16] for all the relevant notation. For any root α in the root system of G_n, we let X_α denote the corresponding root subgroup. We also let s_α denote the corresponding involution in the Weyl group, and let n_α be an element of N mapping onto s_α where N comes from the BN-pair for G_n. We shall denote the simple roots $\alpha_1, \ldots, \alpha_n$ according to Diagram 5. With the exception of the root α_3 of D_3, we know that X_{α_i} and n_{α_i} lie in G_i and commute with G_{i-2}. Consequently, the construction of an FFT using a factorization in terms of the X_{α_i} or n_{α_i} will require that we understand the maximum multiplicity $\mathcal{M}(G_i, G_{i-2})$.

Lemma 5.15. *Let (G, K) be one of the group-subgroup pairs $(A_n(q), A_{n-2}(q))$, $(B_n(q), B_{n-2}(q))$, $(C_n(q), C_{n-2}(q))$ or $(D_n(q), D_{n-2}(q))$. Then for fixed n the maximum multiplicity $\mathcal{M}(G, K)$ is bounded by a function of q of the form $O(q^{\sigma(G,K)})$, where*

$$\sigma(G, K) = \frac{1}{2}\left[\dim G - \operatorname{rank} G - \dim K - \operatorname{rank} K\right].$$

Proof. This is proved in the Appendix A using an argument due to Tom Hagedorn. See also [35]. □

The other piece of information we need concerns the factorization of coset representatives in terms of the elements X_{α_i} and n_{α_i}.

Lemma 5.16. *Let G be a simply connected Chevalley group with Weyl group $\mathfrak{W}$ and let J be any subset of the set of simple roots of $\mathfrak{W}$. Let $\mathfrak{W}_J$ denote the parabolic subgroup corresponding to J and $\mathfrak{W}^J$ the the set of minimal coset representatives for $\mathfrak{W}/\mathfrak{W}_J$. We let N, N_J denote the number of positive roots of $\mathfrak{W}$, $\mathfrak{W}_J$ respectively. Also let P_J denote the parabolic subgroup of G corresponding to J, let L_J and U_J be its reductive and maximal normal unipotent parts, let $Z(L_J)$ be the center of L_J and G_J be the semisimple part of L_J. Then*

(i)

$$G = \left[\coprod_{\substack{w \in \mathfrak{W}^J \\ w = s_{\beta_1} \cdots s_{\beta_k}}} (X_{\beta_1} n_{\beta_1}) \ldots (X_{\beta_k} n_{\beta_k}) \right] \cdot P_J$$

where the $w = s_{\beta_1} \cdots s_{\beta_k}$ *is a reduced expression for* w *in terms of simple reflections.*

(ii) $P_J = U_J \cdot L_J$ *and* $|U_J| = q^{N-N_J}$.

(iii) *If* G *is not of type* G_2*, then there is a sequence,* $\beta_1, \ldots, \beta_m$ *of simple roots such that* $U_J \subseteq \prod_i X_{\beta_i}$ *over any field of odd characteristic.*

(iv) $L_J = Z_{L_J}.G_J$ *and* $|L_J/G_J| = (q-1)^{\operatorname{rank} G - |J|}$.

Proof. Parts (i), (ii) and (iv) follow from the first two chapters in Carter's book [16]. Part (iii) follows from the Steinberg commutator relations. See, e.g., [15] Theorem 12.1.1. □

Proof of Theorem 5.14. We let N_k denote the number of positive roots of G_k. First we assume that $n \geq 2$ in the cases where G_n is of type A, B, or C, and $n \geq 4$ in the case that G has type D. From Lemma 5.15 it is clear that $|G_n/P_n|$ is a polynomial of degree $N_n - N_{n-1}$ in q and that any coset of G_n/P_n has a coset representative of length no more than $N_n - N_{n-1}$ in the generators $(X_\alpha n_\alpha)$. Therefore

$$\begin{aligned} t_{G_n} &\leq (N_n - N_{n-1})\,|G_n/P_n|\,\mathcal{M}(G_n, G_{n-2}) + t_{P_n} \\ &\leq O(q^{N_n - N_{n-1} + \sigma(G_n, G_{n-2})}) + t_{P_n}, \end{aligned}$$

where the 'O' notation indicates a constant independent of q, but which does depend on n. Now let U_n denote the maximal normal unipotent subgroup of P_n and let γ_n be such that U_n is contained in a product of no more than γ_n simple root subgroups, independent of q. This is possible by Lemma 5.16, part (iii). Then

$$\begin{aligned} t_{P_n} &\leq \gamma_n\,|U_n|\,\mathcal{M}(G_n, G_{n-2}) + t_{L_n} \\ &\leq O(q^{N_n - N_{n-1} + \sigma(G_n, G_{n-2})}) + t_{L_n} \\ &\leq O(q^{N_n - N_{n-1} + \sigma(G_n, G_{n-2})}) + t_{G_{n-1}} \end{aligned}$$

and therefore $t_{G_n} \leq O(q^{N_n - N_{n-1} + \sigma(G_n, G_{n-2})}) + t_{G_{n-1}}$. A quick glance at Table 2 verifies that for all the series of groups, $N_n - N_{n-1} + \sigma(G_n, G_{n-2})$ is an increasing function of n, and hence that

$$t_{G_n} \leq O(q^{N_n - N_{n-1} + \sigma(G_n, G_{n-2})}) + t_{G_1}$$

for the series of groups of type A, B, C. For these three series, $G_1 = A_1(q)$ and hence t_{G_1} is bounded by $O(q^3)$ using a naive method of calculating a Fourier transform. For the groups D_n with $n \geq 4$ we have

$$t_{D_n} \leq O(q^{5n-6}) + t_{D_3}$$

and t_{D_3} may be bounded by $O(q^{10})$ using similar techniques. Hence we see that when $n \geq 2$, and type A, B, or C, or $n \geq 4$ and type D, we have

$$t_{G_n} \leq O(q^{N_n - N_{n-1} + \sigma(G_n, G_{n-2})}).$$

□

TABLE 2. COMBINATORIAL DATA FOR CHEVALLEY GROUPS.

G_n	N_n	$\sigma(G_n, G_{n-2})$	$N_n - N_{n-1} + \sigma(G_n, G_{n-2})$
A_n	$\frac{1}{2}n(n+1)$	$n+1$	$2n+1$
B_n	n^2	$3n-2$	$5n-3$
C_n	n^2	$3n-2$	$5n-3$
D_n	n^2-n	$3n-4$	$5n-6$

6. Homogeneous spaces

For statistical applications, data on homogeneous spaces is often of interest, rather than data on the full group. In brief, a homogeneous space for a finite group is simply a set on which the group acts transitively as permutations. In other words it is a coset space G/K. A common example is the action of the finite affine group on point-line pairs and more generally, the action of an automorphism group of a design on its point-block pairs. In this case generalizations of the standard analysis of variance of data on such sets require the computation of projections of the data vector onto group-invariant subspaces. The techniques to be developed in this section provide speed-ups of the most efficient algorithms currently known (cf. [28] and references therein).

