FAST FOURIER TRANSFORMS FOR INVERSE SEMIGROUPS

A Thesis

Submitted to the Faculty

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

in

Mathematics

by

Martin E. Malandro

DARTMOUTH COLLEGE

Hanover, New Hampshire

May 28, 2008

Examining Committee:

(chair) Daniel Rockmore

Peter Doyle

Peter Winkler

Alexander Russell

Charles K. Barlowe, Ph.D.
Dean of Graduate Studies

Abstract

In this thesis we develop a theory of Fourier analysis and fast Fourier transforms (FFTs) for finite inverse semigroups. Our results generalize results in the theory of Fourier analysis for finite groups. There is a general method for generating the irreducible representations of an inverse semigroup, and we use this method to prove that the problem of creating FFTs for inverse semigroups can be reduced to the problems of creating FFTs for their maximal subgroups and creating fast zeta transforms for their poset structures. We then use this result to create FFTs for certain inverse semigroups of interest—in particular, for the rook monoid and its wreath products by arbitrary finite groups. Along the way, we prove a number of results that are important in the theory of Fourier analysis for inverse semigroups. Finally, we use these results to provide an application to the statistical analysis of partially ranked data. Generally speaking, our tools include elements from group and semigroup representation theory, the theory of partially ordered sets and Möbius inversion, and the theory of noncommutative rings.

Acknowledgments

I would first like to thank my advisor, Dan Rockmore, for his constant reassurance and support as I learned how to conduct mathematical research and as I worked through the ideas involved in this thesis. I would also like to thank Mike Orrison, Peter Doyle, and Peter Winkler for their willingness to listen to me as I developed my ideas, for their insights about related problems, and for their ideas for future research. Thanks also to Benjamin Steinberg for meeting with me and speaking with me about his results in semigroup theory, as many of the results in this thesis are built in one way or another from his results. Major thanks to David Webb, who not only taught me algebra and listened to my ideas, but also showed me the beautiful connections between so many fields of mathematics. Finally, I would like to thank all of my friends and family for their love and support as I was conducting this research and writing this thesis.

Contents

List of Tables

Chapter 1

Introduction

Given a complex-valued function f on a finite group G, we may view f as an element of the group algebra $\mathbb{C}G$ by identifying the natural basis of $\mathbb{C}G$ with the characteristic functions of the elements $g \in G$. That is,

$$f = \sum_{g \in G} f(g)\delta_g$$

corresponds to

$$\sum_{g \in G} f(g)g \in \mathbb{C}G.$$

Because $\mathbb{C}G$ is a semisimple algebra, it is the direct sum of its minimal left ideals M_i:

$$\mathbb{C}G = M_1 \oplus \cdots \oplus M_n.$$

By taking a basis for each of the M_i, we obtain a basis for $\mathbb{C}G$ known as a *Fourier basis*. The *Fourier transform* of a function f is then its re-expression in terms of a Fourier basis.

As an example, let $G = \mathbb{Z}/n\mathbb{Z}$, the cyclic group of order n. An element f of the

group algebra $\mathbb{C}\mathbb{Z}/n\mathbb{Z}$ expressed with respect to the natural basis may be viewed as a signal, sampled at n evenly spaced points in time. In this case, the minimal left ideals of $\mathbb{C}\mathbb{Z}/n\mathbb{Z}$ are all 1-dimensional, and hence a Fourier basis is unique (up to scaling factors), and is indeed the usual basis of exponential functions given by the classical discrete Fourier transform. The re-expression of f in terms of a Fourier basis thus corresponds to a re-expression of f in terms of the frequencies that comprise f. This change of basis may be computed efficiently with the help of the classical fast discrete Fourier transform (FFT).

A naive computation the Fourier transform of $f \in \mathbb{C}G$ requires $|G|^2$ operations. An *operation* is defined to be a complex multiplication followed by a complex addition. The problem of efficiently computing the Fourier transform of an arbitrary $\mathbb{C}$-valued function on G has been considered for a wide range of groups G, and efficient algorithms for computing this change of basis now exist for many finite groups. For a survey of these results see, e.g., [12], [22], [23], [24], [25], or [33]. For example, it is known that the Fourier transform of $f \in \mathbb{C}G$ requires no more than:

- $O(n \log n)$ operations if $G = \mathbb{Z}/n\mathbb{Z}$ (see [2] and [8]),
- $O(|S_n| \log^2 |S_n|)$ operations if $G = S_n$, the symmetric group on n elements (see [21]), and
- $O(|B_n| \log^4 |B_n|)$ operations if $G = B_n$, the hyperoctahedral group (that is, the signed symmetric group) on n elements (see [32]).

We shall define a *fast Fourier transform* (FFT) for (or on) a finite group G to be a procedure for calculating the Fourier transform of an arbitrary complex-valued function on G which compares favorably to the naive algorithm. In general, $O(|G| \log^c |G|)$ algorithms are the goal in group FFT theory, although there exist families of groups G for which there exist greatly improved—yet not

$O(|G| \log^c |G|)$—algorithms, such as the family of matrix groups over a finite field [22].

The classical FFT has revolutionized signal processing. Applications include fast waveform smoothing, fast multiplication of large numbers, and efficient waveform compression, to name just a few [3]. FFTs for more general groups have applications in statistical processing [34]. For example, the FFT on $(\mathbb{Z}/2\mathbb{Z})^k$ [40] allows for efficient 2^k-factorial analysis. That is, it allows for the efficient statistical analysis of an experiment in which each of k variables may take on one of two states; e.g., variable k_1 may or may not be present, variable k_2 may be at a high or low intensity, etc. The FFT on S_n allows for an efficient statistical analysis of votes cast in an election involving n candidates [11].

In this thesis, we develop a theory of Fourier transforms for finite inverse semigroups through a study of their algebras. We provide a method for building FFTs on arbitrary inverse semigroups and we construct $O(|S| \log^c |S|)$ FFTs for particular inverse semigroups S of interest. We also give an application of these algorithms to the statistical analysis of partially ranked voting data. This work marks the first extension of group FFTs to non-group semigroups. Our main results are these:

Theorem (Theorem 5.2.7)**.** *Let S be a finite inverse semigroup with $\mathcal{D}$-classes $D_0, \ldots, D_n$. Let r_k denote the number of idempotents in D_k. Choose an idempotent e_k from each $\mathcal{D}$-class D_k, and let G_k be the maximal subgroup of S at e_k. Then the number of operations required to compute the Fourier transform of an arbitrary $\mathbb{C}$-valued function f on S is bounded by*

$$\mathcal{C}(\zeta_S) + \sum_{k=0}^{n} r_k^2 \mathcal{C}(G_k),$$

where $\mathcal{C}(\zeta_S)$ is the maximum number of operations needed to compute the zeta

transform of f on S and $\mathcal{C}(G_k)$ is the maximum number of operations needed to compute the Fourier transform of an arbitrary $\mathbb{C}$-valued function on G_k.

Theorem (Theorems 7.1.1 and 7.2.3)**.** *If $S = R_n$, the rook monoid on n elements, then the Fourier transform of an arbitrary $\mathbb{C}$-valued function on S may be computed in $O(|S| \log^3 |S|)$ operations.*

Theorem (Theorems 8.2.2 and 8.3.4)**.** *If G is a finite group and $S = G \wr R_n$, the wreath product of R_n with G, then the Fourier transform of an arbitrary $\mathbb{C}$-valued function on S may be computed in $O(|S| \log^4 |S|)$ operations.*

We proceed as follows. In Chapter 2, we review important facts about algebras and modules, we state standard definitions and theorems in semigroup theory, and we introduce the most important finite inverse semigroup, the rook monoid. In Chapter 3, we examine bases of inverse semigroup algebras and we define the Fourier transform on an inverse semigroup. In Chapter 4, we review the classical fast Fourier transform, and we examine tools that are useful for building group FFTs and extend them to inverse semigroups. We also review the FFT for the symmetric group and define the computational complexity of the Fourier transform on an inverse semigroup.

Chapter 5 begins with an important result of B. Steinberg, which provides an explicit isomorphism between the inverse semigroup algebra and a direct sum of matrix algebras over group algebras [39]. We show how this result allows us to generate the representations of an inverse semigroup and how this, in turn, allows us to reduce the problem of creating inverse semigroup FFTs to the problems of creating FFTs for their maximal subgroups and fast zeta transforms for their poset structures. This is our most general FFT result. We then use Steinberg's result to

generate the "natural" representations of the rook monoid (first described by C. Grood) directly.

In Chapter 6, we develop a variety of results that are important in the theory of Fourier analysis for inverse semigroups. Our main result in this chapter is a general Fourier inversion theorem for inverse semigroups.

In Chapters 7 through 9, we create explicit FFTs for particular inverse semigroups of interest. Chapter 7 contains an FFT for the rook monoid and Chapter 8 contains an FFT for its wreath products by arbitrary finite groups. Chapter 9 contains another FFT for the rook monoid. While not as efficient as the FFT presented in Chapter 7, it is built in an entirely different way and demonstrates that group-theoretic techniques can be applied directly to the construction of inverse-semigroup FFTs.

We provide an application of the rook monoid FFT to the statistical analysis of partially ranked voting data in Chapter 10. We conclude with thoughts on further research directions in Chapter 11.

Chapter 2

Finite Semigroups and Representations

In this chapter, we review some facts about algebras and modules and we explore some basic notions in semigroup representation theory. Good references are [13], [15], [20], [31], [35], and [36].

2.1 Algebras and Modules

Let F be a field.

Definition (algebra)**.** An *algebra* A over F (or an F-algebra) is an F-vector space in which the elements of A multiply in a way that is compatible with the multiplication of F. Specifically,

$$(ca_1)a_2 = c(a_1a_2) = a_1(ca_2) \text{ for all } c \in F,\ a_1, a_2 \in A.$$

A is said to be a *unital* algebra if there is an identity element $1_A \in A$ for multiplication in A. In this thesis, we specialize to the case $F = \mathbb{C}$. Furthermore,

every algebra that we consider in this thesis will be finite-dimensional.

Definition (algebra homomorphism)**.** An *algebra homomorphism* from an algebra A to an algebra B is a linear map $\phi : A \to B$ satisfying

$$\phi(ab) = \phi(a)\phi(b) \text{ for all } a, b \in A.$$

An algebra *isomorphism* is an algebra homomorphism which is both one-to-one and onto. Algebras A and B are said to be isomorphic if there is an algebra isomorphism between them.

Let $M_n(\mathbb{C})$ be the algebra of $n \times n$ matrices over $\mathbb{C}$ under the usual operations of matrix addition and multiplication.

Definition (algebra representation)**.** A *representation* of an algebra A is an algebra homomorphism

$$\rho : A \to M_n(\mathbb{C}).$$

A representation ρ is said to be *unital* if $\rho(1_A)$ is the identity matrix (provided A has an identity element). If A has no identity element, then every representation of A is unital.

Definition (module)**.** A *left module* over A (or A-module for short) is a $\mathbb{C}$-vector space M in which elements of M may be multiplied on the left by elements in A in a compatible way. Specifically,

$$\begin{aligned}
&a(m_1 + m_2) = am_1 + am_2, \\
&(a_1 + a_2)m = a_1m + a_2m, \\
&c(am) = a(cm) = (ca)m, \text{ and} \\
&(a_1a_2)m = a_1(a_2m)
\end{aligned}$$

for all $c \in \mathbb{C}$, $a, a_1, a_2 \in A$, and $m, m_1, m_2 \in M$.

Every module which we consider in this thesis will be finite-dimensional. M is said to be *unital* if $1_A \in A$ implies $1_A m = m$ for all $m \in M$. If A has no identity, then every A-module is unital.

A particularly important example of an A-module is the algebra A itself (called the *regular module*), where the action of A on itself is given by the usual multiplication in A.

Definition (module homomorphism)**.** A *module homomorphism* from an A-module M to an A-module N is a linear map $\phi : M \to N$ satisfying

$$a\phi(m) = \phi(am) \text{ for all } a \in A,\ m \in M.$$

If ϕ is both one-to-one and onto, then ϕ is a *module isomorphism*. A-modules M and N are isomorphic if there is a module isomorphism between them.

Definition (submodule)**.** A *submodule* M' of an A-module M is a subspace of M which is closed under the action of A. That is, $am' \in M'$ for all $m' \in M'$ and $a \in A$.

Definition (simple module)**.** M is said to be a *simple module* if its only submodules are M and $\{0\}$.

We will be interested in modules that decompose into sums of simple modules:

Definition (semisimple)**.** M is said to be *semisimple* if it is isomorphic to a direct sum of simple modules. The algebra A is said to be semisimple if it is semisimple as an A-module over itself.

In fact, one can show that A is semisimple if and only if every A-module is semisimple [13].

Let M be an A-module, and choose a basis $\{m_1, \ldots, m_n\}$ of M. For any $a \in A$, the action of a on M defines a linear transformation of M, and we may denote the action of a with respect to the basis $\{m_1, \ldots, m_n\}$ by the $n \times n$ matrix $\rho(a)$. This defines a map $\rho : A \to M_n(\mathbb{C})$, and it is straightforward to show that ρ is a representation of A. Conversely, a representation ρ of A gives the underlying vector space (of dimension d_ρ) the structure of an A-module. For this reason, we call A-modules *representations* of A, and so the A-module A is the *regular representation* of A. It is easy to see that an A-module is unital if and only if the associated representation is unital.

Definition (irreducible representation)**.** A representation ρ of A is said to be *irreducible* if the associated A-module is simple and nontrivial, i.e., $\rho \neq 0$ and there is no invertible matrix X for which

$$X\rho X^{-1} = \begin{pmatrix} \rho_1 & 0 \\ m & \rho_2 \end{pmatrix}$$

for some representations ρ_1, ρ_2 of A.

Theorem 2.1.1. *Every irreducible representation of A is unital.*

Proof. Suppose A has identity 1_A. Every A-module M decomposes as

$$M = AM \oplus M',$$

where $M' = \{m - 1_A m : m \in M\}$, and AM is unital. If M is simple, then either $AM = \{0\}$ or $M' = \{0\}$. If M is also nontrivial, then $M' = \{0\}$ and $M = AM$,

and hence M is unital. $\square$

Definition (equivalence of representations)**.** Representations ρ_1, ρ_2 are *equivalent* if their associated modules are isomorphic (i.e., differ only by a change of basis). In other words, ρ_1 and ρ_2 are equivalent if there is a matrix X for which

$$X\rho_1(a)X^{-1} = \rho_2(a) \text{ for all } a \in A.$$

An important result in module theory is *Wedderburn's theorem*, which states that if A is semisimple, then A is isomorphic to a direct sum of matrix algebras over $\mathbb{C}$, and one may compute this isomorphism by "gluing together" a complete set of inequivalent, irreducible representations of A. Formally,

Theorem 2.1.2 (Wedderburn's theorem)**.** *Let A be a finite-dimensional semisimple $\mathbb{C}$-algebra. Let $\mathcal{Y}$ be a complete set of inequivalent, irreducible representations of A. Then $\mathcal{Y}$ is finite, and the map*

$$\bigoplus_{\rho \in \mathcal{Y}} \rho : A \to \bigoplus_{\rho \in \mathcal{Y}} M_{d_\rho}(\mathbb{C})$$

is an isomorphism of algebras.

An interesting corollary is that a semisimple algebra A necessarily has identity, as the inverse image of the identity matrix in the Wedderburn isomorphism is the identity of A.