The analyses of Sections 3 and 4 can be applied with virtually no changes to the computation of transforms on homogeneous spaces. In particular, the sums that arise in the computation of the Fourier transform of K-invariant functions on a group, G, can be computed using the separation of variables algorithm. We simply need to re-investigate the structure of the matrices, which is more special in the homogeneous space case.

Definition 6.1 (Fourier Transform)**.** Let G be a finite group with subgroup K, and let f be a complex-valued function on G/K.

(i) Let ρ be a matrix representation of G. Then the **Fourier transform of f at ρ** is

$$\hat{f}(\rho) = \sum_{y \in Y} \rho(y) f(yK) P_K \tag{52}$$

where Y is any set of coset representatives for G/K, and P_K is the canonical projection from the representation space, V_ρ, onto the subspace of invariant vectors, V_ρ^K.

(ii) Let $\mathcal{R}$ be a set of matrix representations of G. Then the **Fourier transform of f on $\mathcal{R}$** is the set of Fourier transforms of f at the representations in $\mathcal{R}$.

It is immediate from the definition that the Fourier transform of a function on G/K is the same, up to a fixed scalar multiple, as the Fourier transform of the associated right K-invariant function on G. It is also immediate that $\hat{f}(\rho)$ is in $\mathrm{Hom}_{\mathbf{C}}(V_\rho^K; V_\rho)$, and hence that $\hat{f}(\rho)$ is zero unless V_ρ contains a nontrivial K-invariant vector. Representations containing a nontrivial K-invariant vector are said to be **class 1 relative to** K, and in the following discussion we may always assume that our representations are class 1, if needed.

The analysis of the homogeneous space case mimics most of the previous discussion. By analogy, we set up the following notation.

Definition 6.2 (Complexity)**.** Let G be a finite group with subgroup K, and let $\mathcal{R}$ be any set of matrix representations of G.

(i) Let $T_{G/K}(\mathcal{R})$ denote the minimum number of operations needed to compute the Fourier transform of f on $\mathcal{R}$ via a straight-line program for an arbitrary complex-valued function f defined on G/K.

(ii) Let $t_{G/K}(\mathcal{R}) = T_{G/K}(\mathcal{R}) / |G/K|$.

$T_{G/K}(\mathcal{R})$ is called the **complexity** of the Fourier transform on G/K for the set $\mathcal{R}$, and $t_{G/K}$ is called the **reduced complexity**.

The complexity always satisfies the inequalities

$$|G/K| - 1 \leq T_{G/K}(\mathcal{R}) \leq |G/K|^2 .$$

A brief glance at equation (52) verifies that it has exactly the form required in order to apply the separation of variables approach. Indeed, equation (52) is exactly the same as (12) of Section 4.1, with $F(y) = f(yK)P_K$. Hence, all of Section 4.1 applies to our situation. In the same way as the separation of variables approach implied the results of Section 3, we may obtain analogues of the coset decomposition results for Fourier transforms on homogeneous spaces.

Assume G is a finite group, K is a subgroup of G, Y is a subset of G , and $\mathcal{R}$ is a set of matrix representations of G. Then let $M_{G/K}(Y, \mathcal{R})$ denote the minimum number of operations required to compute the sums

$$\sum_{y \in Y} \rho(y) F(y, \rho)$$

for all ρ in $\mathcal{R}$, where for each $y \in Y$ and $\rho \in \mathcal{R}$, $F(y, \rho)$ is an arbitrary matrix in $\mathrm{Hom}_{\mathbf{C}}(V_\rho^K; V_\rho)$. Let $m_{G/K}(Y, \mathcal{R}) = M_{G/K}(Y, \mathcal{R}) / |G/K|$.

The analogue of Theorem 3.1 in this case is

Theorem 6.1. *Let $G \geq H \geq K$ be a chain of finite groups, and let $\mathcal{R}$ be a complete chain-adapted set of inequivalent irreducible matrix representations of G. Let Y be a set of coset representatives for G/H. Then*

$$t_{G/K}(\mathcal{R}) \leq t_{H/K}(\mathcal{R}_H) + m_{G/K}(Y, R).$$

Proof. Let Y be a set of coset representatives for G/H and Z be a set of coset representatives for H/K. Then $Y \cdot Z$ is a set of coset representatives for G/K, so

$$\begin{aligned} \hat{f}(\rho) &= \sum_{y \in Y} \rho(y) \left[\sum_{z \in Z} \rho(z) f_y(z.K) P_K \right] \\ &= \sum_{y \in Y} \rho(y) \hat{f}_y(\rho \downarrow H) \end{aligned} \tag{53}$$

where for each $y \in Y$, f_y is the function of H/K defined by $f_y(zK) = f(yzK)$. We have reduced the computation of $\hat{f}(\rho)$ to a collection of transforms on H/K and the sum (53). The result follows easily. □

As in the group case, this bound may be improved using the block diagonal structure of $\hat{f}_y(\rho \downarrow H)$.

The sum (53) also has the form of (12), which is required for separation of variables to apply. The only way in which the analysis for homogeneous spaces differs from that for groups is in the structure of the matrices $F(y)$ that occur. We now analyze the matrix multiplications that occur when we apply separation of variables to $F(y) = f(yK)P_K$.

Theorem 6.2. *Assume $G \geq H \geq L$ is a chain of finite groups, K is a subgroup of G, and ρ is a matrix representation of G which is adapted relative to all these groups. Suppose that a is in the centralizer of L in H, and F is a matrix in $\mathrm{Hom}_{\mathbf{C}}(V_\rho^K, V_\rho)$. Then the matrix multiplication $\rho(a) \cdot F$ can be computed in no more than*

$$\mathcal{M}(\rho, 1_K) \sum_{\alpha, \beta} \mathcal{M}(\rho, \alpha) \mathcal{M}(\alpha, \beta)^2 d_\beta$$

scalar operations, where α, β range over irreducible representations of H and L respectively.

Proof. The condition on a implies that $\rho(a)$ is in $\mathrm{End}_L(\rho \downarrow H)$, so that the statement of the theorem has a form similar to that of Theorem 4.5. The proof follows the same lines as that of Theorem 4.5, but we need instead consider a diagram like Diagram 6.

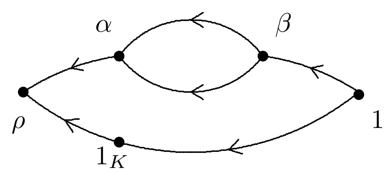

Diagram 6

The procedure for obtaining the complexity bound is the same. □

Note that $\mathcal{M}(\rho, 1_K) = \dim V_\rho^K$.

Lemma 6.3. *In the notation of Theorem 6.2,*

$$\mathcal{M}(\rho, 1_K) \sum_{\alpha,\beta} \mathcal{M}(\rho,\alpha)\mathcal{M}(\alpha,\beta)^2 d_\beta \leq \mathcal{M}(\rho, 1_K) d_\rho \mathcal{M}(H, L).$$

Now we can apply the separation of variables technique, as in Section 4.1, to get:

Corollary 6.4. *Let $\mathcal{R}$ be a complete set of inequivalent irreducible matrix representations for G, adapted to the subgroup chain $K_n \geq \cdots \geq K_0$ and to the chain $G \geq H \geq K$. Let Y be a complete set of coset representatives for G/H, and X be a set of factorizations of elements of Y in terms of elements from a subset $S \subset G$. Let γ be the maximum length of any word in X. Then*

$$m_{G/K}(Y, \mathcal{R}) \leq \sum_{k=0}^{\gamma-1} \sum_{s_n \cdots s_0 \in \tilde{X}_k} \mathcal{M}(s_0) \tag{54}$$

$$\leq \mathcal{M}(S) \left[\sum_{k=0}^{\gamma-1} \left| \tilde{X}_k \right| \right] \tag{55}$$

where $\tilde{X}_k$ is obtained from X_0 by deleting k elements from the right of each word and then deleting all words with an identity element at the far right.

Proof. This comes from summing the bound of Lemma 6.3 over all the matrix multiplications in the separation of variables algorithm. We are able to cancel the $|G/K|^{-1}$ factor in the definition of $m_{G/K}$ using the identity

$$\sum_\rho d_\rho \mathcal{M}(\rho, 1_K) = |G/K|$$

where ρ ranges over the (class 1) representations of G. □

Comparing Corollary 6.4 with Corollary 4.9, we see that the upper bounds we have derived for $m_{G/K}$ and m_G are the same. It follows that the bounds for m_G in Corollary 4.10 are also bounds for $m_{G/K}$. In fact, we can tighten the results of Corollary 6.4 even further: As we are only dealing with class 1 representations of G, we could replace the $\mathcal{M}(s_0)$ terms in Corollary 6.4, with maximum multiplicities taken only over representations that occur in restrictions of class 1 representations of G.

6.1. Weyl group homogeneous spaces: S_n/S_{n-k}, B_n/B_{n-k}, and D_n/D_{n-k}.

Theorem 6.5. *Let $\mathcal{R}$ be any complete set of irreducible representations of S_n, adapted to the chain of subgroups $S_n > S_{n-1} > \cdots > S_1$. Then*

$$T_{S_n/S_{n-k}}(\mathcal{R}) \leq k\left(n^2 - kn + \frac{k^2-1}{3}\right)|S_n/S_{n-k}|.$$

Proof. From Corollary 6.4 and the discussion in Section 5.2, we see that $t_{S_n/S_{n-k}} \leq t_{S_{n-1}/S_{n-k}} + n(n-1)$. This is easily iterated to give the desired result. □

This result is only useful for $k \geq 3$. When $k = 2$ a naive approach to the computation requires $n(n-1)\,|S_n/S_{n-2}|$ scalar operations, whereas an algorithm based on Theorem 6.5 requires $2(n-1)^2\,|S_n/S_{n-2}|$ operations. When $k = 1$, the left S_{n-1}-invariant matrix coefficients factor, and may be written in a very simple form. This leads to an efficient algorithm for computing Fourier transforms on S_n/S_{n-1}.

Theorem 6.6. *If $\mathcal{R}_B$ and $\mathcal{R}_D$ are complete sets of irreducible representations of B_n and D_n respectively, each set adapted to the appropriate chain of subgroups, $B_n > B_{n-1}\cdots > B_1$ or $D_n > D_{n-1}\cdots > D_1$, then*

$$\begin{aligned} T_{B_n/B_{n-k}}(\mathcal{R}_B) &\leq k\left(4n^2 - (4k-2)n + \frac{(k-1)(4k+1)}{3}\right)|B_n/B_{n-k}|, \\ T_{D_n/D_{n-k}}(\mathcal{R}_D) &\leq 4k\left(n^2 - kn + \frac{1}{3}k^2 - \frac{1}{3}\right)|D_n/D_{n-k}|. \end{aligned}$$

Proof. From the proof of Theorem 5.3 we deduce that $t_{B_n/B_{n-k}} \leq t_{B_{n-1}/B_{n-k}} + 2n(2n-1)$ and $t_{D_n/D_{n-k}} \leq t_{D_{n-1}/D_{n-k}} + 4n(n-1)$. □

6.2. Matrix group homogeneous spaces: GL_n/GL_{n-k} and U_n/U_{n-k}.

Theorem 6.7. *Let $\mathcal{R}_G$ and $\mathcal{R}_U$ be complete sets of irreducible representations of $GL_n(q)$ and $U_n(q^2)$, adapted to the chains of subgroups* (43) *and* (48) *respectively. Then*

$$T_{GL_n/GL_{n-k}}(\mathcal{R}_G) \leq \tfrac{1}{2}2^{2n}q^{2n-2}\,|GL_n/GL_{n-k}|$$

and

$$T_{U_n/U_{n-k}}(\mathcal{R}_U) \leq 3(1+\tfrac{18}{7q})B_1(q^{-1})q^{3n-3}\,|U_n/U_{n-k}|.$$

Proof. The bounds on the reduced complexity of the Fourier transforms on GL_n and U_n that we obtained in Sections 5.4 and 5.5 are also bounds for the reduced complexity of transforms on the homogeneous spaces. □

7. Further directions and improvements

Theorem 4.8 and Corollary 4.9 are particularly easy to use but are by no means the best results possible. We now briefly describe some of the improvements we have obtained, which will appear in the second part of this paper [50].

7.1. **Variations on the main results.** In many cases, further savings can be realized if the Fourier transform is treated as a collection of scalar equations rather than as a matrix equation. The separation of variables idea still applies to the scalar setting, but now we obtain a recursive sum of products of numbers, as opposed to matrices. These products may be computed in any order. Consequently, the scalar separation of variables algorithm possesses a flexibility which is not present in the matrix separation of variables algorithm: the ability to choose the order in which the factors are summed over. Roughly speaking, this flexibility allows us to sum over factors with a low value of $\mathcal{M}$ first, successively building the complete computation. In practice the first summations we perform occur the most times in the separation of variables algorithm (in the matrix case, this amounts to saying that the sets X_k get smaller as k increases), so by ensuring these sums are done more quickly, we make the whole algorithm more efficient.