2.2 Semigroup Representations

A *semigroup* S is a nonempty set together with an associative, binary operation, which we will write multiplicatively. A sub-semigroup of S is a nonempty subset

of S which is itself a semigroup. A subgroup of S is a nonempty subset of S which is itself a group.

If S has an identity element, then S is called a *monoid*. If S does not have an identity element, then we may formally attach one to create a monoid. Let S^1 be S if S has an identity element, and let S^1 be S with an identity formally attached if S does not have an identity element. For the rest of this thesis, S will denote a finite semigroup.

Definition (semigroup homomorphism)**.** A semigroup *homomorphism* from a semigroup S to a semigroup T is a map $\phi : S \to T$ which satisfies

$$\phi(s_1 s_2) = \phi(s_1)\phi(s_2) \text{ for all } s_1, s_2 \in S.$$

ϕ is an *isomorphism* if it is both one-to-one and onto, and S and T are said to be isomorphic if there exists an isomorphism between them.

Definition (semigroup representation)**.** A *representation* of S (over $\mathbb{C}$) is a homomorphism ρ from S to the semigroup of $d_\rho \times d_\rho$ matrices with entries in $\mathbb{C}$ under multiplication.

Remark: If S is a group, this definition does not agree with the usual group definition, which requires that ρ associate the identity element of S to the identity matrix. Such representations are called *unital*. However, there is only a trivial difference between a representation and a unital representation of S^1: a representation is called *null* if $\rho(s) = 0$ for all $s \in S$, and every representation of S^1 is either unital, null, or a direct sum of a unital and a null representation ([31], Fact 1.10).

Definition (semigroup algebra)**.** Let S be a finite semigroup. The *semigroup algebra* of S over $\mathbb{C}$, denoted $\mathbb{C}S$, is the formal $\mathbb{C}$-span of the symbols $\{s\}_{s\in S}$. Multiplication in $\mathbb{C}S$, denoted by $*$, is given by convolution (i.e., the linear extension of the semigroup operation via the distributive law): Suppose $f, g \in \mathbb{C}S$, with

$$f = \sum_{r\in S} f(r)r, \quad g = \sum_{t\in S} g(t)t.$$

Then

$$f * g = \sum_{r\in S} f(r)r \sum_{t\in S} g(t)t = \sum_{s\in S} \sum_{r,t\in S:rt=s} f(r)g(t)s.$$

Remark: If S is a group, then convolution may be written in the familiar way:

$$f * g = \sum_{s\in S} \sum_{r\in S} f(r)g(r^{-1}s)s.$$

The basis $\{s\}_{s\in S}$ of the semigroup algebra $\mathbb{C}S$ is called the *semigroup basis*. It is one of two "natural" bases of the semigroup algebra. We will define the other natural basis in Section 3.2.

Let S be a finite semigroup. Any semigroup representation of S extends linearly to a representation of its algebra, and any representation of a semigroup algebra yields a representation of the underlying semigroup by restricting to the semigroup basis. In this way, representations of semigroups and representations of their algebras are in one-to-one correspondence. Semigroup representations are said to be *irreducible* (resp. *inequivalent*) if their associated algebra representations are irreducible (resp. inequivalent). When S has an identity, it is immediate that a representation of S is unital if and only if the associated algebra representation is unital.

2.3 Inverse Semigroups and Algebras

In this thesis, we are concerned with a particular class of semigroups known as *inverse semigroups*.

Definition (inverse semigroup)**.** An *inverse semigroup* is a semigroup S such that, for each $x \in S$, there is a *unique* $y \in S$ such that

$$xyx = x \text{ and } yxy = y.$$

In this case, we write $y = x^{-1}$.

We remark that the condition that y be unique is necessary for this definition. An element $x \in S$ is said to be *regular* or *Von-Neumann regular* if there is at least one $y \in S$ satisfying $xyx = x$ and $yxy = y$, and S is said to be *regular* if every element of S is regular. Consider the full transformation semigroup X on the set $\{1, 2, \ldots, n\}$; that is, all maps from $\{1, 2, \ldots, n\}$ to itself under composition. It is easy to see that X is regular, and that (for $n \geq 2$) there exist elements $x \in X$ for which there are *multiple* elements $y \in X$ satisfying $xyx = x$ and $yxy = y$. X is therefore not inverse. An equivalent characterization of inverse semigroups (see, e.g., [20]) is as follows.

Definition (inverse semigroup)**.** An *inverse semigroup* is a semigroup S which is regular and for which all idempotents of S commute.

Let S be a finite inverse semigroup. In [29] (Theorem 4.4), Munn proves that the semigroup algebra $\mathbb{C}S$ is semisimple, and we therefore have:

Theorem 2.3.1 (decomposing representations)**.** *Any representation of S (resp. $\mathbb{C}S$) is equivalent to a direct sum of irreducible and null representations of S (resp. $\mathbb{C}S$).*

Furthermore, Wedderburn's theorem applies to $\mathbb{C}S$:

Theorem 2.3.2. *Let S be a finite inverse semigroup. Let $\mathcal{Y}$ be a complete set of inequivalent, irreducible representations of $\mathbb{C}S$. Then $\mathcal{Y}$ is finite, and the map*

$$\bigoplus_{\rho\in\mathcal{Y}} \rho : \mathbb{C}S \to \bigoplus_{\rho\in\mathcal{Y}} M_{d_\rho}(\mathbb{C}) \tag{2.1}$$

is an isomorphism of algebras. Explicitly, let $f \in \mathbb{C}S$ where $f = \sum_{s\in S} f(s)s$. Then

$$f \mapsto \bigoplus_{\rho\in\mathcal{Y}} \sum_{s\in S} f(s)\rho(s).$$

We also have the following formula for the sum of the squares of the dimensions of the irreducible representations of S:

Corollary 2.3.3. *Let S be a finite inverse semigroup, and let $\mathcal{Y}$ be a complete set of inequivalent, irreducible representations of S. Then*

$$|S| = \sum_{\rho\in\mathcal{Y}} d_\rho^2. \tag{2.2}$$

Proof. The formula (2.2) is just the $\mathbb{C}$-dimensionality of the algebras appearing in (2.1). □

2.4 The Rook Monoid

The most important example of an inverse semigroup is the *rook monoid* (also called the *symmetric inverse semigroup*) on n elements, which we denote by R_n. It is the semigroup of all injective partial functions from $\{1, \ldots, n\}$ to itself under the operation of partial function composition. In this thesis, we adopt the convention

that maps act on the left of sets, and so, for $g, f \in R_n$, $g \circ f$ is defined for precisely the elements x for which $x \in \mathrm{dom}(f)$ and $f(x) \in \mathrm{dom}(g)$. R_n is called the rook monoid because it is isomorphic to the semigroup of all $n \times n$ matrices with the property that at most one entry in each row is 1 and at most one entry in each column is 1 (the rest being 0) under multiplication. Such matrices, called *rook matrices*, correspond to the set of all possible placements of non-attacking rooks on an $n \times n$ chessboard. For example, consider the element $\sigma \in R_4$ defined by

$$\sigma(2) = 1, \quad \sigma(4) = 4.$$

Then, viewed as a partial permutation, σ is

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ - & 1 & - & 4 \end{pmatrix}$$

where the dash indicates that the above entry is not mapped to anything. As a rook matrix, we have

$$\sigma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

We quickly prove that R_n is indeed an inverse semigroup:

Theorem 2.4.1. *R_n is an inverse semigroup.*

Proof. For an element $\sigma \in R_n$, define $\gamma \in R_n$ by

$$\mathrm{dom}(\gamma) = \mathrm{ran}(\sigma), \text{ and, for } x \in \mathrm{dom}(\gamma), \ \gamma(x) = \sigma^{-1}(x).$$

It is easy to see that $\gamma\sigma\gamma = \gamma$ and $\sigma\gamma\sigma = \sigma$, and that γ is the only element of R_n satisfying both equations. $\square$

The rook monoid is a generalization of the symmetric group, and it plays the same role for finite inverse semigroups as the symmetric group does for finite groups in this variation of Cayley's theorem (see, e.g. [20], p. 36-37):

Theorem 2.4.2. *Let S be a finite inverse semigroup. Then there exists $n \in \mathbb{Z}$ for which S is isomorphic to a sub-semigroup of R_n.*

It is often useful to consider subsets of R_n whose elements have domains of a certain size:

Definition (rank)**.** Given an element $\sigma \in R_n$, the *rank* of σ, denoted $\mathrm{rk}(\sigma)$, is defined to be $\mathrm{rk}(\sigma) = |\mathrm{dom}(\sigma)| = |\mathrm{ran}(\sigma)|$. It is clear that the rank of σ is the same as the rank of the associated rook matrix.

The elements of R_n are divided into two classes, the elements of rank n (i.e., the permutations) and the elements of rank less than n. In his analysis of the representation theory of R_n, Munn [29] introduced what he called *cycle-link notation* for the elements of R_n. As an example, consider the element $\sigma \in R_4$

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & - & 2 & 4 \end{pmatrix}.$$

In cycle-link notation, a cycle $(a_1, a_2, \ldots, a_k)$ means that

$$a_1 \mapsto a_2, a_2 \mapsto a_3, \ldots, a_{k-1} \mapsto a_k, \text{ and } a_k \mapsto a_1,$$

and a link $[b_1, b_2, \ldots, b_k]$ means that

$$b_1 \mapsto b_2, b_2 \mapsto b_3, \ldots, b_{k-1} \mapsto b_k, \text{ and } b_k \text{ goes nowhere.}$$

The element σ, expressed in cycle-link notation, would be $[1, 3, 2](4)$. Note that the cycle-link representation of a permutation always consists of cycles only, and elements of rank less than n always contain links when written in cycle-link notation.

We have the following theorem on the size of R_n:

Theorem 2.4.3.

$$|R_n| = \sum_{k=0}^{n} \binom{n}{k}^2 k!$$

Proof. For any particular rank k, there are $\binom{n}{k}$ choices for the domain and $\binom{n}{k}$ choices for the range of an element of R_n, and for any particular choice of domain and range, there are $k!$ ways of mapping the domain to the range. □

We also have the recursive formula

Theorem 2.4.4. *For $n \geq 3$,*

$$|R_n| = 2n|R_{n-1}| - (n-1)^2|R_{n-2}|.$$

Proof. See Section 7.2. □

Chapter 3

Fourier Transforms for Inverse Semigroups

In this chapter, we review the poset structure of an inverse semigroup, define natural and Fourier bases for inverse semigroup algebras, and define the Fourier transform of a function on an inverse semigroup. Good references are [13], [23], [37], and [39].

3.1 The Poset Structure of an Inverse Semigroup

Definition (poset structure of S)**.** Let S be a finite inverse semigroup. For $s, t \in S$, define

$$\begin{aligned} s \leq t &\iff s = et \text{ for some idempotent } e \in S \\ &\iff s = tf \text{ for some idempotent } f \in S. \end{aligned}$$

For R_n, the idempotents are the restrictions of the identity map, and this

ordering is the same as the ordering:

$$s \leq t \iff t \text{ extends } s \text{ as a partial function.}$$

Remark: If S is a group, then its poset structure is trivial.

If P is a finite poset, then the *zeta function* ζ of P is given by:

$$\zeta : P \times P \to \{0, 1\}$$

$$\zeta(x, y) = \begin{cases} 1 & \text{if } x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

The zeta function is invertible over any field F (in fact, over any ring), and its inverse is called the *Möbius function* μ. The Möbius function for R_n over $\mathbb{C}$ is well known ([37], [39]). It is, for $x \leq y$,

$$\mu(x, y) = (-1)^{\mathrm{rk}(y)-\mathrm{rk}(x)}.$$

3.2 Natural Bases for Inverse Semigroup Algebras

Let S be a finite inverse semigroup. We have already seen one natural basis for the semigroup algebra $\mathbb{C}S$, the semigroup basis $\{s\}_{s \in S}$. Multiplication in $\mathbb{C}S$ with respect to this basis is just the linear extension of the multiplication in S. We claim that there is another "natural" basis of $\mathbb{C}S$. To motivate this new basis, recall that every finite inverse semigroup is isomorphic to a sub-semigroup of a rook monoid and can therefore be viewed as a collection of partial functions. There is another model for composing partial functions: only allow the composition if the range of

the first function "lines up" with the domain of the second. For example, if

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & - & 1 & - \end{pmatrix}, \qquad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & - & - \end{pmatrix},$$

then the idea is that the composition $\pi \circ \sigma$ is

$$\pi \circ \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & - & 4 & - \end{pmatrix},$$

and the composition $\sigma \circ \pi$ is disallowed. The *groupoid basis* of $\mathbb{C}S$ encodes this.

Definition (groupoid basis)**.** Define, for each $s \in S$, the element $\lfloor s \rfloor \in \mathbb{C}S$ by

$$\lfloor s \rfloor = \sum_{t \in S : t \leq s} \mu(t, s) t.$$

Theorem 3.2.1. *The collection $\{\lfloor s \rfloor\}_{s \in S}$ is a basis for $\mathbb{C}S$. Multiplication in $\mathbb{C}S$ relative to this basis is given by the linear extension of*

$$\lfloor s \rfloor \lfloor t \rfloor = \begin{cases} \lfloor st \rfloor & \text{if } \mathrm{dom}(s) = \mathrm{ran}(t), \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, the change of basis to the $\{s\}_{s \in S}$ basis of $\mathbb{C}S$ is given by Möbius inversion:

$$s = \sum_{t \in S : t \leq s} \lfloor t \rfloor . \tag{3.1}$$

Proof. This is [39], Lemma 4.1 and Theorem 4.2, using our convention that maps act on the left of sets. □

The notions of $\mathrm{dom}(s)$ and $\mathrm{ran}(s)$ may also be defined intrinsically in terms

of S, i.e., without reference to an embedding of S into R_n. Specifically, for any element $s \in S$, ss^{-1} and $s^{-1}s$ are idempotent, and one may define

$$\mathrm{dom}(s) = s^{-1}s$$
$$\mathrm{ran}(s) = ss^{-1}.$$

If we use this definition for $s \in R_n$, we see that $\mathrm{dom}(s)$ is actually *not* the domain of s, but is rather the map which is the identity on the domain of s and undefined elsewhere, and likewise for $\mathrm{ran}(s)$. This means that we are abusing the distinction between the domain and range of a map and the corresponding partial identities. Under this definition, we have that the groupoid basis multiplies as follows:

$$\lfloor s \rfloor \lfloor t \rfloor = \begin{cases} \lfloor st \rfloor & \text{if } s^{-1}s = tt^{-1}, \\ 0 & \text{otherwise.} \end{cases}$$

The viewpoint then is that we have two "natural bases" for $\mathbb{C}S$, the semigroup basis $\{s\}_{s \in S}$ and the groupoid basis $\{\lfloor s \rfloor\}_{s \in S}$. Note that if S is a group, then $s = \lfloor s \rfloor \in \mathbb{C}S$ for all $s \in S$, so group algebras only have one natural basis.

3.3 Fourier Bases for Inverse Semigroup Algebras

Let S be a finite inverse semigroup. $\mathbb{C}S$ is semisimple, so $\mathbb{C}S$ decomposes into a direct sum of simple (i.e., irreducible) submodules:

$$\mathbb{C}S = \bigoplus L_i.$$

Let $\mathcal{Y}$ be a complete set of inequivalent, irreducible matrix representations of $\mathbb{C}S$. Then, according to Theorem 2.1.2,

$$\bigoplus_{\rho\in\mathcal{Y}} \rho : \mathbb{C}S \to \bigoplus_{\rho\in\mathcal{Y}} M_{d_\rho}(\mathbb{C})$$

is an isomorphism of algebras.