The sums that occur in the scalar separation of variables algorithm are generalizations of the sum (22), and the factors that appear are indexed by collections of representations which satisfy relations generalizing the relations represented by Diagram 1. The diagrammatic methods used in the proof of Theorem 4.5 generalize to this situation, so complexity bounds for the new algorithms may be obtained explicitly. A useful combinatorial tool here is to treat the indices as injections from the diagrams describing the relations (e.g., Diagrams 2–4, and 6) into the Bratteli diagram for the subgroup chain. The explicit expressions for the complexity of the algorithm has a form similar to, but generalizing, the expressions in Theorems 4.5, 4.6, and 4.8.

We use the techniques just described to refine the results we have already obtained in Section 5. For example, we get a better bound for the complexity of the Fourier transform on $GL_n(q)$ using the same bases as in Section 5.

Theorem 7.1. *Let $\mathcal{R}$ be any complete set of irreducible representations of $GL_n(q)$ adapted to the chain of subgroups*

$$GL_n > GL_{n-1} \times GL_1 > GL_{n-1} > \cdots > GL_1 \times GL_1 > GL_1.$$

Then for any $n \geq 2$, there is a positive constant K_n such that, for $q \geq 2$,

$$T_{GL_n(q)}(\mathcal{R}) \leq K_n q^n \left| GL_n(q) \right|.$$

Similar improvements hold for the unitary groups and other finite groups of Lie type. We also prove a general theorem, bounding t_G in terms of the complexities of two subgroups and the number of double cosets. This result works particularly nicely when the subgroups are abelian, and in that case it yields new results for $SL_2(q)$ and for the symmetric groups.

7.2. **Homogeneous spaces.** The scalar separation of variables algorithm generalizes easily to the context of homogeneous spaces. The idea in the improved algorithms is to write the associated spherical functions of the homogeneous space as a sum of products, with a small number of terms in the sum. The separation of variables algorithm then amounts to calculating the inner product of a function and an associated spherical function by summing over one factor in the product at a time. This provides speed-ups of the algorithms in Section 6 (cf. [28] and references therein). This material will also appear in [50].

Appendix A. Proofs of the technical lemmas

Now we shall prove the remaining lemmas used in the explicit calculations of Section 5. These concern estimates for the number of conjugacy classes of the general linear and unitary groups, the derivation of the factorization for coset representatives of U_n/U_{n-1}, and bounds for the multiplicities of restrictions between Chevalley groups.

A.1. **Conjugacy classes.** In Lemmas 5.9 and 5.12 we stated upper bounds for the number of conjugacy classes in $GL_n(q)$ and $U_n(q^2)$ respectively. These two results are simple corollaries of the following discussion.

The generating functions for the number of conjugacy classes of $GL_n(q)$, the number of canonical forms of $n \times n$ matrices over $\mathbf{F}_q$, and the number of conjugacy classes of $U_n(q^2)$ are closely related. Define

$$F_\alpha(q,t) = \prod_{n=1}^{\infty} \frac{1+\alpha t^n}{1-qt^n}$$

and let $f_n(q;\alpha)$ be the coefficient of t^n in the expansion of $F_\alpha(q,t)$ considered as a power series in t. Then by results of [60] and [64] $f_n(q;-1)$ is the number of conjugacy classes of $GL_n(q)$, $f_n(q;0)$ is the number of canonical forms of $n \times n$ matrices over $\mathbf{F}_q$, and $f_n(q;1)$ is the number of conjugacy classes of $U_n(q^2)$; the key to bounding $f_n(q;\alpha)$ is to first understand the coefficients of $f_n(q;-1)$.

Lemma A.1. (i) *For $n \geq 1$,*

$$f_n(q;-1) = q^n - \left(q^{\lfloor \frac{n-1}{2} \rfloor} + \cdots + q^{\lfloor \frac{n}{3} \rfloor}\right) + \left(\text{terms of degree} \leq \left\lfloor \frac{n-1}{3} \right\rfloor - 1\right).$$

(ii) *For $n \geq 1$ and $q \geq 2$, we have $f_n(q;-1) < q^n$.*

Proof. Let $p(n,k)$ denote the number of partitions of n into exactly k parts, let $p(n)$ denote the total number of partitions of n, and let $r(n,k)$ be the coefficient of q^k in $f_n(q;-1)$. Then $p(n,k)$ is the coefficient of q^k in $f_n(q;0)$, so using Euler's pentagonal number theorem, it is easy to see that

$$r(n,k) = \sum_{l \in \mathbf{Z}} p(n - 3l(l+1)/2, k).$$

Hence $r(n,k)$ satisfies the recurrence relation, $r(n,k) = r(n+1,k+1) - r(n-k,k+1)$, which follows from the same recurrence relation for $p(n,k)$. Suppose now that $\frac{n}{2} \leq k < n$. Then $p(n,k) = p(n-k)$, and so

$$r(n,k) = p(n-k) - p(n-k-1) - p(n-k-2) + p(n-k-5) + \cdots = 0.$$

Trivially, $r(n,n) = 1$, and for $k > n$, we have $r(n,k) = 0$. Thus we know $r(n,k)$ for all $k \geq \frac{n}{2}$, and it is easy to see that this data, together with the recurrence relation for r, determines $r(n,k)$ for all n and k. If $\frac{n}{3} \leq k < \frac{n}{2}$, then $\frac{n-k}{2} \leq k < (n-k)$, so using the recurrence relation gives

$$\begin{aligned} r(n,k) &= r(n+1,k+1) - r(n-k,k+1) = r(n+1,k+1) = \cdots \\ &= r(2(n-k)+1,(n-k)+1) = r(2(n-k),n-k) - r(n-k,n-k) \\ &= 0 - 1 = -1. \end{aligned}$$

The same idea can be used to find $r(n,k)$ for any n and k. For example, if $\frac{n}{4} \leq k < \frac{n}{3}$, then $r(n,k) = \lceil \frac{n-3k-4}{2} \rceil$. This proves part (i) of the theorem. For part (ii), we start by noting that $r(n,k) \leq p(n,k)$, so that when $q \geq 2$ and $n \geq 1$,

$$f_n(q;-1) < q^n - q^{\lfloor \frac{n-1}{2} \rfloor} + q^{\lfloor \frac{n-1}{3} \rfloor - 1} p(\lfloor (n-1)/3 \rfloor - 1).$$

For $n \geq 55$ the result now follows by using the simple upper bound $p(n) \leq e^{\pi (2/3)^{\frac{1}{2}} (\sqrt{n}-1)}$ to show that $p(n) \leq 2^{\lfloor \frac{n}{2} \rfloor}$. For $n < 55$ the result follows by inspection. □

To extend Lemma A.1 to $f_n(q;\alpha)$ for $\alpha \geq 1$, we need to consider the function

$$B_\alpha(t) = \sum_{n=1}^{\infty} \frac{1+\alpha t^n}{1-t^n} = F_\alpha(1,t) = \sum_{n=0}^{\infty} f_n(1;\alpha) t^n.$$

Theorem A.2. *For $n \geq 1$, $q \geq 2$, $\alpha \geq -1$,*

$$f_n(q;\alpha) < B_\alpha(q^{-1}) q^n.$$

Proof. First note that $F_\alpha(q,t) = B_\alpha(t) F_{-1}(q,t)$. Multiplying this out and equating coefficients of t^n gives

$$f_n(q;\alpha) = \sum_{m=0}^{n} f_m(1;\alpha) f_{n-m}(q;-1).$$

But $f_m(1,\alpha)$ is a polynomial in $\alpha+1$ with nonnegative integral coefficients, so if $\alpha \geq -1$ then $f_m(1,\alpha) \geq 0$ and $f_0(1,\alpha) = 1$. Hence

$$f_n(q;\alpha) < \sum_{m=0}^{n} f_m(1;\alpha) q^{n-m} \leq q^n \sum_{m=0}^{\infty} f_m(1;\alpha) q^{-m} = q^n B_\alpha(q^{-1}).$$

□

Remark. As a consequence of the proof of Theorem A.2, we see that for $\frac{n}{2} \leq k$, the coefficient of q^k in $f_n(q,\alpha)$ is the same as the coefficient of q^k in $q^n B_\alpha(q^{-1})$. I.e., it is $f_{n-k}(1;\alpha)$. From this, it is easy to show that, for fixed q, the constant $B_\alpha(q^{-1})$ appearing in Theorem A.2 is the best possible. In fact, a slight extension of the asymptotics due to Stong [60] shows that for fixed q and fixed $\alpha \geq -1$,

$$f_n(q;\alpha) = B_\alpha(q^{-1}) q^n + \frac{1}{2} \left[\frac{B_\alpha(q^{-\frac{1}{2}})}{1-q^{\frac{1}{2}}} + (-1)^n \frac{B_\alpha(-q^{-\frac{1}{2}})}{1+q^{\frac{1}{2}}} \right] q^{\frac{n}{2}} + O(q^{\frac{n}{3}})$$

as n tends to infinity.

Corollary A.3. *For $n \geq 1$, $q \geq 2$, $\alpha \geq -1$, $l \geq 0$,*

$$f_n(q;\alpha) < \left[\sum_{m=0}^{l-1} f_n(1;\alpha) q^{n-m} \right] + K_{\alpha,l} q^{n-l}$$

where $K_{\alpha,l} = 2^l \left(B_\alpha(2^{-1}) - \sum_{m=0}^{l-1} f_m(1;\alpha) 2^{-m} \right)$.

Proof. The expression

$$q^l \left(B_\alpha(q^{-1}) - \sum_{m=0}^{l-1} f_m(1;\alpha) q^{-m} \right) = \sum_{m=l}^{\infty} f_m(1;\alpha) q^{l-m}$$

is a decreasing function of q. Thus its maximum value in the permitted range of q occurs when $q = 2$, and is precisely $K_{\alpha,l}$. Rearranging, we obtain

$$B_\alpha(q^{-1}) \leq \left[\sum_{m=0}^{l-1} f_n(1,\alpha)q^{-m}\right] + q^{-l}K_{\alpha,l}.$$

The result now follows from Theorem A.2 □

For the purposes of this paper, we are only interested in the cases $\alpha = -1$ or $\alpha = 1$. Some simple bounds give an approximate value of $B_1(2^{-1}) = 8.25599$. The following tables display various values of $B_1(q^{-1})$ and $K_{1,l}$. We also give the first few terms of the power series for $B_\alpha(t)$ and $B_1(t)$.

l	0	1	2	3	4
$f_l(1;1)$	1	2	4	8	14
$f_l(1;\alpha)$	1	$\alpha+1$	$2(\alpha+1)$	$(\alpha+1)(\alpha+3)$	$(\alpha+1)(2\alpha+5)$
$K_{1,l}$	8.256	14.51	25.02	42.05	68.10

q	2	3	5	7	11	17	23	53	1009
$B_1(q^{-1})$	8.256	2.794	1.658	1.398	1.222	1.133	1.095	1.039	1.002

A.2. **Coset representatives for U_n/U_{n-1}.** We now give a proof of Lemma 5.13. We assume we are working in odd characteristic.

Proof of Lemma 5.13. The group $U_n(q^2)$ acts transitively on the unitary unit $(n-1)$-sphere, consisting of all column vectors $(x_1, \ldots, x_n)^T$ with entries in $\mathbf{F}_{q^2}$ such that $\sum_{k=1}^n x_k^{1+q} = 1$. The stabilizer of the point $(0, \ldots, 0, 1)^T$ is U_{n-1}. To obtain a factorization of coset representatives according to Lemma 5.13, it suffices to show how to use the inverses of elements of the forms (A), (B), or (C), referred to there, to rotate an arbitrary vector in the unitary sphere onto $(0, \ldots, 0, 1)^T$.

As in the statement of Lemma 5.13, we let N denote the group homomorphism $N(\alpha) = \alpha^{1+q}$. Note that N is an epimorphism onto the group of nonzero elements of the subfield of q elements. We let R be a complete set of coset representatives of $\mathbf{F}_{q^2}^\times / \ker N$.

Now consider an arbitrary element, $x = (x_1, \ldots x_n)^T$ of the unitary unit sphere. If the vector (x_1, x_2) has nonzero unitary norm, then choose an element $s \in R$ such that $N(s) = x_1^{1+q} + x_2^{1+q}$. Hence $(x_1/s, x_2/s)$ is a unit vector and so by the transitivity of U_2 on the unitary 2-sphere, it is clear that we can choose $y \in \mathbf{F}_{q^2}$ and $r \in R$ such that $u_2(r, y)^{-1}$ maps (x_1, x_2) onto a multiple of $(0, 1)$.

In the case where $x_1^{1+q} + x_2^{1+q} = 0$ it is possible that either $x_1^{1+q} + x_3^{1+q}$ or $x_2^{1+q} + x_3^{1+q}$ is nonzero. In the first case multiplying x by t_3 brings the vector into a form where the vector of the first two components has a nonzero norm, and in the second case multiplication by $t_3 \cdot t_2$ achieves this. If none of these three cases holds, then the vector (x_1, x_2, x_3) must be zero (this requires that the characteristic is not 2). Therefore it is always possible to map x onto a vector with zero first component, using the inverse of a matrix of form (A), (B) or (C), provided n is greater than 2.