There is a natural basis for the algebra on the right: the set of matrices in the algebra with the property that exactly one entry is 1 (the rest being 0). The inverse image of this set is a basis for $\mathbb{C}S$ called the *dual matrix coefficient basis for* $\mathcal{Y}$, or the *Fourier basis for* $\mathbb{C}S$ *according to* $\mathcal{Y}$. When we refer to a Fourier basis for $\mathbb{C}S$, we mean any basis of $\mathbb{C}S$ that can arise in this manner by choosing an appropriate $\mathcal{Y}$. Note that there is a unique Fourier basis for $\mathbb{C}S$ (up to ordering) if and only if every irreducible representation of S has dimension 1.

Consider the natural basis for the algebra on the right. The preimage of a single column of these elements from the ρ^{th} block is a basis B for a submodule of $\mathbb{C}S$, and each element of $\mathbb{C}S$ acts on B exactly as described by ρ. We therefore have that B is a basis for an irreducible submodule of $\mathbb{C}S$ (isomorphic to the $\mathbb{C}S$-module described by ρ). Since the map above was an isomorphism, the preimages of distinct columns have intersection $\{0\}$, and we therefore have the well-known fact:

Theorem 3.3.1. *Each irreducible submodule of* $\mathbb{C}S$ *occurs in the decomposition of* $\mathbb{C}S$ *into irreducibles exactly as many times as its dimension.*

We therefore also have that a Fourier basis for $\mathbb{C}S$ is a basis which realizes the decomposition of $\mathbb{C}S$ into irreducible submodules L_i.

We are now ready to define the Fourier transform on an inverse semigroup.

3.4 The Fourier Transform on an Inverse Semigroup

Let S be a finite inverse semigroup. The natural basis of the space V of complex-valued functions on S is denoted by $\{\delta_s\}_{s\in S}$. That is, for $f \in V$,

$$f = \sum_{s\in S} f(s)\delta_s.$$

We would like to associate this basis to one of the natural bases of the semigroup algebra $\mathbb{C}S$, and then define the Fourier transform on S as the change of basis in $\mathbb{C}S$ (and hence in V, by association) to a Fourier basis. At present, it is not clear which association is preferable, but as we shall see in Chapter 10, both associations are useful, so we shall not discard either one.

Definition (semigroup association)**.** Under the *semigroup basis association*, we associate, for $s \in S$,

$$s \in \mathbb{C}S \longleftrightarrow \delta_s \in V,$$

and so, for $f = \sum_{s\in S} f(s)\delta_s \in V$, $f \in \mathbb{C}S$ is written as

$$f = \sum_{s\in S} f(s)s.$$

Definition (groupoid association)**.** Under the *groupoid basis association*, we associate, for $s \in S$,

$$\lfloor s \rfloor \in \mathbb{C}S \longleftrightarrow \delta_s \in V,$$

and so, for $f = \sum_{s \in S} f(s)\delta_s \in V$, $f \in \mathbb{C}S$ is written as

$$f = \sum_{s \in S} f(s) \lfloor s \rfloor .$$

These associations are **not** compatible unless the poset structure of S is trivial (for example, if S is a group). Now, let $\mathcal{Y}$ be a complete set of inequivalent, irreducible representations of $\mathbb{C}S$.

Definition. The isomorphism

$$\mathbb{C}S \cong \bigoplus_{\rho \in \mathcal{Y}} M_{d_\rho}(\mathbb{C}), \text{ where } f \in \mathbb{C}S \mapsto \bigoplus_{\rho \in \mathcal{Y}} \rho(f) \tag{3.2}$$

is called a *Fourier transform* on $\mathbb{C}S$.

Definition. Let $f \in \mathbb{C}S$. The *Fourier transform* of f is the image of f in the matrix algebra

$$\bigoplus_{\rho \in \mathcal{Y}} M_{d_\rho}(\mathbb{C})$$

via the isomorphism (3.2). Equivalently, the Fourier transform of f is the re-expression of f in $\mathbb{C}S$ in terms of a Fourier basis for $\mathbb{C}S$.

Let $f \in \mathbb{C}S$ be given with respect to one of the natural bases, i.e., either $f = \sum_{s \in S} f(s)s$ or $f = \sum_{s \in S} f(s) \lfloor s \rfloor$. We shall sometimes say "calculating the Fourier transform on $\mathbb{C}S$" to mean calculating the Fourier transform of an arbitrary element $f \in \mathbb{C}S$ given with respect to one of the natural bases, where the choice of $\mathcal{Y}$ is understood. We now define the Fourier transform of f at a representation of $\mathbb{C}S$:

Definition. Let ρ be a representation of $\mathbb{C}S$. Define

$$\hat{f}(\rho) = \rho(f) = \begin{cases} \sum_{s \in S} f(s)\rho(s) & \text{if } f = \sum_{s \in S} f(s)s, \\ \sum_{s \in S} f(s)\rho(\lfloor s \rfloor) & \text{if } f = \sum_{s \in S} f(s) \lfloor s \rfloor . \end{cases}$$

For $\rho \in \mathcal{Y}$, $\hat{f}(\rho)$ is therefore just the ρ^{th} block in the image of f in the isomorphism (3.2). Thus, we can write

$$\mathbb{C}S \cong \bigoplus_{\rho \in \mathcal{Y}} M_{d_\rho}(\mathbb{C}), \text{ where } f \in \mathbb{C}S \mapsto \bigoplus_{\rho \in \mathcal{Y}} \hat{f}(\rho).$$

Remark: Since the map in (3.2) is an isomorphism, it respects multiplication, and hence the Fourier transform turns convolution of elements of $\mathbb{C}S$ into multiplication of block-diagonal matrices. It turns convolution into pointwise multiplication if and only if all irreducible representations of S have degree one, which, for example, is the case for $S = \mathbb{Z}/n\mathbb{Z}$.

Chapter 4

Fast Fourier Transforms

In this chapter, we show how the classical discrete Fourier transform fits into our general framework and we revisit tools for computing FFTs on non-Abelian groups and extend them to semigroups. In particular, we explain Clausen's algorithm for the FFT on the symmetric group and explore complexity considerations for inverse semigroups. Good references are [4], [18], [21], [22], [23], and [25].

4.1 The Discrete Fourier Transform

The classical discrete Fourier transform is defined as follows:

Definition. Let $f(0), \ldots, f(n-1)$ be a sequence of n complex numbers. The *discrete Fourier transform* of this sequence is the sequence $\hat{f}(0), \ldots, \hat{f}(n-1)$, where

$$\hat{f}(k) = \sum_{t=0}^{n-1} f(t) e^{\frac{-2\pi i k t}{n}}.$$

For $t \in \{0, \ldots, n-1\}$, we then have

$$f(t) = \sum_{k=0}^{n-1} \frac{1}{n} \hat{f}(k) e^{\frac{2\pi i k t}{n}}.$$

As motivation, suppose that $\{f(0), f(\frac{1}{n}), f(\frac{2}{n}), \ldots, f(\frac{n-1}{n})\}$ is a collection of n equispaced samples of a continuous waveform f on $[0,1]$. Suppose further that f is band-limited of bandwidth n, so that its Fourier expansion is

$$f(t) = \sum_{k=0}^{n-1} \hat{f}(k) e^{2\pi i k t}$$

for $0 \leq t < 1$. The Fourier coefficients $\hat{f}(k)$ are given by the integrals

$$\hat{f}(k) = \int_0^1 f(t) e^{-2\pi i k t} dt. \tag{4.1}$$

The assumption that f is bandlimited allows us to compute the coefficients $\hat{f}(k)$ by finite sums which discretize the integrals (4.1). In particular, it is well known that if f is band-limited of bandwidth n, then the following formula for $\hat{f}(k)$ holds (see, e.g., [34]):

$$\hat{f}(k) = \frac{1}{n} \sum_{t=0}^{n-1} f(\frac{t}{n}) e^{\frac{-2\pi i k t}{n}}.$$

Thus, it is possible to reconstruct the original waveform f from the finite collection of equispaced samples $\{f(\frac{t}{n})\}_{t=0}^{n-1}$, in the sense that this information is sufficient to compute its Fourier coefficients. Notice now that the collection $\{\hat{f}(k)\}_{k=0}^{n-1}$ appearing here is just the scaled (by a factor of n) output of the discrete Fourier transform of the sequence $f(0), f(\frac{1}{n}), \ldots, f(\frac{n-1}{n})$. Thus, the discrete Fourier transform of a sampled band-limited waveform allows us to express the underlying waveform in terms of the frequencies of which it is comprised.

Now, let $\mathbb{Z}/n\mathbb{Z}$ denote the cyclic group of order n. It is well-known that all of the complex irreducible representations of $\mathbb{Z}/n\mathbb{Z}$ are 1-dimensional, so they can be put into one-to-one correspondence with the elements of $\mathbb{Z}/n\mathbb{Z}$. In fact, the following theorem is immediate:

Theorem 4.1.1. *Let* $\{\phi_0, \ldots, \phi_{n-1}\}$ *be the sequence of representations defined by:*

$$\phi_k(t) = e^{\frac{-2\pi ikt}{n}}.$$

Then $\phi_k(t_1 + t_2) = \phi_k(t_1)\phi_k(t_2)$ *for all* $t_1, t_2 \in \mathbb{Z}/n\mathbb{Z}$, *and* $\{\phi_0, \ldots, \phi_{n-1}\}$ *is a complete set of inequivalent, irreducible representations of* $\mathbb{C}\mathbb{Z}/n\mathbb{Z}$.

Let $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{C}$. The group algebra $\mathbb{C}\mathbb{Z}/n\mathbb{Z}$ has only one natural basis, so $f \in \mathbb{C}\mathbb{Z}/n\mathbb{Z}$ by

$$f = \sum_{t=0}^{n-1} f(t)t,$$

and it is immediate that our definition of $\hat{f}(\phi_k)$ is the same as the coefficient $\hat{f}(k)$ from the discrete Fourier transform. The Fourier transform of f then expresses f in terms of the Fourier basis of $\mathbb{C}\mathbb{Z}/n\mathbb{Z}$, i.e., the basis $\{b_k\}_{k=0}^{n-1}$, where

$$b_k = \frac{1}{n} \sum_{t=0}^{n-1} e^{\frac{2\pi ikt}{n}} t,$$

and we see from this that the Fourier basis of $\mathbb{C}\mathbb{Z}/n\mathbb{Z}$ is the usual basis of exponential functions.

Calculating the Fourier transform of an arbitrary function $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{C}$ by naive methods requires $n^2 = |\mathbb{Z}/n\mathbb{Z}|^2$ operations, where an *operation* is a complex multiplication together with a complex addition. As n grows large, this cost becomes prohibitive, and fast algorithms are therefore necessary.

Here is the idea behind the famous *Cooley-Tukey FFT* (i.e., the classical fast discrete Fourier transform), expressed in the language of groups (see also [8] and [23]). Let $f : \mathbb{Z}/n\mathbb{Z} \to \mathbb{C}$. Suppose n is composite, $n = pq$ with p prime. Then

$\mathbb{Z}/n\mathbb{Z}$ has a subgroup H isomorphic to $\mathbb{Z}/q\mathbb{Z}$:

$$H = \{px : x \in \{0, \ldots, q-1\}\}.$$

H then partitions $\mathbb{Z}/n\mathbb{Z}$ into cosets; $\mathbb{Z}/n\mathbb{Z} = \cup_{j=0}^{p-1} j + H$. For all $k \in \mathbb{Z}/n\mathbb{Z}$, we would like to compute

$$\hat{f}(k) = \hat{f}(\phi_k) = \sum_{t=0}^{n-1} f(t)\phi_k(t).$$

Since ϕ_k is a homomorphism, by our partitioning we may rewrite this as

$$\sum_{t=0}^{n-1} f(t)\phi_k(t) = \sum_{j=0}^{p-1} \phi_k(j) \sum_{h \in H} f_j(h)\phi_k(h), \tag{4.2}$$

where $f_j(h) = f(j+h)$.

The point is that the inner sums in (4.2) are themselves Fourier transforms on H. To compute the Fourier transform on $\mathbb{Z}/n\mathbb{Z}$, then, we may proceed by calculating the inner sums (by computing p Fourier transforms on $H \cong \mathbb{Z}/q\mathbb{Z}$) and then multiplying the results by the $\phi_k(j)$ and adding them up.

For any group G, let $\mathcal{C}(G)$ denote the minimal number of operations required to compute the Fourier transform of an arbitrary $\mathbb{C}$-valued function on any group isomorphic to G. Then we have

$$\mathcal{C}(\mathbb{Z}/n\mathbb{Z}) \leq p\mathcal{C}(\mathbb{Z}/q\mathbb{Z}) + np.$$

Since the inner sums in (4.2) are themselves Fourier transforms, if q is composite, then we can iterate this algorithm to amplify our savings. This algorithm is most efficient when n is a power of 2, say $n = 2^k$. In this case, we have:

Theorem 4.1.2 (Cooley-Tukey [8])**.**

$$\mathcal{C}(\mathbb{Z}/2^k\mathbb{Z}) \leq 2^k + k2^{k+1} = O(n \log n).$$

4.2 Representations and Schur's Lemma

The reduction argument involved in the Cooley-Tukey FFT is also a central idea in the construction of non-Abelian group and semigroup FFTs. However, more tools are necessary to create these more general FFTs. In particular, we need to understand how multidimensional irreducible representations behave when restricted to sub-semigroups. Let $X_0 < X_1 < \cdots < X_n$ be a chain of finite semigroups whose semigroup algebras $\mathbb{C}X_i$ are semisimple. Even if a representation of X_i is irreducible, its restriction to X_{i-1} typically will not be. However, by the semisimplicity of the algebras involved, it will be equivalent to (though not necessarily *equal* to) a direct sum of irreducible and null representations of X_{i-1}.

Definition (adapted representations)**.** Let $\mathcal{Y}_i$ be a set of inequivalent, irreducible representations for $\mathbb{C}X_i$. The collection $\{\mathcal{Y}_i\}_{i=0}^n$ is said to be *chain-adapted to the chain* $X_0 < X_1 < \cdots < X_n$ if, for every $i \geq 1$ and every $\rho \in \mathcal{Y}_i$, $\rho|_{X_{i-1}}$ is *equal* to a direct sum of null representations and irreducible representations in $\mathcal{Y}_{i-1}$. Note that this definition forces the irreducible representations appearing in such a restriction to be equal whenever they are equivalent.

Induction shows that a representation from a chain-adapted set of representations may be restricted further down the chain with the same equality results. In particular, if $\{\mathcal{Y}_i\}_{i=0}^n$ is chain-adapted to $X_0 < X_1 < \cdots < X_n$, then $\{\mathcal{Y}_j\}_{j\in J}$ is chain adapted to $<_{j\in J} X_j$ for any nonempty subset $J \subseteq \{1, \ldots, n\}$.

We shall often simply ask that a complete set of irreducible representations $\mathcal{Y}_n$ for X_n be adapted to the chain $X_0 < X_1 < \cdots < X_n$. In this case, the choices of the $\mathcal{Y}_i$ are understood (and, in fact, are often completely determined) by the choice of $\mathcal{Y}_n$.

Remark: While adapted representations are very important in the construction of FFTs, the requirement that a set of representations be adapted is in no way limiting. Given any chain $X_0 < X_1 < \cdots < X_n$ of semigroups whose semigroup algebras $\mathbb{C}X_i$ are semisimple, a straightforward induction argument shows that a complete set of inequivalent, irreducible, chain-adapted matrix representations always exists. Explicitly describing such sets, however, is frequently a challenging endeavor.