Now we may apply the same method to map x onto a vector with the first two entries zero, the first three zero, and so on. Finally we obtain a vector with only the last two entries nonzero. Clearly we can use the inverse of an element of form (A) to map this vector onto a vector with only the last entry nonzero. As the vector so obtained is a unit vector, it must have the form $(0,\dots 0,\varepsilon)$ for some ε with $\varepsilon^{1+q}=1$. □

A.3. **Multiplicities of restrictions in Chevalley groups.** Now we shall prove Lemma 5.15 on the multiplicities of restrictions in Chevalley groups. The proof we use was suggested by Tom Hagedorn and follows the line of argument of his thesis [35]. We shall limit ourselves to a brief sketch of this argument. As usual we shall always assume the characteristic is odd.

Proof of Lemma 5.15. Recall that for Chevalley groups $G \geq H \geq K$, we have $\sigma(G,K) = \sigma(G,H) + \sigma(H,K) + \operatorname{rank} H$ where σ is as in Lemma 5.15. On the other hand,

$$\mathcal{M}(G,K) \leq \mathcal{M}(G,H)\mathcal{M}(H,K)\left|\hat{H}\right|$$

where $\hat{H}$ denotes the set of equivalence classes of irreducible representations of H. Thus, we reduce the problem to bounding $\mathcal{M}(G,H)$ for pairs (G,H) of the form $(A_n(q),A_{n-1}(q))$, $(B_n(q),B_{n-1}(q))$, $(C_n(q),C_{n-1}(q))$ and $(D_n(q),D_{n-1}(q))$.

The problem of bounding multiplicities can also be reduced, as follows, to bounding the pairing of a Deligne-Lusztig character of G, restricted to H, with a Deligne-Lusztig character of H: Let us say that a linear combination is bounded if the number of terms may be bounded independently of q and the coefficients may also be bounded independently of q. Then for any irreducible character χ of G (or of H) there is a bounded linear combination of Deligne-Lusztig characters which is the character of a representation containing χ.

We shall now let G and H denote connected reductive algebraic groups of classical type over an algebraically closed field of odd characteristic, and we let F be a Frobenius map. Suppose T and T' are F-stable maximal tori of G and H respectively, and θ and θ' are irreducible characters of T^F and $(T')^F$ respectively. As usual, $R_{T,\theta}$ denotes the Deligne-Lusztig character associated to T and θ (cf. [25] for the complete definitions). Then the pairing of $R_{T,\theta}\downarrow H^F$ with $R_{T',\theta'}$ has the form

$$\left\langle R_{T,\theta}\downarrow H^F, R_{T',\theta'}\right\rangle = \sum_{s}\sum_{u}\sum_{w,w'} a(s,w,w',u)\frac{Q_{T_w}^{C_G^0(s)}(u)Q_{T'_{w'}}^{C_H^0(s)}(u)}{\left|C_{C_H^0(s)^F}(u)\right|}, \tag{56}$$

where s varies over H^F conjugacy classes of elements in $(T')^F$, u varies over unipotent conjugacy classes of the connected centralizer $C_H^0(s)^F$; T_w, $T'_{w'}$ are F-stable maximal tori in $C_G^0(s)$, $C_H^0(s)$ respectively, and the Q's are Green polynomials. For a given s, u, T_w, $T'_{w'}$, the term

$$\frac{Q_{T_w}^{C_G^0(s)}(u)Q_{T'_{w'}}^{C_H^0(s)}(u)}{\left|C_{C_H^0(s)^F}(u)\right|} \tag{57}$$

is a function of q that can be bounded by $O(q^{\sigma(G,H)-\alpha(s,u)})$ where

$$\alpha(s,u) = \frac{1}{2}\left[\dim G - \dim H - \left(\dim C_{C_G^0(s)}(u) - \dim C_{C_H^0(s)}(u)\right)\right]$$

and for any given s the inner summations in (56) are a bounded linear combination of terms of the form (57); $a(s,w,w',u)$ may also be bounded independently of s.

Note that there are only finitely many (a number bounded independently of q) different forms that the term (57) can have given G and H.

Therefore, to obtain a bound for the pairing (56) it suffices to bound $\alpha(s,u)$ and then determine the number of elements s this bound applies to.

To bound $\dim C_{C^0_G(s)}(u) - \dim C_{C^0_H(s)}(u)$ we can reduce to the case where G and H come from one of the series of classical groups: $SL(n)$, $SO(2n+1)$, $Sp(2n)$, or $SO(2n)$. In this case the connected centralizer $C^0_H(s)$ is determined up to isomorphism by the characteristic polynomial of s considered as an element of H; up to isogeny it is simply a product of groups corresponding to different eigenvalues of s. The characteristic polynomial of s considered as an element of G may be obtained from its characteristic polynomial as an element of H by multiplying by either 1 or 2 factors of $(t-1)$; 1 factor in the case of restricting from A_n to A_{n-1} and 2 in the cases of restricting from B_n to B_{n-1}, C_n to C_{n-1} and D_n to D_{n-1}. Hence the centralizer, $C^0_G(s)$, only differs from $C^0_H(s)$ in the factor that corresponds to the eigenvalue 1 of s. Having obtained the form of the centralizers, the formulas in [16], p. 398 (see also the article of Springer and Steinberg in [14]) may be used to compute the dimensions of centralizers of unipotent elements, and hence to bound $\dim C_{C^0_G(s)}(u) - \dim C_{C^0_H(s)}(u)$ in terms of the multiplicity m of 1 as an eigenvalue of s. We call this bound β_m.

Hence we can bound $\alpha(s,u)$ from below by a function α_m of m and the number of s in $(T')^F$ with a given m can be bounded by $O(q^{\gamma_m})$ for some easily determined function γ_m. To prove the theorem we need only verify that $\alpha_m - \gamma_m \geq 0$ for all possible values of m. This is verified by the summary in Table 3.

TABLE 3. VERIFICATION OF LEMMA 5.15.

(G,H)	β_m	α_m	γ_m	$\alpha_m - \gamma_m$	MAXIMUM m
(A_n, A_{n-1})	$2m+1$	$n-m$	$\max\{n-m, 0\}$	1 OR 0	n
(B_n, B_{n-1})	$2m+1$	$2n-m$	$n-1-\frac{m-1}{2}$	$n-\frac{m}{2}-\frac{1}{2}$	$2n-1$
(C_n, C_{n-1})	$2m+3$	$2n-m-2$	$n-1-\frac{m}{2}$	$n-\frac{m}{2}-1$	$2n-2$
(D_n, D_{n-1})	$2m+1$	$2n-m-2$	$n-1-\frac{m}{2}$	$n-\frac{m}{2}-1$	$2n-2$

For the proof to make sense for the pair (D_2, D_1) we have to replace D_1 by a one-dimensional torus. □

REFERENCES

[1] E. Assmus Jr. and J. Key, *Designs and their codes*, Cambridge Univ. Press, Cambridge, 1992. MR **93j:**51003

[2] M. Atkinson, *The complexity of group algebra computations*, Theoret. Comput. Sci. **5** (1977), 205–209. MR **80a:**16020

[3] L. Auslander and R. Tolimieri, *Is computing with the fast Fourier transform pure or applied mathematics?*, Bull. Amer. Math. Soc. (N.S.) **1** (1979), 847–897. MR **81e:**42020

[4] L. Babai, K. Freidl and M. Stricker, *Decomposition of *-closed algebras in polynomial time*, Proc. 18th ACM ISSAC (1993), 86–94.

[5] L. Babai, E. Luks, and A. Seress, *Fast management of permutation groups*, Proc. 28th IEEE FOCS (1988), 272–282.

[6] L. Babai and L. Rónyai, *Computing irreducible representations of finite groups*, Math. Comp. **55** (1990), 705–722. MR **91d:**68057

[7] U. Baum, *Existence and efficient construction of fast Fourier transforms for supersolvable groups*, Comput. Complexity **1** (1991), 235–256. MR **93d:**20031

[8] U. Baum and M. Clausen, *Some lower and upper complexity bounds for generalized Fourier transforms and their inverses*, SIAM J. Comput. **20**(3) (1991), 451–459. MR **91m:**68100

[9] U. Baum, M. Clausen and B. Tietz, *Improved upper complexity bounds for the discrete Fourier transform*, Appl. Algebra Engrg. Comm. Comput. **2** (1991), 35–43. MR **94f:**65129

[10] R. Beals and L. Babai, *Las Vegas algorithms for matrix groups*, Proc. 34$^{\text{th}}$ IEEE FOCS (1993), 427–436.

[11] L. Beckett and P. Diaconis, *Spectral analysis for discrete longitudinal data*, Adv. Math. **103** (1994), 107–128. MR **95d:**62151

[12] T. Beth, *On the computational complexity of the general discrete Fourier transform*, Theoret. Comput. Sci. **51** (1987), 331–339. MR **88i:**20009

[13] ———, *Verfahren der schnellen Fourier-Transformation*, Teubner Studienbücher, Stuttgart, 1984. MR **86g:**65002

[14] A. Borel, R. Carter, C. Curtis, N. Iwahori, T. Springer and R. Steinberg, *Seminar on algebraic groups and related finite groups*, Lecture Notes in Math., vol. 131, Springer-Verlag, New York, 1970. MR **41:**3486

[15] R. Carter, *Simple groups of Lie type*, Wiley-Interscience, New York, 1989. MR **90g:**20001

[16] ———, *Finite groups of Lie type: conjugacy classes and characters*, Wiley-Interscience, New York, 1985. MR **87d:**20060

[17] M. Clausen, *Fast Fourier transforms for metabelian groups*, SIAM J. Comput. **18** (1989), 584–593. MR **90e:**94002

[18] ———, *Fast generalized Fourier transforms*, Theoret. Comput. Sci. **67** (1989), 55–63. MR **91f:**68081

[19] ———, *Beiträge zum Entwurf schneller Spektraltransformationen*, Habilitationsschrift, Fakultät für Informatik der Universität Karlsruhe (TH), 1988.

[20] M. Clausen and U. Baum, *Fast Fourier transforms*, Wissenschaftsverlag, Mannheim, 1993. MR **96i:**68001

[21] ———, *Fast Fourier transforms for symmetric groups, theory and implementation*, Math. Comp. **61** (204) (1993), 833–847. MR **94a:**20028

[22] J. Cooley, *The re-discovery of the fast Fourier transform algorithm*, Mikrochimica Acta **III** (1987), 33–45.

[23] ———, *How the FFT gained acceptance*, Proceedings of the ACM Conference on the History of Scientific and Numeric Computation, Princeton, NJ May 13–15, 1987, 133–140. CMP 93:07

[24] J. Cooley and J. Tukey, *An algorithm for machine calculation of complex Fourier series*, Math. Comp. **19** (1965), 297–301. MR **31:**2843

[25] P. Deligne and G. Lusztig, *Representations of reductive groups over finite fields*, Ann. of Math. **103** (1976), 103–161. MR **52:**14076

[26] P. Diaconis, *A generalization of spectral analysis with applications to ranked data*, Ann. Statist. **17** (1989), 949–979. MR **91a:**60025

[27] ———, *Group representations in probability and statistics*, IMS, Hayward, CA, 1988. MR **90a:**60001

[28] P. Diaconis and D. Rockmore, *Efficient computation of isotypic projections for the symmetric group*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **11**, L. Finkelstein and W. Kantor (eds.), 1993, 87–104. MR **94g:**20022

[29] ———, *Efficient computation of the Fourier transform on finite groups*, J. Amer. Math. Soc. **3**(2) (1990), 297-332. MR **92g:**20024

[30] J. R. Driscoll and D. Healy. *Computing Fourier transforms and convolutions on the 2-sphere*, (extended abstract) Proc. 34$^{\text{th}}$ IEEE FOCS, (1989), 344–349; Adv. in Appl. Math. **15** (1994), 202–250. MR **95h:**65108

[31] D. F. Elliott and K. R. Rao, *Fast transforms: algorithms, analyses, and applications*, Academic Press, New York, 1982. MR **85e:**94001

[32] A. Fässler and E. Stiefel, *Group theoretical methods and their applications*, Birkhäuser, Boston, MA, 1992. MR **93a:**20023

[33] I. M. Gel'fand and M. Tsetlin, *Finite dimensional representations of the group of unimodular matrices*, Dokl. Akad. Nauk SSSR **71** (1950), 825–828 (Russian).

[34] F. Goodman, P. de la Harpe and V. Jones, *Coxeter graphs and towers of algebras*, Springer-Verlag, New York, 1989. MR **91c:**46082

[35] T. Hagedorn, *Multiplicities in restricted representations*, Ph.D. Thesis, Department of Mathematics, Harvard University, 1994.

[36] F. Harary, *Graph theory*, Addison-Wesley, Reading, MA, 1972. MR **41:**1566

[37] D. Healy, D. Maslen, S. Moore, D. Rockmore and M. Taylor, *Applications of a fast convolution algorithm on the 2-sphere*, Technical Report, Department of Mathematics and Computer Science, Dartmouth College, 1994.