Remark: Chain-adapted representations are also sometimes referred to as *semi-normal* or *Gelfand-Tsetlin* representations ([16], [21]).

One reason that adapted sets of representations are useful in the construction of FFTs is that they are sparse and structured when evaluated at certain elements. The specifics are given by the following computational version of Schur's Lemma (Lemma 4.2.2). While Schur's Lemma appears in many forms, the most widely recognized is probably this:

Lemma 4.2.1 (Schur's Lemma)**.** *Let R be a ring and let M, N be simple R-modules. Then every R-module homomorphism $\phi : M \to N$ is either 0 or an isomorphism.*

Proof. See, e.g. [13], p. 30-32. □

Here is the computational version of Schur's Lemma. This result is an extension

of the version of Schur's Lemma in [22]. We derive it as a consequence of Lemma 4.2.1.

Lemma 4.2.2 (Schur's Lemma)**.** *Suppose that $A < B < C$ are finite semigroups, that $\mathcal{Y}_A$, $\mathcal{Y}_B$, and $\mathcal{Y}_C$ are complete sets of inequivalent, irreducible matrix representations for A, B, C respectively, adapted to the chain, and that $\mathbb{C}A, \mathbb{C}B$, and $\mathbb{C}C$ are semisimple. Let $\rho \in \mathcal{Y}_C$. Say $\rho|_B = \rho^1 \oplus \cdots \oplus \rho^k$ (with each ρ^j either in $\mathcal{Y}_B$ or, without loss of generality, null of dimension 1) and $\rho^j|_A = \rho^j_1 \oplus \cdots \oplus \rho^j_{g(j)}$ (with each ρ^j_i either in $\mathcal{Y}_A$ or, again without loss of generality, null of dimension 1). Let $\sigma \in B$ such that σ commutes with A. Then $\rho(\sigma)$ is a block matrix*

$$\rho(\sigma) = \begin{pmatrix} \rho^1(\sigma) & 0 & 0 & \cdots & 0 \\ 0 & \rho^2(\sigma) & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \rho^k(\sigma) \end{pmatrix}$$

where, for $1 \leq j \leq k$, the block $\rho^j(\sigma)$ is itself a block matrix, with blocks indexed by all ordered pairs from $\{\rho^j_1, \ldots, \rho^j_{g(j)}\}$:

$$\rho^j(\sigma) = \begin{array}{c} \\ \rho^j_1 \\ \rho^j_2 \\ \vdots \\ \rho^j_{g(j)} \end{array} \begin{array}{c} \begin{array}{cccc} \rho^j_1 & \rho^j_2 & \cdots & \rho^j_{g(j)} \end{array} \\ \begin{pmatrix} \lambda^j_{1,1} I & \lambda^j_{1,2} I & \cdots & \lambda^j_{1,g(j)} I \\ \lambda^j_{2,1} I & \lambda^j_{2,2} & \cdots & \lambda^j_{2,g(j)} I \\ \vdots & \vdots & \ddots & \vdots \\ \lambda^j_{g(j),1} I & \lambda^j_{g(j),2} I & \cdots & \lambda^j_{g(j),g(j)} I \end{pmatrix} \end{array}$$

where the $\lambda^j_{a,b}$ are scalars, I is the appropriately-sized identity matrix for the block in which it appears, and the block in position ρ^j_a, ρ^j_b is non-zero only if ρ^j_a and ρ^j_b

are equivalent representations of A.

Proof. To prove this, we need only show that $\rho^j(\sigma)$ has the block form indicated above. Let V^j be the representation module for ρ^j on B. Viewing V^j as a $\mathbb{C}A$-module, we have $V^j = V_1^j \oplus \cdots \oplus V_{g(j)}^j$, and the V_i^j are simple $\mathbb{C}A$-modules where the module action on V_i^j is given by ρ_i^j. Note that a basis $\mathcal{B}$ of V^j has already been chosen by virtue of the fact that ρ^j is in matrix form, and likewise, bases $\mathcal{B}_i$ for the V_i^j are given since ρ_i^j is in matrix form. Since ρ^j is adapted, we have that $\mathcal{B}_i \subseteq \mathcal{B}$ for all i, $\mathcal{B}_i \cap \mathcal{B}_k = \emptyset$ for $i \neq k$, the $\mathcal{B}_i$ are ordered as subsets of $\mathcal{B}$, and $\cup \mathcal{B}_i = \mathcal{B}$.

Since σ commutes with A, we have that, for all $v \in V^j$ and $x \in A$,

$$\rho^j(\sigma) \cdot x \cdot v = \rho^j(\sigma x) \cdot v = \rho^j(x\sigma) \cdot v = x \cdot \rho^j(\sigma) \cdot v,$$

i.e., $\rho^j(\sigma)$ is a $\mathbb{C}A$-linear map from V^j to itself. Hence

$$\rho^j(\sigma) \in \mathrm{Hom}_{\mathbb{C}A}(V^j, V^j) = \mathrm{Hom}_{\mathbb{C}A}(V_1^j \oplus \cdots \oplus V_{g(j)}^j, V_1^j \oplus \cdots \oplus V_{g(j)}^j).$$

According to our basis $\mathcal{B}$ for V^j, we have

$$\rho^j(\sigma) \in \begin{pmatrix} \mathrm{Hom}_{\mathbb{C}A}(V_1^j, V_1^j) & \mathrm{Hom}_{\mathbb{C}A}(V_2^j, V_1^j) & \cdots & \mathrm{Hom}_{\mathbb{C}A}(V_{g(j)}^j, V_1^j) \\ \mathrm{Hom}_{\mathbb{C}A}(V_1^j, V_2^j) & \mathrm{Hom}_{\mathbb{C}A}(V_2^j, V_2^j) & \cdots & \mathrm{Hom}_{\mathbb{C}A}(V_{g(j)}^j, V_2^j) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Hom}_{\mathbb{C}A}(V_1^j, V_{g(j)}^j) & \mathrm{Hom}_{\mathbb{C}A}(V_2^j, V_{g(j)}^j) & \cdots & \mathrm{Hom}_{\mathbb{C}A}(V_{g(j)}^j, V_{g(j)}^j) \end{pmatrix}.$$

Let $X \in \mathrm{Hom}_{\mathbb{C}A}(V_a^j, V_b^j)$ be given in matrix form with respect to the bases $\mathcal{B}_a$, $\mathcal{B}_b$. By Lemma 4.2.1, X is either 0 or an isomorphism (i.e. X is either 0 or ρ_a^j and ρ_b^j are equivalent representations of A). Suppose then that X is an isomorphism.

The goal, then, is to show that X is a diagonal matrix. Since X is $\mathbb{C}A$-linear, for every $x \in \mathbb{C}A, v \in V_a^j$ we have

$$Xx \cdot v = x \cdot Xv.$$

With respect to $\mathcal{B}_a$, $\mathcal{B}_b$ this is

$$X\rho_a^j(x)v = \rho_b^j(x)Xv,$$

and hence for every $x \in \mathbb{C}A$

$$X\rho_a^j(x) = \rho_b^j(x)X.$$

Since ρ_a^j and ρ_b^j are equivalent, and are either null or part of an adapted set of matrix representations, we have $\rho_a^j = \rho_b^j$, and thus

$$X\rho_a^j(x) = \rho_a^j(x)X \tag{4.3}$$

for all $x \in \mathbb{C}A$.

Now, either ρ_a^j is null of dimension 1 (in which case so is ρ_b^j), which means that X is 1-dimensional and we're done, or ρ_a^j is an irreducible representation of A. So, suppose ρ_a^j is irreducible. Then by Burnside's theorem (Theorem 1.14 of [31]), we have

$$\rho_a^j(\mathbb{C}A) = M_{|\mathcal{B}_a|}(\mathbb{C}),$$

and therefore (4.3) says that X is in the center of $M_{|\mathcal{B}_a|}(\mathbb{C})$, i.e. X is diagonal. $\square$

Remark/Notation: Schur's Lemma says that adapted representations cause the

matrix $\rho(\sigma)$ to be sparse and structured. Under the same hypotheses as above, let $\mathcal{M}(B, A)$ be the maximum multiplicity of an irreducible or dimension-1 null representation of A occurring in the restriction, from B to A, of an irreducible representation of B. Since $\rho(\sigma)$ is a block matrix of the form

$$\rho(\sigma) = \begin{pmatrix} \rho^1(\sigma) & 0 & 0 & \cdots & 0 \\ 0 & \rho^2(\sigma) & 0 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & \rho^k(\sigma) \end{pmatrix},$$

and the block $\rho^j(\sigma)$ is of the form

$$\rho^j(\sigma) = \begin{array}{c} \\ \rho_1^j \\ \rho_2^j \\ \vdots \\ \rho_{g(j)}^j \end{array} \begin{array}{c} \begin{array}{cccc} \rho_1^j & \rho_2^j & \cdots & \rho_{g(j)}^j \end{array} \\ \begin{pmatrix} \lambda_{1,1}^j I & \lambda_{1,2}^j I & \cdots & \lambda_{1,g(j)}^j I \\ \lambda_{2,1}^j I & \lambda_{2,2}^j & \cdots & \lambda_{2,g(j)}^j I \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{g(j),1}^j I & \lambda_{g(j),2}^j I & \cdots & \lambda_{g(j),g(j)}^j I \end{pmatrix} \end{array}$$

where $\lambda_{a,b}^j \neq 0$ implies ρ_a^j and ρ_b^j are equivalent representations of A, we see that the matrix $\rho(\sigma)$ contains at most $\mathcal{M}(B, A)$ non-zero entries per row and column. The computational implication of Schur's Lemma, then, is that for an arbitrary $d_\rho \times d_\rho$ matrix H, it requires no more than $\mathcal{M}(B, A)d_\rho^2$ complex multiplications and additions to perform each of the matrix multiplications $\rho(\sigma)H$ and $H\rho(\sigma)$, as opposed to the upper bound of d_ρ^3 multiplications and additions that would be necessary if $\rho(\sigma)$ were arbitrary. See [22] for a variety of applications of this consequence of Schur's Lemma in the construction of group FFTs.

4.3 An FFT for the Symmetric Group

With the tools developed in Section 4.2, we are now ready to extend the reduction idea used in the Cooley-Tukey FFT to create an FFT for the symmetric group S_n. The algorithm that we describe in this section is due to Clausen [4]. We will use symmetric group FFTs as sub-algorithms in the FFT for the rook monoid which we develop in Chapter 7, and we will generalize Clausen's construction to create a different FFT for the rook monoid in Chapter 9. We begin by reviewing the necessary representation theory of the symmetric group. A standard reference is [18].

4.3.1 Seminormal Representations of the Symmetric Group

Clausen's algorithm for the FFT on the symmetric group S_n requires a set of inequivalent, irreducible, chain-adapted representations relative to the chain

$$S_n > S_{n-1} > \cdots > S_1 = \{e\},$$

where e is the identity of S_n and S_k is the subgroup of S_n which fixes pointwise $k+1$ through n. Two such sets of representations are *Young's seminormal* and *Young's orthogonal* forms [41]. We describe Young's seminormal form here. The description given here may be found in [30], and this is the description that is used in generalizations to seminormal representations of Iwahori-Hecke algebras (see, e.g., [16] and [30]). The matrices described here are rescaled versions of the ones found in the description of Young's seminormal form appearing in [4] and in Chapter 3 of [18].

The reader should compare the seminormal representations of S_n described here

to the seminormal representations of R_n described in Section 9.3.

Definition (partition)**.** A *partition* λ of a nonnegative integer k (written $\lambda \vdash k$) is a weakly decreasing sequence of nonnegative integers whose sum is k. We consider two partitions to be equal if and only if they only differ by the number of 0's they contain, and we identify a partition λ with its Young diagram.

For example, $\lambda = (5, 5, 3, 1)$ is a partition of 14, and

$$\lambda = (5, 5, 3, 1) = (5, 5, 3, 1, 0) = \ydiagram{5,5,3,1}.$$

It is well-known that a complete set of inequivalent, irreducible representations for S_n is indexed by the partitions of n (see, for example, [18]).

Definition (tableau, standard tableau)**.** For $\lambda \vdash n$, define L to be a *tableau of shape* λ if it is a filling of the diagram for λ with numbers from $\{1, 2, \ldots n\}$ such that each number in L appears exactly once. L is a *standard tableau* if, furthermore, the entries in each column of L increase from top to bottom and the entries in each row of L increase from left to right.

Fix λ. Let T^λ denote the set of standard tableaux of shape λ. The symmetric group acts on tableaux by permuting their entries. If L is a tableau, then $(i-1, i)L$ is the tableau that is obtained from L by swapping $i-1$ and i. Note that $L \in T^\lambda$ need not imply $(i-1, i)L \in T^\lambda$ (that is, $(i-1, i)L$ need not be standard).

Let $\{v_L : L \in T^\lambda\}$ be a set of independent vectors. We form

$$V^\lambda = \mathbb{C}\text{-span}\{v_L : L \in T^\lambda\}.$$

As such, the symbols v_L, are a basis for the vector space V^λ. Young defined an

action of S_n on V^λ in such a way that (extending by linearity) V^λ is an irreducible $\mathbb{C}S_n$-module and such that, as λ ranges over all partitions of n, the V^λ constitute a complete set of inequivalent, irreducible representations of S_n. We first describe this action, and we then describe an ordering of the bases for the V^λ so that the resulting matrix representations are chain-adapted to $S_n > S_{n-1} > \cdots > S_1$.

Definition (content)**.** If b is a box of λ in position (i, j), then the *content* of b is defined to be

$$ct(b) = j - i.$$

Let $L \in T^\lambda$. If $i-1, i \in L$, then let $L(i-1)$ and $L(i)$ denote the box in L containing $i-1$ and i, respectively.

To define the action of S_n on V^λ, it is sufficient to define the action of a set of generators of S_n on V^λ.

Definition (action of S_n on V^λ)**.** Define the action of the transpositions $t_i = (i-1, i)$, for $2 \leq i \leq n$, as follows:

$$t_i v_L = \frac{1}{ct(L(i)) - ct(L(i-1))} v_L + (1 + \frac{1}{ct(L(i)) - ct(L(i-1))}) v_{L'}$$

where

$$v_{L'} = \begin{cases} v_{t_i L} & \text{if } t_i L \text{ is standard,} \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 4.3.1 (Young)**.** *As λ varies over all partitions of n, the V^λ under the above actions of $\mathbb{C}S_n$ constitute a complete set of inequivalent, irreducible representations of $\mathbb{C}S_n$.*

Definition (corner of a partition)**.** A corner is a box c of λ for which λ contains no box to the right or below c. In other words, the corners are the possible positions

of n in a standard tableau of shape λ.

We now record the Branching Theorem for S_n (see, e.g., Chapter 2 of [18]).

Theorem 4.3.2 (Branching Theorem)**.** *As a $\mathbb{C}S_{n-1}$ module,*

$$V^\lambda \cong \bigoplus_{\mu \in \lambda^-} V^\mu,$$

where λ^- is the set of all partitions $\mu \vdash n-1$ such that μ is obtained by removing a corner from λ.

Now, for purposes of chain-adaptation, we order the basis $\{v_L\}$ for V^λ using Young's famous *last-letter ordering*.

Definition (last letter ordering)**.** Consider two standard tableaux L_1 and L_2 of the same shape λ. If n is in a higher row in L_1 than in L_2, then declare $L_1 < L_2$. Otherwise, n is in the same row of both L_1 and L_2 (and is necessarily the last entry in that row), so delete the box containing n from the tableaux and consider the position of $n-1$, etc.