[38] M. T. Heideman, D. H. Johnson, and C. S. Burrus, *Gauss and the history of the fast Fourier transform*, Archive for History of Exact Sciences **34**(3) (1985), 265–277. MR **87f:**01018

[39] H. Hiller, *The geometry of coxeter groups*, Research Notes in Mathematics vol. 54, Pitman, 1982. MR **83h:**14045

[40] G. James, *The representation theory of the symmetric groups*, Lecture Notes in Math., vol. 682, Springer-Verlag, New York, 1978. MR **80g:**20019

[41] ———, *Representations of general linear groups*, LMS vol. 94, Cambridge Univ. Press, Cambridge, 1984. MR **86j:**20036

[42] M. Karpovsky, *Fast Fourier transforms on finite nonabelian groups* IEEE Trans. Comput. **26**(10) (1977), 1028–1030. MR **58:**15646

[43] M. Karpovsky and E. Trachtenberg, *Filtering in a communication channel by Fourier transforms over finite groups*, Spectral Techniques and Fault Detection, M. Karpovsky (ed.) Academic Press, NY (1985), 179–212. MR **86i:**94081

[44] A. Kerber, *Representations of permutations groups.* I, II, Lecture Notes in Math., vols. 240, 495, Springer-Verlag, Berlin (1971), (1975). MR **48:**4098; MR **53:**13376

[45] A. Klimyk and N. Vilenkin, *Representations of Lie groups and special functions.* I, II, and III, Mathematics and Its Applications (Soviet Series), vols. 72, 74, 75, Kluwer, Boston, 1991, 1993. MR **93h:**33009; MR **94m:**22001; MR **96f:**33043

[46] J. Lafferty and D. Rockmore, *Fast Fourier analysis for SL_2 over a finite field and related numerical experiments*, J. Exp. Math. **1** (1992) 115–139. MR **94f:**11131

[47] S. Linton, G. Michler, and J. Olsson, *Fourier transforms with respect to monomial representations*, Math. Ann. **297** (1993), 253–268. MR **94i:**20015

[48] D. Maslen, *Efficient computation of Fourier transforms on compact groups*, Max-Planck preprint MPI/95-8 (1995).

[49] D. Maslen and D. Rockmore *Generalized FFTs - a survey of some recent results*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Groups and Computation, II, L. Finkelstein and W. Kantor (eds.), (to appear).

[50] ———, *Separation of variables and the efficient computation of Fourier transforms on finite groups, II*, (in preparation).

[51] ———, *Adapted diameters and the efficient computation of Fourier transforms of finite groups*, Proceedings of the 1995 ACM-SIAM Symposium on Discrete Algorithms, 253–262. MR **95k:**65131

[52] J. Nelder, *The analysis of randomized experiments with orthogonal block structure, I and II*, Proc. of the Royal Acad., Series A **283** (1965), 147–162 and 163–178. MR **31:**848; MR **30:**4363

[53] A. Oppenheim and R. Schafer, *Discrete-time signal processing.* Prentice Hall, NJ, 1989.

[54] D. Rockmore *Some applications of generalized FFTs*, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., Groups and Computation, II, L. Finkelstein and W. Kantor (eds.), (to appear).

[55] ———, *Fast Fourier transforms for wreath products*, Appl. Comput. Harmon. Anal. **4** (1995), 34–55.

[56] ———, *Efficient computation of Fourier inversion for finite groups*, J. Assoc. Comput. Mach. **41**(1) (1994), 31–66. MR **96j:**65153

[57] ———, *Fast Fourier analysis for abelian group extensions*, Adv. in Appl. Math. **11** (1990), 164–204. MR **91k:**68091

[58] J. Serre, *Linear representations of finite groups*, Springer-Verlag, New York, 1977. MR **56:**8675

[59] C. Sims, *Computational methods in the study of permutation groups*, Computational Problems in Abstract Algebra, J. Leech (ed.), Pergamon Press 1970, 169–183. MR **41:**1856

[60] R. Stong, *Some asymptotic results on finite vector spaces*, Adv. in Appl. Math. **9** (1988), 167–199. MR **89c:**05007

[61] E. Thoma, *Die Einschrankung der Charaktere von $GL(n,q)$ auf $GL(n-1,q)$*, Math. Z. **119** (1971), 321–338. MR **44:**5388

[62] R. Tolimieri, M. An, and C. Lu, *Algorithms for discrete Fourier transform and convolution,* Springer-Verlag, New York, 1989. MR **93i:**65131

[63] C. Van Loan, *Computational framework for the fast Fourier transform*, SIAM, Philadelphia, 1992. MR **93a:**65186

[64] G. Wall, *On the conjugacy classes in the unitary, symplectic and orthogonal groups*, J. Austral. Math. Soc. **3** (1963), 1–62. MR **27:**212

[65] A. Willsky, *On the algebraic structure of certain partially observable finite-state Markov processes*, Inform. Contr. **38** (1978), 179–212.

[66] S. Winograd, *Arithmetic complexity of computations*, SIAM CBMS-NSF Series, **33**, SIAM, Philadelphia PA, 1980. MR **81k:**68039

[67] C. M. Woodman. *The symmetry groups of non-rigid molecules as semi-direct products*, Mol. Phys. (6)**19** (1970), 753–780.

[68] A. Zelevinsky. *Representations of finite classical groups*, Lecture Notes in Math., vol. 869, Springer-Verlag, New York, 1981. MR **83k:**20017

Abstract. This paper introduces new techniques for the efficient computation of a Fourier transform on a finite group. We present a divide and conquer approach to the computation. The divide aspect uses factorizations of group elements to reduce the matrix sum of products for the Fourier transform to simpler sums of products. This is the separation of variables algorithm. The conquer aspect is the final computation of matrix products which we perform efficiently using a special form of the matrices. This form arises from the use of subgroup-adapted representations and their structure when evaluated at elements which lie in the centralizers of subgroups in a subgroup chain. We present a detailed analysis of the matrix multiplications arising in the calculation and obtain easy-to-use upper bounds for the complexity of our algorithm in terms of representation theoretic data for the group of interest.

Our algorithm encompasses many of the known examples of fast Fourier transforms. We recover the best known fast transforms for some abelian groups, the symmetric groups and their wreath products, and the classical Weyl groups. Beyond this, we obtain greatly improved upper bounds for the general linear and unitary groups over a finite field, and for the classical Chevalley groups over a finite field. We also apply these techniques to obtain analogous results for homogeneous spaces.

This is part I of a two part paper. Part II will present a refinement of these techniques which yields further savings.

Max-Planck-Institut für Mathematik, 53225 Bonn, Germany
Current address: Department of Mathematics, Universiteit Utrecht, 3584 CD, Utrecht, Netherlands
E-mail address: maslen@math.ruu.nl

Department of Mathematics, Dartmouth College, Hanover, New Hampshire 03755
E-mail address: rockmore@cs.dartmouth.edu