It is now easy to see, under this ordering of the bases for the V^λ, that the matrix representations afforded by the action of S_n on the V^λ are chain-adapted to the chain $S_n > \cdots > S_1$. Let c_j denote the j^{th} corner of the partition λ and let $V_j^\lambda = \mathbb{C}\text{-span}(v_L : L \text{ contains } n \text{ in box } c_j)$. It is clear that $S_{n-1}V_j^\lambda \subseteq V_j^\lambda$ for all j and, by the definition of the action of S_n on V^λ, that each V_j^λ is an irreducible $\mathbb{C}S_{n-1}$ module. The same argument is used to perform induction down the chain $S_n > \cdots > S_1$, and is trivial because we ordered our basis for V^λ inductively according to the same rule.

We now restate the Branching Theorem for S_n with respect to the matrix representations obtained from this action and last-letter ordering:

Theorem 4.3.3 (Branching Theorem)**.** *Let ρ^λ be the matrix representation associated to V^λ with respect to the basis $\{v_L\}$, with the basis ordered according to Young's last-letter ordering. Then*

$$\rho^\lambda|_{S_{n-1}} = \bigoplus_{\mu \in \lambda^-} \rho^\mu,$$

where λ^- is the set of all partitions $\mu \vdash n-1$ such that μ is obtained by removing a corner from λ. The first $\mu \in \lambda^-$ is the one that removes the highest corner, the next μ is the one that removes the second highest corner, etc.

The matrix representations of S_n afforded by this basis constitute Young's seminormal representations. With these representations in hand, we are now ready to explain Clausen's algorithm [4] for the Fourier transform on S_n.

4.3.2 An FFT for the Symmetric Group

Consider S_n and the subgroup chain $S_n > S_{n-1} > \cdots > S_1 = \{e\}$, where e is the identity of S_n and S_k is the subgroup of S_n which fixes pointwise $k{+}1, k{+}2, \ldots, n$. For any complete set of inequivalent, irreducible representations $\mathcal{Y}$ of $\mathbb{C}S_n$, let $\mathcal{T}_\mathcal{Y}(S_n)$ denote the minimal number of operations needed to compute the Fourier transform of an arbitrary complex-valued function f on S_n using $\mathcal{Y}$. That is, $\mathcal{T}_\mathcal{Y}(S_n)$ is the minimal number of operations necessary to compute, for all $\rho \in \mathcal{Y}$, the sums

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

For the purpose of our analysis, we assume that all representations $\rho \in \mathcal{Y}$ are precomputed (i.e., evaluated at all elements of S_n) and stored in memory.

Notice that a naive algorithm gives the bound

$$\mathcal{T}_{\mathcal{Y}}(S_n) \leq \sum_{\rho \in \mathcal{Y}} |S_n| d_\rho^2,$$

which, by (2.2), is $|S_n|^2$. Now, let us define $\mathcal{C}(S_n)$ to be the minimal value of the $\mathcal{T}_{\mathcal{Y}}(S_n)$, as $\mathcal{Y}$ varies over all complete sets of inequivalent, irreducible representations. Our goal is to give a bound for $\mathcal{C}(S_n)$ which compares favorably to $|S_n|^2$, and we will do so by giving such a bound on $\mathcal{T}_{\mathcal{Y}}(S_n)$ for a computationally advantageous set $\mathcal{Y}$. In particular, let $\mathcal{Y} = \mathcal{Y}_n$ be any complete set of inequivalent, irreducible representations for S_n adapted to this chain. One well-known choice is Young's seminormal form, as described in Section 4.3.1.

S_{n-1} partitions S_n into n cosets. That is, $S_n = \cup_{i=1}^{n} T_i S_{n-1}$ for some coset representatives T_i. Let t_j be the transposition $(j-1, j)$, and let us use the following set of coset representatives:

$$\{T_i : 1 \leq i \leq n, T_i = \underbrace{e \cdots e}_{i} t_{i+1} t_{i+2} \cdots t_n\}.$$

Thus, for example, $T_1 = t_2 t_3 \cdots t_n$, $T_{n-1} = t_n$, and $T_n = e$. Let us define functions f_{T_i}:

$$f_{T_i} : S_{n-1} \to \mathbb{C}$$
$$\text{by } f_{T_i}(s) = f(T_i s).$$

For any representation ρ, since ρ is a homomorphism, it is straightforward that

$$\hat{f}(\rho) = \sum_{s \in S_n} f(s)\rho(s) = \sum_{i=1}^{n} \rho(T_i) \sum_{s \in S_{n-1}} f_{T_i}(s)\rho(s). \tag{4.4}$$

Notice that, just as in the Cooley-Tukey FFT, the inner sums in (4.4) are themselves Fourier transforms on S_{n-1}. If we knew all of the $\hat{f}_{T_i}(\gamma)$ for all $\gamma \in \mathcal{Y}_{n-1}$ and for all i, then we could reconstruct the inner sums in (4.4) based on how ρ splits when restricted to S_{n-1}. This splitting is described by the Branching Theorem (Theorem 4.3.3), and this reconstruction can be done for free because ρ is adapted to $S_n > S_{n-1}$. In particular, say $\rho|_{S_{n-1}} = \rho_1 \oplus \cdots \oplus \rho_k$, $\rho_j \in \mathcal{Y}_{n-1}$. Then

$$\sum_{s \in S_{n-1}} f_{T_i}(s)\rho(s) = \sum_{s \in S_{n-1}} f_{T_i}(s)\left[\rho_1(s) \oplus \cdots \oplus \rho_k(s)\right] = \hat{f_{T_i}}(\rho_1) \oplus \cdots \oplus \hat{f_{T_i}}(\rho_k).$$

Once we have computed the inner sums in (4.4), to finish computing $\hat{f}(\rho)$, we multiply the inner sums by the $\rho(T_i)$ and add the results. Since the number of operations this last step requires depends on the coset representatives chosen, let us denote the number of operations it requires by $M_{S_n}(\{T_i\}_{i=1}^n)$. Equation (4.4) then implies

$$\mathcal{T}_{\mathcal{Y}_n}(S_n) \leq n\mathcal{T}_{\mathcal{Y}_{n-1}}(S_{n-1}) + M_{S_n}(\{T_i\}_{i=1}^n). \tag{4.5}$$

We now turn to analyzing the $M_{S_n}(\{T_i\}_{i=i}^n)$ term in (4.5). Notice that, for $j > 2$, $t_j \in S_j$ and t_j commutes with S_{j-2}. Furthermore, $t_2 \in S_2$ and t_2 commutes with S_1. By the combinatorics of Young tableaux and the Branching Theorem (Theorem 4.3.3), the maximum multiplicity occurring in the restriction of any irreducible representation of S_j to S_{j-2} is 2. By Schur's Lemma, then, for any j ($2 \leq j \leq n$) and any $\rho \in \mathcal{Y}_n$, $\rho(t_j)$ contains at most 2 nonzero entries per row and column.

Fix $\rho \in \mathcal{Y}_n$. Since $\rho(T_i) = \rho(t_{i+1})\rho(t_{i+2})\cdots\rho(t_n)$, computing $\rho(T_i)A_{T_i}(\rho)$ for an arbitrary matrix $A_{T_i}(\rho)$ may be accomplished by multiplying $A_{T_i}(\rho)$ on the left by $\rho(t_n)$, multiplying the result on the left by $\rho(t_{n-1})$, multiplying the result of that

on the left by $\rho(t_{n-2})$, etc. It therefore takes a maximum of $2(n-i)d_\rho^2$ operations to perform the multiplication $\rho(T_i)A_{T_i}(\rho)$ (keeping in mind that multiplying by the identity matrix may be done for free), and once $\rho(T_i)A_{T_i}(\rho)$ has been computed for all T_i, it takes a maximum of $(n-1)d_\rho^2$ operations to add the results to give $\sum_{i=1}^n \rho(T_i)A_{T_i}(\rho)$. Letting ρ vary over $\mathcal{Y}_n$, then, implies

$$\begin{aligned} M_{S_n}(\{T_i\}_{i=i}^n) &\leq \sum_{\rho\in\mathcal{Y}_n}\sum_{i=1}^n 2(n-i)d_\rho^2 + \sum_{\rho\in\mathcal{Y}_n}(n-1)d_\rho^2 \\ &= (n-1)(n)|S_n| + (n-1)|S_n| \\ &= (n^2-1)|S_n|. \end{aligned}$$

Putting this together with (4.5), we obtain:

$$\mathcal{T}_{\mathcal{Y}_n}(S_n) \leq n\mathcal{T}_{\mathcal{Y}_{n-1}}(S_{n-1}) + (n^2-1)|S_n|.$$

The algorithm described here is the heart of Clausen's FFT on S_n [4]. Induction on n then yields:

Theorem 4.3.4 (Clausen [4])**.** $\mathcal{T}_{\mathcal{Y}_n}(S_n) \leq \frac{2}{3}n(n+1)^2 n!$.

This is of order $n!(\log n!)^3 = |S_n|\log^3|S_n|$. By a more careful analysis of the matrix multiplications involved, D. Maslen has obtained an algorithm for the FFT on S_n of complexity $O(|S_n|\log^2|S_n|)$. Here is Theorem 1.1 of [21]:

Theorem 4.3.5 (Maslen)**.** *Let $\mathcal{Y}_n$ denote a complete set of inequivalent, irreducible representations for S_n in Young's orthogonal or seminormal form. Then*

$$\mathcal{T}_{\mathcal{Y}_n}(S_n) \leq \frac{3}{4}n(n-1)|S_n|.$$

4.4 Complexity for Inverse Semigroups

If G is a finite group, then $\{g\}_{g\in G}$ indexes the natural basis of $\mathbb{C}G$. If S is a finite inverse semigroup, then $\mathbb{C}S$ has two natural bases, the semigroup basis $\{s\}_{s\in S}$, and the groupoid basis $\{\lfloor s \rfloor\}_{s\in S}$. We therefore define two notions of computational complexity for the Fourier transform on S, one being the minimal number of operations necessary to change from the semigroup basis of its algebra to a Fourier basis, and the other being the minimal number of operations necessary to change from the groupoid basis of its algebra to a Fourier basis.

Definition (Computational complexity)**.** Let $\mathcal{Y}$ be a complete set of inequivalent, irreducible representations of $\mathbb{C}S$. For an arbitrary element $f \in \mathbb{C}S$ expressed with respect to the semigroup basis

$$f = \sum_{s\in S} f(s)s,$$

the minimal number of operations to compute the Fourier transform of f, i.e., to compute

$$\hat{f}(\rho) = \sum_{s\in S} f(s)\rho(s) \tag{4.6}$$

for all $\rho \in \mathcal{Y}$, is denoted by $\mathcal{T}_{\mathcal{Y}}^{\text{semigroup}}(S)$.

For an arbitrary element $f \in \mathbb{C}S$ expressed with respect to the groupoid basis

$$f = \sum_{s\in S} f(s) \lfloor s \rfloor ,$$

the minimal number of operations to compute the Fourier transform of f, i.e., to compute

$$\hat{f}(\rho) = \sum_{s\in S} f(s)\rho(\lfloor s \rfloor) \tag{4.7}$$

for all $\rho \in \mathcal{Y}$, is denoted by $\mathcal{T}_{\mathcal{Y}}^{\text{groupoid}}(S)$.

As usual, an *operation* is defined to be a single complex multiplication followed by a complex addition. For the purposes of our analysis, we assume that all representations in $\mathcal{Y}$ are precomputed (i.e., evaluated at every semigroup basis element of $\mathbb{C}S$ or at every groupoid basis element of $\mathbb{C}S$) and stored in memory.

Now, let $\mathcal{Y}$ vary over all complete sets of inequivalent, irreducible representations for $\mathbb{C}S$. We define

$$\mathcal{C}^{\text{semigroup}}(S) = \min_{\mathcal{Y}}(\mathcal{T}_{\mathcal{Y}}^{\text{semigroup}}(S)),$$

$$\mathcal{C}^{\text{groupoid}}(S) = \min_{\mathcal{Y}}(\mathcal{T}_{\mathcal{Y}}^{\text{groupoid}}(S)), \text{ and}$$

$$\mathcal{C}(S) = \max(\mathcal{C}^{\text{groupoid}}(S), \mathcal{C}^{\text{semigroup}}(S)).$$

If G is a group, the semigroup and the groupoid bases of $\mathbb{C}G$ are identical, so we drop the superscripts on the complexity notation. For example,

$$\mathcal{T}_{\mathcal{Y}}^{\text{semigroup}}(G) = \mathcal{T}_{\mathcal{Y}}^{\text{groupoid}}(G) = \mathcal{T}_{\mathcal{Y}}(G).$$

Now, let S be a finite inverse semigroup and let $\mathcal{Y}$ be any complete set of inequivalent, irreducible representations of $\mathbb{C}S$. A naive implementation of the Fourier transform on S, i.e., computing (4.6) and (4.7) directly, gives

$$\mathcal{C}(S) \leq \sum_{\rho \in \mathcal{Y}} |S| d_\rho^2,$$

which, by (2.2), gives

Theorem 4.4.1. $\mathcal{C}(S) \leq |S|^2$.

We define a *fast Fourier transform* on an inverse semigroup S to be an algorithm (or a collection of algorithms) which yields a complexity result for $\mathcal{C}(S)$ that compares favorably to $|S|^2$. Since every group is an inverse semigroup, we also have $\mathcal{C}(G) \leq |G|^2$ for any group G, although, as we have already seen, many families of groups enjoy results along the lines of $\mathcal{C}(G) = O(|G| \log^c |G|)$. Indeed, such upper bounds remain the goal in group FFT theory. There are also groups G for which there currently exist greatly improved (but not $O(|G| \log^c |G|)$) algorithms, such as matrix groups over finite fields or, more generally, finite groups of Lie type [22]. It is conjectured [25] that there are universal constants c_1, c_2 such that for any group G, $\mathcal{C}(G) \leq c_1 |G| \log^{c_2} |G|$.

In Chapters 7 and 8, we demonstrate that such bounds are also attainable for certain (non-group) inverse semigroups of interest. In particular, we exhibit algorithms for the rook monoid R_n and for its wreath products $G \wr R_n$ by arbitrary finite groups G, which in turn yield the following complexity results:

Theorem 4.4.2. *We have*

$$\mathcal{C}^{\text{groupoid}}(R_n) \leq \frac{3}{4} n(n-1)|R_n| = O(|R_n| \log^2 |R_n|)$$

and

$$\mathcal{C}^{\text{semigroup}}(R_n) \leq \frac{2}{3} n^3 |R_n| + \frac{3}{4} n(n-1)|R_n| = O(|R_n| \log^3 |R_n|).$$

Hence

$$\mathcal{C}(R_n) = O(|R_n| \log^3 |R_n|).$$

Also, if h is the number of inequivalent, irreducible representations of G, then

$$\begin{aligned}\mathcal{C}^{\text{groupoid}}(G \wr R_n) &\leq \left[\frac{\mathcal{C}(G)}{|G|} \cdot \frac{n(n+1)}{2} + 2^h \frac{n^2(n+1)^2}{4} + 1\right] \cdot |G \wr R_n| \\ &= O(|G \wr R_n| \log^4 |G \wr R_n|),\end{aligned}$$

and

$$\begin{aligned}\mathcal{C}^{\text{semigroup}}(G \wr R_n) &\leq \mathcal{C}^{\text{groupoid}}(G \wr R_n) + \left[|G|n^2 + |G|^2 \frac{n^3}{3}\right] \cdot |G \wr R_n| \\ &= O(|G \wr R_n| \log^4 |G \wr R_n|).\end{aligned}$$

Hence

$$\mathcal{C}(G \wr R_n) = O(|G \wr R_n| \log^4 |G \wr R_n|).$$

Chapter 5

Inverse Semigroups and Groupoid Algebras

In this chapter, we use a theorem of B. Steinberg (Theorem 4.6 in [39]) to reduce the problem of creating FFTs on inverse semigroups to the problems of creating FFTs on their maximal subgroups and fast zeta transforms on their poset structures. We explain Steinberg's result in Section 5.1. In Section 5.2, we use this result to give a complete description of the irreducible representations of a finite inverse semigroup in terms of the representations of its maximal subgroups, and we use this description in turn to obtain our result on Fourier transforms. We provide another application of Steinberg's theorem in Section 5.3, where we use it to derive C. Grood's "natural" representations of the rook monoid [15] in a straightforward and elementary manner.

5.1 Matrix Algebras Over Group Algebras

Let S be a finite inverse semigroup. Recall from Section 3.2 that the groupoid basis of $\mathbb{C}S$ is defined as follows:

Definition (groupoid basis)**.** Define, for each $s \in S$, the element $\lfloor s \rfloor \in \mathbb{C}S$ by

$$\lfloor s \rfloor = \sum_{t \in S : t \leq s} \mu(t, s) t,$$

so that

$$s = \sum_{t \in S : t \leq s} \lfloor t \rfloor .$$

The collection $\{\lfloor s \rfloor\}_{s \in S}$ is a basis for $\mathbb{C}S$. Multiplication in $\mathbb{C}S$ relative to this basis is given by the linear extension of

$$\lfloor s \rfloor \lfloor t \rfloor = \begin{cases} \lfloor st \rfloor & \text{if } \mathrm{dom}(s) = \mathrm{ran}(t) \iff s^{-1}s = tt^{-1}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.1}$$

Given an element $s \in S$, it is natural to think of s as an "isomorphism" from $\mathrm{dom}(s)$ to $\mathrm{ran}(s)$. We use this to define the notion of *isomorphic idempotents*.

Definition. Let $a, b \in S$ be idempotent. a and b are said to be *isomorphic* idempotents if there is an "isomorphism" from a to b, that is, if there is an element $s \in S$ such that $a = s^{-1}s$ and $b = ss^{-1}$.

Let us now define two idempotents in S to be $\mathcal{D}$*-related* if they are isomorphic. For the rook monoid R_n, the idempotents are the restrictions of the identity map, and two idempotents are isomorphic if and only if they have the same rank. We can extend $\mathcal{D}$ to an equivalence relation on S by defining $s \mathcal{D} t$ if $s^{-1}s$ is isomorphic to $t^{-1}t$ (or, equivalently, if ss^{-1} is isomorphic to tt^{-1}). This is Green's famous

$\mathcal{D}$-relation ([6], [14]), and the equivalence classes of S with respect to this relation are called the $\mathcal{D}$-classes of S. We mention that an equivalent characterization of $\mathcal{D}$ is that $s\mathcal{D}t$ if and only if s and t generate the same two-sided ideal in S. For R_n, there are $n+1$ $\mathcal{D}$-classes. They are $D_0, D_1, \ldots, D_n$, where D_k is the set of elements of R_n of rank k.

Let $e \in S$ be idempotent. Let G_e be the *maximal subgroup at* e, that is, the largest subset of S which contains e and which is also a group. It is easy to see that

$$G_e = \{s \in S : s^{-1}s = ss^{-1} = e\},$$

and that e is the identity of G_e. If a, b are isomorphic idempotents, it is straightforward to show that $G_a \cong G_b$. For R_n, the maximal subgroup at an idempotent e of rank k is isomorphic to the permutation group S_k.

Now, let us describe the decomposition of the semigroup algebra $\mathbb{C}S$ into a direct sum of matrix algebras over group algebras. Let $D_0, \ldots, D_n$ be the $\mathcal{D}$-classes of S. Let $\mathbb{C}D_k$ be the $\mathbb{C}$-span of $\{\lfloor s \rfloor : s \in D_k\}$. It is immediate from (5.1) that $\mathbb{C}S = \bigoplus_{k=0}^{n} \mathbb{C}D_k$. We now show

Theorem 5.1.1 (B. Steinberg)**.** *Let r_k indicate the number of idempotents in D_k, and let e_k be any idempotent in D_k. Denote the maximal subgroup of S at e_k by G_k. Then, as algebras, $\mathbb{C}D_k \cong M_{r_k}(\mathbb{C}G_k)$.*

Proof. We already know that G_a and G_b are isomorphic for any idempotents $a, b \in D_k$. Now, fix an idempotent $e_k \in D_k$, and for every idempotent $a \in D_k$, fix an element $p_a \in S$ such that ${p_a}^{-1}p_a = e_k$ and $p_a{p_a}^{-1} = a$ (that is, p_a is an isomorphism from e_k to a). Let us take $p_{e_k} = e_k$. It is easy to show that, in fact, $p_a \in D_k$. We view our $r_k \times r_k$ matrices as being indexed by pairs of idempotents in D_k. We now define our isomorphism by defining it on the basis $\{\lfloor s \rfloor : s \in D_k\}$

of $\mathbb{C}D_k$ and extending linearly. So, for an element $\lfloor s \rfloor \in \mathbb{C}D_k$ with $s^{-1}s = a$ and $ss^{-1} = b$, define

$$\phi(\lfloor s \rfloor) = {p_b}^{-1} s p_a E_{b,a},$$

where $E_{b,a}$ is the standard $r_k \times r_k$ matrix with a 1 in the b, a position and 0 elsewhere.

A quick calculation shows that ${p_b}^{-1} s p_a \in G_k$ by construction. We now show that ϕ is a homomorphism.

$$\begin{aligned}
\phi(\lfloor s \rfloor)\phi(\lfloor t \rfloor) &= \left({p_{ss^{-1}}}^{-1} s p_{s^{-1}s} E_{ss^{-1},s^{-1}s}\right)\left({p_{tt^{-1}}}^{-1} t p_{t^{-1}t} E_{tt^{-1},t^{-1}t}\right) \\
&= \begin{cases} {p_{ss^{-1}}}^{-1} s p_{s^{-1}s} {p_{tt^{-1}}}^{-1} t p_{t^{-1}t} E_{ss^{-1},t^{-1}t} & \text{if } s^{-1}s = tt^{-1}, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

Supposing then that $s^{-1}s = tt^{-1}$, we have

$$\begin{aligned}
{p_{ss^{-1}}}^{-1} s p_{s^{-1}s} {p_{tt^{-1}}}^{-1} t p_{t^{-1}t} E_{ss^{-1},t^{-1}t} &= {p_{ss^{-1}}}^{-1} s p_{s^{-1}s} {p_{s^{-1}s}}^{-1} t p_{t^{-1}t} E_{ss^{-1},t^{-1}t} \\
&= {p_{ss^{-1}}}^{-1} s s^{-1} s t p_{t^{-1}t} E_{ss^{-1},t^{-1}t} \\
&= {p_{ss^{-1}}}^{-1} s t p_{t^{-1}t} E_{ss^{-1},t^{-1}t} \\
&= {p_{(st)(st)^{-1}}}^{-1} s t p_{(st)^{-1}(st)} E_{(st)(st)^{-1},(st)^{-1}(st)} \\
&= \phi(\lfloor st \rfloor).
\end{aligned}$$

It is now easy to see that ϕ is an isomorphism, with the inverse induced by, for $s \in G_k$,

$$s E_{b,a} \mapsto \left\lfloor p_b s {p_a}^{-1} \right\rfloor.$$

□

The corollary is:

Corollary 5.1.2. $\mathbb{C}S \cong \bigoplus_{k=0}^{n} M_{r_k}(\mathbb{C}G_k)$.

A dimensionality count thus establishes

$$|S| = \sum_{k=0}^{n} r_k^2 |G_k|. \tag{5.2}$$

We now explain what the isomorphism constructed in the proof of Theorem 5.1.1 translates into when $S = R_n$, the rook monoid. Fix a $\mathcal{D}$-class D_k (that is, the subset of elements of R_n of rank k), and let us take $e_k \in D_k$ to be the partial identity on $\{1, \ldots, k\}$, that is:

$$e_k = \begin{pmatrix} 1 & 2 & \cdots & k & k+1 & \cdots & n \\ 1 & 2 & \cdots & k & - & \cdots & - \end{pmatrix}.$$

We then have

$$G_k = \{s \in R_n : \mathrm{dom}(s) = \mathrm{ran}(s) = \{1, \ldots, k\}\}.$$

Let us identify G_k with the permutation group S_k in the obvious manner.

For an idempotent $a \in D_k$ (that is, a rank-k restriction of the identity map), let us take p_a to be the unique order-preserving bijection from $\{1, \ldots, k\}$ to $\mathrm{dom}(a) = \mathrm{ran}(a)$. For an element $s \in R_n$ of rank k, let us define the *permutation type* of s, $\mathrm{perm}(s)$, to be, informally, the "arrows" from $\mathrm{dom}(s)$ to $\mathrm{ran}(s)$, expressed as a permutation in $G_k = S_k$. For example, if

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & - & 1 & 2 \end{pmatrix}, \text{ then } \mathrm{perm}(s) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

because s sends the first element of its domain to the third element of its range, the second element of its domain to the first element of its range, and the third element of its domain to the second element of its range.

Formally, we define

$$\mathrm{perm}(s) = {p_{\mathrm{ran}(s)}}^{-1} s p_{\mathrm{dom}(s)}$$

where $\mathrm{dom}(s)$ and $\mathrm{ran}(s)$ are once again understood to be the corresponding partial identities in R_n, so that $p_{\mathrm{dom}(s)}$ is the unique order preserving bijection from $\{1, \ldots, k\}$ to $\mathrm{dom}(s)$ and ${p_{\mathrm{ran}(s)}}^{-1}$ is the unique order preserving bijection from $\mathrm{ran}(s)$ to $\{1, \ldots, k\}$.

The isomorphism ϕ defined in the proof of Theorem 5.1.1 now works as follows. We have $\binom{n}{k} \times \binom{n}{k}$ matrices, so let us index their rows and columns by the k-subsets of $\{1, \ldots, n\}$. We have

$$\mathbb{C}D_k \cong M_{\binom{n}{k}}(\mathbb{C}S_k)$$
$$\text{by } \phi(\lfloor s \rfloor) = \mathrm{perm}(s) E_{\mathrm{ran}(s),\mathrm{dom}(s)}.$$

Let us state the corollary for reference.

Corollary 5.1.3. $\mathbb{C}R_n \cong \bigoplus_{k=0}^{n} M_{\binom{n}{k}}(\mathbb{C}S_k)$.

5.2 Representations of Inverse Semigroups

Let S be a finite inverse semigroup. In this section, we explain how Theorem 5.1.1 allows us to describe a complete set of inequivalent, irreducible representations of $\mathbb{C}S$ in terms of the representations of its maximal subgroups, and we state and prove our most general inverse semigroup FFT complexity result. We remark that

Theorems 5.2.1 through 5.2.5 in this section are essentially folklore results from semigroup theory and Morita theory (see, e.g., [13] and [39]). We make them completely explicit for the computational results that follow.

We begin with a finite-dimensional algebra A (with identity 1_A) over $\mathbb{C}$ and the following theorem.

Theorem 5.2.1. *A is semisimple if and only if $M_n(A)$ is semisimple.*

Proof. Let $J(A)$ denote the *Jacobson radical* of the algebra A. By p. 70 of [13], $J(M_n(A))$ is equal to $M_n(J(A))$. One characterization of semisimplicity for finite-dimensional algebras is that an algebra is semisimple if and only if it has Jacobson radical $\{0\}$. The theorem follows. □

Suppose now that A is semisimple. Let $\rho_1, \rho_2 : A \to M_k(\mathbb{C})$ be irreducible representations.

Theorem 5.2.2. *For $s \in A$, Define $\bar{\rho_1}, \bar{\rho_2} : M_n(A) \to M_{nk}(\mathbb{C})$ by*

$$\bar{\rho_i}(sE_{i,j}) = E_{i,j} \otimes \rho_i(s),$$

where $E_{i,j}$ is the standard $n \times n$ matrix with a 1 in the i, j position and 0 elsewhere, and extend linearly to representations of $M_n(A)$. Then $\bar{\rho_1}$ and $\bar{\rho_2}$ are irreducible, and are equivalent if and only if ρ_1 and ρ_2 are equivalent.

Proof. The proof here is essentially straightforward linear algebra. Let $\rho = \rho_i$ for $i = 1$ or 2. We begin by showing $\bar{\rho}$ is irreducible. Let V be the underlying vector space on which A acts via ρ, with the basis $\{v_1, \ldots, v_k\}$ giving rise to the matrix representation ρ. To say that ρ is irreducible is to say that, for all $v \in V$ with $v \neq 0$, $\rho(A)v = V$ (that is, every non-zero element of V generates V as an A-module).

The underlying vector space on which $M_n(A)$ acts via $\bar{\rho}$ may be described as a vector space $\bar{V}$ with basis $\{v_1^1, \ldots, v_k^1, v_1^2, \ldots, v_k^2, \ldots, v_1^n, \ldots, v_k^n\}$. In other words, $\bar{V}$ is, as a vector space, the direct sum of n copies of V. Let us superscript them according to this ordering; $\bar{V} = V^1 \oplus \cdots \oplus V^n$, where $V^j = \mathbb{C}\text{-span}(v_1^j, \ldots, v_k^j)$. Then $\bar{\rho}$ defines an action of $M_n(A)$ on $\bar{V}$ under which $sE_{i,j}$ sends vectors in V^j to V^i as described by $\rho(s)$ (and sends vectors outside of V^j to 0). If $\bar{v} \in \bar{V}$ with $\bar{v} \neq 0$, then, say

$$\bar{v} = \sum_{p=1}^{n} \sum_{q=1}^{k} y_q^p v_q^p,$$

and, for some j,

$$\sum_{q=1}^{k} y_q^j v_q^j \neq 0.$$

If $x \in \bar{V}$ is any vector, say

$$x = \sum_{p=1}^{n} \sum_{q=1}^{k} x_q^p v_q^p,$$

then by the irreducibility of ρ, we may find elements $s_1, \ldots, s_n \in A$ for which

$$\bar{\rho}(s_i E_{i,j})\bar{v} = \bar{\rho}(s_i E_{i,j}) \sum_{q=1}^{k} y_q^j v_q^j = \sum_{q=1}^{k} x_q^i v_q^i,$$

and hence

$$\bar{\rho}\left(\sum_{i=1}^{n} s_i E_{i,j}\right)\bar{v} = \sum_{i=1}^{n} \bar{\rho}(s_i E_{i,j})\bar{v} = \sum_{i=1}^{n} \sum_{q=1}^{k} x_q^i v_q^i = x,$$

which proves that $\bar{\rho}$ is irreducible.

Now, suppose that ρ_1 is equivalent to ρ_2. Then there is a matrix X for which $X\rho_1 X^{-1} = \rho_2$. It is straightforward that $(I_n \otimes X)\,\bar{\rho_1}\,(I_n \otimes X^{-1}) = \bar{\rho_2}$, i.e., $\bar{\rho_1}$ is equivalent to $\bar{\rho_2}$.

Now, suppose that $\bar{\rho_1}$ is equivalent to $\bar{\rho_2}$. Then there is a matrix X for which

$X\bar{\rho}_1 X^{-1} = \bar{\rho}_2$. Write X as a block matrix:

$$X = \begin{bmatrix} X_{1,1} & \cdots & X_{1,n} \\ \vdots & & \vdots \\ X_{n,1} & \cdots & X_{n,n} \end{bmatrix},$$

with the $X_{i,j}$ $k \times k$ matrices. Since $\bar{\rho}_1(1_A E_{i,i}) = \bar{\rho}_2(1_A E_{i,i}) = E_{i,i} \otimes I_k$, we have

$$X\,(E_{i,i} \otimes I_k) = (E_{i,i} \otimes I_k)\,X,$$

and thus $X_{i,j} = 0$ unless $i = j$. Furthermore, $\bar{\rho}_1(1_A E_{i,j}) = \bar{\rho}_2(1_A E_{i,j}) = E_{i,j} \otimes I_k$, so we have

$$X(E_{i,j} \otimes I_k) = (E_{i,j} \otimes I_k)X,$$

and since $X = \sum_{l=1}^{n} E_{l,l} \otimes X_{l,l}$, we have

$$E_{i,j} \otimes X_{i,i} = E_{i,j} \otimes X_{j,j}$$

and so $X_{i,i} = X_{j,j}$ for all i and j. Thus $X = I_n \otimes X_{1,1}$, and since X is invertible, so is $X_{1,1}$, and $X^{-1} = I_n \otimes {X_{1,1}}^{-1}$. It now follows that $X_{1,1}\rho_1 {X_{1,1}}^{-1} = \rho_2$, i.e., ρ_1 is equivalent to ρ_2.

□

Let G be a finite group. Let us denote by $\mathrm{IRR}(G)$ a complete set of inequivalent, irreducible matrix representations of $\mathbb{C}G$.

Corollary 5.2.3. *Let $n \in \mathbb{N}$. Then $M_n(\mathbb{C}G)$ is semisimple, and the elements of $\mathrm{IRR}(G)$ are in one-to-one correspondence with the inequivalent, irreducible representations of $M_n(\mathbb{C}G)$. Explicitly, for each $\rho \in \mathrm{IRR}(G)$, let us form the represen-*

tation $\bar{\rho}$ of $M_n(\mathbb{C}G)$ by "tensoring up." That is, $\bar{\rho}$ is given by the linear extension of

$$\bar{\rho}(gE_{i,j}) = E_{i,j} \otimes \rho(g)$$

for $g \in G$. Then the $\bar{\rho}$ form a complete set of inequivalent, irreducible representations of $M_n(\mathbb{C}G)$.

Proof. The semisimplicity of $M_n(\mathbb{C}G)$ follows from the semisimplicity of $\mathbb{C}G$ and Theorem 5.2.1. The only thing left to show here is that the set of representations obtained in this manner is complete, and we accomplish this by a dimensionality count:

$$\dim(M_n(\mathbb{C}G)) = n^2|G|$$

and

$$\sum_{\rho \in \mathrm{IRR}(G)} d_{\bar{\rho}}^2 = \sum_{\rho \in \mathrm{IRR}(G)} (nd_\rho)^2 = n^2 \sum_{\rho \in \mathrm{IRR}(G)} d_\rho = n^2|G|.$$

□

We now turn to sums of algebras. The following theorem is immediate.

Theorem 5.2.4. *Let A and B be semisimple algebras. Let ρ_1 be an irreducible representation of A, and let ρ_2 be an irreducible representation of B. Then we may extend ρ_i to an irreducible representation of the algebra $A \oplus B$ by defining it to be zero on the other summand. The representations ρ_1 and ρ_2 extend to inequivalent representations of $A \oplus B$ in this manner.*

Let S be a finite inverse semigroup with $\mathcal{D}$-classes $D_0, \ldots, D_n$. We know

$$\mathbb{C}S \cong \bigoplus_{k=0}^{n} M_{r_k}(\mathbb{C}G_k), \tag{5.3}$$

where the G_k are the maximal subgroups of S. $\mathbb{C}S$ is semisimple (by Theorem 4.4 in [29]—alternatively, this follows directly from Theorem 5.1.1 and Corollary 5.2.3), and by combining Theorem 5.1.1 with the results so far in this section, we have the following procedure for generating all of the irreducible representations of $\mathbb{C}S$:

Theorem 5.2.5. *The irreducible representations of $\mathbb{C}S$ are in one-to-one correspondence with $\uplus_{k=0}^{n}\mathrm{IRR}(G_k)$. Specifically, given an irreducible representation of $\mathbb{C}G_k$, we tensor it up to an irreducible representation of $M_{r_k}(\mathbb{C}G_k)$ and extend to $\mathbb{C}S$ by letting it be 0 on the other summands of $\bigoplus_{k=0}^{n} M_{r_k}(\mathbb{C}G_k)$ and composing with the isomorphism* (5.3)*. The resulting set of representations is a complete set of inequivalent, irreducible representations of $\mathbb{C}S$.*

Proof. The only thing to show here is that the resulting set of representations is complete, and this follows immediately from a dimensionality count and (5.2). □

Let us now turn to Fourier transforms. Let S be a finite inverse semigroup. We argue now that the set of representations for $\mathbb{C}S$ constructed from the representations of the maximal subgroups of S in Theorem 5.2.5 is a computationally advantageous set of representations, and we exhibit a construction which translates into a fast change of basis from the groupoid basis of $\mathbb{C}S$ to a Fourier basis when fast Fourier transforms on all of the maximal subgroups of S are known.

So, suppose S has $\mathcal{D}$-classes $D_0, \ldots, D_n$. For each $\mathcal{D}$-class D_k, choose an idempotent e_k, and let G_k be the maximal subgroup of S at e_k. Let r_k denote the number of idempotents in D_k. For each idempotent $a \in D_k$, fix an element $p_a \in S$ such that ${p_a}^{-1}p_a = e_k$ and $p_a{p_a}^{-1} = a$ (and take $p_{e_k} = e_k$). By the proof of Theorem 5.1.1, this defines the isomorphism (5.3). Let $\mathcal{Y}$ be the complete set of inequivalent, irreducible representations of $\mathbb{C}S$ given by tensoring up the represen-

tations in $\uplus_{k=0}^{n}\mathrm{IRR}(G_k)$ to representations of $\bigoplus_{k=0}^{n} M_{r_k}(\mathbb{C}G_k)$ and following the isomorphism (5.3). That is, if $\bar{\rho} \in \mathcal{Y}$ was obtained by tensoring up a representation $\rho \in \mathrm{IRR}(G_k)$, then $\bar{\rho}$ is given by the linear extension of

$$\bar{\rho}(\lfloor s \rfloor) = \begin{cases} \bar{\rho}(p_{ss^{-1}}{}^{-1} s p_{s^{-1}s} E_{ss^{-1},s^{-1}s}) = E_{ss^{-1},s^{-1}s} \otimes \rho(p_{ss^{-1}}{}^{-1} s p_{s^{-1}s}) & \text{if } s \in D_k, \\ 0 & \text{otherwise.} \end{cases}$$

Consider an arbitrary element $v \in \mathbb{C}S$, expressed with respect to the groupoid basis:

$$v = \sum_{s \in S} v(s) \lfloor s \rfloor .$$

Now, let $\bar{\rho} \in \mathcal{Y}$, and suppose $\bar{\rho}$ was obtained by tensoring up a representation in $\mathrm{IRR}(G_k)$. Then

$$\hat{v}(\bar{\rho}) = \bar{\rho}(v) = \sum_{s \in S} v(s) \bar{\rho}(\lfloor s \rfloor) = \sum_{s \in D_k} v(s) \bar{\rho}(\lfloor s \rfloor),$$

the last equality arising from the fact that $\bar{\rho}$ is identically zero off of $\mathbb{C}D_k$. Now, let us view $\bar{\rho}(v)$ as an $r_k \times r_k$ matrix with entries in $d_\rho \times d_\rho$ matrices (so we are viewing the rows and columns of $\bar{\rho}(v)$ as indexed by the idempotents in D_k). For idempotents $a, b \in D_k$, the b, a entry of $\bar{\rho}(v)$ is then the $d_\rho \times d_\rho$ matrix

$$\bar{\rho}(v)_{b,a} = \sum_{\substack{s \in D_k: \\ ss^{-1}=b \\ s^{-1}s=a}} v(s) \rho(p_b{}^{-1} s p_a).$$

By the proof of Theorem 5.1.1, this is the same as

$$\sum_{s \in G_k} v(p_b s p_a{}^{-1}) \rho(s). \tag{5.4}$$

If we define a function $h_{b,a} : G_k \to \mathbb{C}$ by

$$h_{b,a}(s) = v(p_b s {p_a}^{-1}),$$

we see that (5.4) is just the Fourier transform of the function $h_{b,a}$ on the group G_k at ρ. Notice that this holds regardless of the matrix realizations of the representations in $\mathrm{IRR}(G_k)$, so we can choose $\mathrm{IRR}(G_k)$ to be a computationally advantageous set of representations for these group Fourier transforms. Furthermore, once $\mathrm{IRR}(G_k)$ is chosen, it is clear from the above argument that the collection $\{\widehat{h_{b,a}}(\rho) : \rho \in \mathrm{IRR}(G_k)\}$ consists exactly of the blocks that compose $\{\hat{v}(\bar{\rho}) : \rho \in \mathrm{IRR}(G_k)\}$. An algorithm for computing the Fourier transform of v thus presents itself: for each $\mathcal{D}$-class D_k, run r_k^2 Fourier transforms on G_k, and then arrange the results into block form to construct the $\hat{v}(\bar{\rho})$. The latter step can be done for free in our computational model because it requires no operations. We therefore have:

Theorem 5.2.6. *Let S be a finite inverse semigroup with $\mathcal{D}$-classes $D_0, \ldots, D_n$. Let r_k denote the number of idempotents in D_k. For each D_k, choose an idempotent e_k and let G_k be the maximal subgroup at e_k. Let $\mathrm{IRR}(G_k)$ be a complete set of inequivalent, irreducible representations of G_k, and let $\mathcal{Y}$ be the set of inequivalent, irreducible representations of $\mathbb{C}S$ obtained by tensoring up the representations in $\biguplus_{k=0}^{n} \mathrm{IRR}(G_k)$. Then*

$$\mathcal{T}_{\mathcal{Y}}^{\text{groupoid}}(S) \leq \sum_{k=0}^{n} r_k^2 \mathcal{T}_{\mathrm{IRR}(G_k)}(G_k),$$

and since we may choose the $\mathrm{IRR}(G_k)$ at will, we have

$$\mathcal{C}^{\text{groupoid}}(S) \leq \sum_{k=0}^{n} r_k^2 \mathcal{C}(G_k).$$

Let us denote by $\mathcal{C}(\zeta_S)$ the maximal number of operations necessary to compute the change of basis in $\mathbb{C}S$ from the semigroup basis to the groupoid basis. We choose this notation because this change of basis is just the zeta transform on $\mathbb{C}S$, where S is viewed as a poset. Then, under the same hypothesis as in Theorem 5.2.6, we have:

Theorem 5.2.7.

$$\mathcal{C}^{\text{semigroup}}(S) \leq \mathcal{C}(\zeta_S) + \sum_{k=0}^{n} r_k^2 \mathcal{C}(G_k).$$

The problem of building fast Fourier transforms for finite inverse semigroups therefore reduces to the problems of building fast zeta transforms on their posets and building fast Fourier transforms on their maximal subgroups.

We end this section by considering the case when "good" FFTs (that is, $c|G|\log^d|G|$-complexity FFTs) are known for every maximal subgroup G of S.

Corollary 5.2.8. *Suppose S is an inverse semigroup with $\mathcal{D}$-classes $D_0, \ldots, D_n$. Let r_k be the number of idempotents in D_k. Choose an idempotent e_k from each $\mathcal{D}$-class D_k, and let G_k be the maximal subgroup at e_k. If $\mathcal{C}(G_k) \leq c_k|G_k|\log^{d_k}|G_k|$ for all k, then for some $d \in \mathbb{Z}$ we have*

$$\mathcal{C}^{\text{groupoid}}(S) \leq O(|S|\log^d|S|).$$

Proof. Let $c = \max_{k=0}^{n} c_k$, and let $d = \max_{k=0}^{n} d_k$. Then

$$\begin{aligned}
\mathcal{C}^{\text{groupoid}}(S) &\leq \sum_{k=0}^{n} r_k^2 c|G_k|\log^d|G_k| \\
&\leq c\log^d|S| \sum_{k=0}^{n} r_k^2|G_k| \\
&= c|S|\log^d|S|,
\end{aligned}$$

the last equality arising from (5.2). □

5.3 Natural Representations of the Rook Monoid

There are many ways to describe the irreducible representations of the symmetric group S_n. Indeed, we have already seen one description in Section 4.3.1. There is another description which gives rise to *Young's natural representations*, in which the irreducible representations are realized as certain submodules (called *Specht modules*) of certain easily described $\mathbb{C}S_n$-modules. In [15], C. Grood generalizes the notion of a Specht module and provides an analogous "natural" description of the irreducible representations of the rook monoid R_n. In this section, we show how Theorem 5.1.1 allows us to recover Grood's description directly from Young's natural representations of S_n. The upshot of our approach is that we do not have to prove (as Grood does) that the representations obtained are indeed representations, that they are irreducible or inequivalent, or that the resulting set of representations is complete, as these results follow from the discussion in Section 5.2. We begin by reviewing Young's natural representations, as described in Chapter 7 of [18].

Let λ be a partition of n. As in Section 4.3.1, we identify λ with its Young diagram.

Definition (tabloid)**.** A *tabloid* of shape λ is an equivalence class of tableaux of shape λ, where two tableaux represent the same tabloid if and only if they contain the same entries in every row. That is, a tabloid is just a tableau with unordered row entries. If t is a tableau, we denote its tabloid by $[t]$.

Now, let M^λ be the $\mathbb{C}$-span of the tabloids of shape λ. The action of $\mathbb{C}S_n$ on tableaux gives rise to a well-defined action of $\mathbb{C}S_n$ on tabloids, and hence M^λ is a $\mathbb{C}S_n$-module. The irreducible representation corresponding to λ that we seek (that

is, the Specht module S^λ) will be a submodule of M^λ. Before we can describe S^λ, however, we need another definition.

Definition (polytabloid)**.** Let t be a tableau of shape λ. Suppose λ has l columns. Let S_{c_i} be the permutation group on the elements in the i^{th} column of λ, and form $G = S_{c_1} \times S_{c_2} \times \cdots \times S_{c_l}$. That is, G is the group of permutations on the entries of t which permute entries within their columns. The element $E_t \in M^\lambda$, given by

$$E_t = \sum_{\sigma \in G} \mathrm{sgn}(\sigma)[\sigma(t)],$$

is called a *polytabloid* of shape λ.

For example, suppose t is the tableau $t = \begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array}$. Then

$$E_t = \left[\begin{array}{|c|c|c|}\hline 1 & 2 & 5 \\ \hline 3 & 4 \\ \cline{1-2}\end{array}\right] - \left[\begin{array}{|c|c|c|}\hline 3 & 2 & 5 \\ \hline 1 & 4 \\ \cline{1-2}\end{array}\right] - \left[\begin{array}{|c|c|c|}\hline 1 & 4 & 5 \\ \hline 3 & 2 \\ \cline{1-2}\end{array}\right] + \left[\begin{array}{|c|c|c|}\hline 3 & 4 & 5 \\ \hline 1 & 2 \\ \cline{1-2}\end{array}\right].$$

Note that the polytabloid E_t depends on the tableau t, not just the tabloid $[t]$.

Theorem 5.3.1 (Young)**.** *Let T be the set of standard tableaux of shape λ. Then $\{E_t\}_{t \in T}$ is a basis for an irreducible $\mathbb{C}S_n$-submodule S^λ of M^λ. Moreover, as λ varies across all partitions of n, the S^λ form a complete set of inequivalent, irreducible representations of S_n.*

The matrix representations of S_n afforded by this basis comprise Young's natural representations.

Now, let us turn to the rook monoid R_n. We begin by generalizing the above definitions, and we proceed as in [15].

Definition (n-tableau, n-standard tableau)**.** Let $\lambda \vdash k$. An n-tableau t is a filling of the Young diagram of λ with numbers from $\{1, \ldots, n\}$ such that each number

in t appears at most once. We say that t is *n-standard* if, furthermore, the entries in each column of t increase from top to bottom and the entries in each row of t increase from left to right.

Definition (n-tabloid)**.** Let $\lambda \vdash k$. An *n-tabloid* of shape λ is an equivalence class of n-tableaux of shape λ, where two n-tableaux represent the same n-tabloid if they contain the same entries in every row. That is, an n-tabloid is just an n-tableau with unordered row entries. If t is an n-tableau, we denote its n-tabloid by $[t]$.

Let $\bar{M}^\lambda$ denote the $\mathbb{C}$-span of the n-tabloids of shape λ. S_n acts on tableaux by permuting their entries, and this action naturally extends to an action of R_n on n-tableaux. Suppose that t is an n-tableau and $\sigma \in R_n$. If t contains an entry that is not in $\mathrm{dom}(\sigma)$, then we set $\sigma \cdot t$ to the zero vector. Otherwise, $\sigma \cdot t$ is just the n-tableau obtained by replacing the entries in t with their images under σ. This action of R_n on n-tableaux gives rise to a well-defined action on n-tabloids, making $\bar{M}^\lambda$ a $\mathbb{C}R_n$-module.

Definition (n-polytabloid)**.** Let t be an n-tableau of shape λ. Suppose λ has l columns. Let S_{c_i} be the permutation group on the elements in the i^{th} column of λ, and form $G = S_{c_1} \times S_{c_2} \times \cdots \times S_{c_l}$. That is, G is the group of permutations on the entries of t which permute entries within their columns. The element $E_t \in \bar{M}^\lambda$, given by

$$E_t = \sum_{\sigma \in G} \mathrm{sgn}(\sigma)[\sigma(t)],$$

is called an *n-polytabloid* of shape λ.

Theorem 5.3.2 (Grood)**.** *Let T be the set of n-standard tableaux of shape λ. Then $\{E_t\}_{t \in T}$ is a basis for an irreducible $\mathbb{C}R_n$-submodule $\bar{S}^\lambda$ of $\bar{M}^\lambda$. Moreover, as λ varies across all partitions of $\{0, 1, \ldots, n\}$, the $\bar{S}^\lambda$ form a complete set of inequivalent, irreducible representations of R_n.*

The matrix representations afforded by this basis comprise Grood's natural representations of R_n. Here is our proof of Theorem 5.3.2:

Proof. For a $\mathcal{D}$-class D_k of R_n (that is, the subset of elements of R_n of rank k), and let us take $e_k \in D_k$ to be the partial identity on $\{1, \ldots, k\}$. Let G_k be the maximal subgroup at e_k, so we have

$$G_k = \{s \in R_n : \mathrm{dom}(s) = \mathrm{ran}(s) = \{1, \ldots, k\}\}.$$

Let us identify G_k with the permutation group S_k in the obvious manner. For an idempotent $a \in D_k$ (that is, a rank-k restriction of the identity map), let us take p_a to be the unique order-preserving bijection from $\{1, \ldots, k\}$ to $\mathrm{dom}(a) = \mathrm{ran}(a)$. By Theorem 5.1.1, this defines our isomorphism

$$\mathbb{C}R_n \cong \bigoplus_{k=0}^{n} M_{\binom{n}{k}}(\mathbb{C}S_k). \tag{5.5}$$

Let $\mathcal{Y}_k$ denote Young's natural representations of S_k. As in Section 5.2, let us take $\mathcal{Y}$ to be the complete set of inequivalent, irreducible representations of $\mathbb{C}R_n$ obtained by tensoring up the representations in $\biguplus_{k=0}^{n} \mathcal{Y}_k$ to representations of $\bigoplus_{k=0}^{n} M_{\binom{n}{k}}(\mathbb{C}S_k)$ and following the isomorphism (5.5). This defines a complete set of inequivalent, irreducible representations of $\mathbb{C}R_n$ by defining them on the groupoid basis. To obtain representations of R_n, we need to express these representations on the semigroup basis. Once we do so, it will be apparent that the representations we have obtained are exactly those described in the theorem.

Let $\rho \in \mathcal{Y}_k$, and let $\bar{\rho} \in \mathcal{Y}$ be the corresponding representation of $\mathbb{C}R_n$. Suppose ρ corresponds to the partition λ of k. Let S^λ be the vector space on which S_k acts via ρ. Let $T = \{t_1, t_2, \ldots, t_m\}$ be the set of standard tableaux of shape λ, so that

the natural basis $\{E_{t_1}, E_{t_2}, \ldots, E_{t_m}\}$ of S^λ gives rise to the matrix representation ρ. Note that elements of T contain entries in $\{1, \ldots, k\}$. Let $\bar{S^\lambda}$ be the underlying vector space on which $\mathbb{C}R_n$ acts via $\bar{\rho}$. We denote the basis of $\bar{S^\lambda}$ giving rise to $\bar{\rho}$ by $\{E_{t_1}^{\{1,\ldots,k\}}, \ldots, E_{t_m}^{\{1,\ldots,k\}}, \ldots, E_{t_1}^{\{n-k+1,\ldots,n\}}, \ldots, E_{t_m}^{\{n-k+1,\ldots,n\}}\}$, where the top indices run across all k-subsets of $\{1, \ldots, n\}$. Let us first remark why we are justified in calling this space $\bar{S^\lambda}$. Suppose X is a k-subset of $\{1, \ldots, n\}$, and let p_X be the unique order-preserving bijection from $\{1, \ldots, k\}$ to X. If t_j is a standard tableau of shape λ, then $p_X \cdot t_j$ is an n-standard tableau of shape λ (it is the tableau obtained by replacing the entries in t_j with those in X, in order), so the basis element $E_{t_j}^X$ may be identified with the n-polytabloid $E_{p_X \cdot t_j}$. The basis of $\bar{S^\lambda}$ that we have obtained is therefore exactly the collection of n-polytabloids E_t, as t runs across the set of all n-standard tableaux. Thus, we will be done as soon as we can show that the action of R_n on $\bar{S^\lambda}$ given by $\bar{\rho}$ is the same as the natural action of R_n on n-polytabloids of shape λ.

So, how does the groupoid basis of $\mathbb{C}R_n$ act on this basis of $\bar{S^\lambda}$ via $\bar{\rho}$? Suppose $s \in R_n$, let X be a k-subset of $\{1, \ldots, n\}$, and let us examine $\bar{\rho}(\lfloor s \rfloor) \cdot E_{t_j}^X$. First, it is clear that $\bar{\rho}(\lfloor s \rfloor) \cdot E_{t_j}^X$ is 0 unless $\mathrm{rk}(s) = k$. So, suppose $\mathrm{rk}(s) = k$. Again, it is clear that $\bar{\rho}(\lfloor s \rfloor) \cdot E_{t_j}^X$ is 0 unless $X = \mathrm{dom}(s)$. So, suppose $\mathrm{dom}(s) = X$. Then $\bar{\rho}(\lfloor s \rfloor) \cdot E_{t_j}^{\mathrm{dom}(s)}$ is given by $\rho(\mathrm{perm}(s))$. Explicitly, it is

$$\bar{\rho}(\lfloor s \rfloor) \cdot E_{t_j}^{\mathrm{dom}(s)} = \sum_{i=1}^{m} \rho(\mathrm{perm}(s))_{i,j} \, E_{t_i}^{\mathrm{ran}(s)}. \tag{5.6}$$

Consider the following action of $\mathbb{C}R_n$ on n-tableaux, which we temporarily denote by $*$. Suppose that t is an n-tableau and $s \in R_n$. If the set of entries of t is not precisely $\mathrm{dom}(s)$, set $\lfloor s \rfloor * t$ to the zero vector. Otherwise, set $\lfloor s \rfloor * t$ equal to the n-tableau obtained by replacing the entries in t with their images under s.

Extend this linearly to an action of $\mathbb{C}R_n$. This gives rise to a well-defined action of $\mathbb{C}R_n$ on n-tabloids, and hence to an action of $\mathbb{C}R_n$ on $\bar{M}^\lambda$. If $[t]$ is an n-tabloid, we have

$$s = \sum_{\substack{x \in R_n: \\ x \leq s}} \lfloor x \rfloor \,,$$

so

$$s * [t] = \sum_{\substack{x \in R_n: \\ x \leq s}} \lfloor x \rfloor * [t],$$

and we see that $*$ is just the linear extension of the natural action of R_n on n-tabloids. It follows from this and (5.6) that the action of R_n on $\bar{S}^\lambda$ given by $\bar{\rho}$ is just the natural action of R_n on n-polytabloids of shape λ. Thus, the collection $\mathcal{Y}$ is a complete set of inequivalent, irreducible representations of R_n in Grood's "natural" form. $\square$

Chapter 6

More on Fourier Bases of Inverse Semigroup Algebras

Let S be a finite inverse semigroup. In this chapter, we embark on a detailed study of Fourier bases of the semigroup algebra $\mathbb{C}S$. In Section 6.1, we exploit Theorem 5.1.1 to explicitly describe a Fourier basis of $\mathbb{C}S$ in terms of Fourier bases of its maximal subgroup algebras $\mathbb{C}G_k$. In Section 6.2, we consider two "natural" inner products on $\mathbb{C}S$. We prove that the isotypic subspaces of $\mathbb{C}S$ are mutually orthogonal with respect to one of them and, in general, are not mutually orthogonal with respect to the other. Finally, in Section 6.3, we state and prove a general Fourier inversion theorem for finite inverse semigroups.

6.1 Explicit Fourier Basis Descriptions

Let S be a finite inverse semigroup. In this section, we explicitly describe a Fourier basis for the semigroup algebra $\mathbb{C}S$ in terms of Fourier bases for the group algebras $\mathbb{C}G_k$ of the maximal subgroups G_k of S. Fourier bases for these group algebras

may be found using the Fourier inversion theorem for groups (Theorem 6.3.1).

Let $D_0, \ldots, D_n$ be the $\mathcal{D}$-classes of S, and let r_k denote the number of idempotents in D_k. Pick an idempotent e_k in each $\mathcal{D}$-class D_k, and let G_k be the maximal subgroup of S at e_k. As in the proof of Theorem 5.1.1, for every idempotent $a \in D_k$, fix an element $p_a \in S$ such that ${p_a}^{-1} p_a = e_k$ and $p_a {p_a}^{-1} = a$ (and take $p_{e_k} = e_k$). Now let $\mathrm{IRR}(G_k)$ be a complete set of inequivalent, irreducible representations of $\mathbb{C}G_k$. For each $\rho \in \mathrm{IRR}(G_k)$, let $\bar{\rho}$ denote its extension (via "tensoring up," as discussed in Section 5.2) to $\bigoplus_{k=0}^{n} M_{r_k}(\mathbb{C}G_k)$, and hence to $\mathbb{C}S$. We take

$$\mathcal{Y} = \{\bar{\rho} : \rho \in \uplus_{k=0}^{n} \mathrm{IRR}(G_k)\},$$

so that $\mathcal{Y}$ is a complete set of inequivalent, irreducible representations of $\mathbb{C}S$.

It is now easy to explicitly describe the Fourier basis for $\mathbb{C}S$ according to $\mathcal{Y}$ in terms of the groupoid basis. That is, if $B \subseteq \mathbb{C}S$ is the set of inverse images of the natural basis of $\bigoplus_{\bar{\rho} \in \mathcal{Y}} M_{d_{\bar{\rho}}}(\mathbb{C})$ in the isomorphism

$$\bigoplus_{\bar{\rho} \in \mathcal{Y}} \bar{\rho} : \mathbb{C}S \to \bigoplus_{\bar{\rho} \in \mathcal{Y}} M_{d_{\bar{\rho}}}(\mathbb{C}), \tag{6.1}$$

then for each $y \in B$,

$$y = \sum_{s \in S} y(s) \lfloor s \rfloor .$$

We will now describe the coefficients $y(s)$.

We begin by assuming we have an explicit description of a Fourier basis for $\mathbb{C}G_k$ for each $k \in \{0, \ldots, n\}$. That is, if C is the set of inverse images of the natural basis of the algebra on the right in the isomorphism

$$\bigoplus_{\rho \in \mathrm{IRR}(G_k)} \rho : \mathbb{C}G_k \to \bigoplus_{\rho \in \mathrm{IRR}(G_k)} M_{d_\rho}(\mathbb{C}), \tag{6.2}$$

then, for each $c \in C$,

$$c = \sum_{x \in G_k} c(x)x,$$

and we assume that we know the coefficients $c(x)$. They may be found, for example, by using the standard Fourier inversion theorem for groups (Theorem 6.3.1).

Now, fix $\rho \in \mathrm{IRR}(G_k)$. Fix $c_{i,j} \in \mathbb{C}G_k$,

$$c_{i,j} = \sum_{x \in S_k} c_{i,j}(x)x,$$

to be the inverse image in the isomorphism (6.2) of the element of

$$\bigoplus_{\rho \in \mathrm{IRR}(G_k)} M_{d_\rho}(\mathbb{C})$$

that is 1 in the i, j position in the ρ block and 0 elsewhere. $\bar{\rho}$ is a block matrix whose rows and columns are indexed by the idempotents in D_k, and whose entries are themselves $d_\rho \times d_\rho$ matrices. Let a, b be idempotents in D_k.

Under the above hypothesis, we have the following description of a Fourier basis for $\mathbb{C}S$:

Theorem 6.1.1. *Let X be a $d_\rho \times d_\rho$ matrix with a 1 in the i, j position and 0 elsewhere. For idempotents $a, b \in D_k$, let $E_{b,a}$ be an $r_k \times r_k$ matrix with a 1 in the b, a position and 0 elsewhere. The inverse image in the isomorphism (6.1) of the element of $\bigoplus_{\bar{\rho} \in \mathcal{Y}} M_{d_{\bar{\rho}}}(\mathbb{C})$ that is $E_{b,a} \otimes X$ in the $\bar{\rho}$ block and 0 elsewhere is*

$$\lfloor p_b \rfloor \left(\sum_{x \in G_k} c_{i,j}(x) \lfloor x \rfloor \right) \lfloor {p_a}^{-1} \rfloor .$$

Proof. Suppose $\bar{\gamma} \in \mathcal{Y}$, $\bar{\gamma} \neq \bar{\rho}$, and $\gamma \in \mathrm{IRR}(G_k)$. Then

$$\bar{\gamma}\left(\lfloor p_b \rfloor \left(\sum_{x \in G_k} c_{i,j}(x)\, \lfloor x \rfloor\right) \left\lfloor {p_a}^{-1} \right\rfloor\right) = 0$$

because

$$\begin{aligned}
\bar{\gamma}\left(\sum_{x \in G_k} c_{i,j}(x)\, \lfloor x \rfloor\right) &= E_{e_k,e_k} \otimes \left(\sum_{x \in G_k} c_{i,j}(x)\gamma(x)\right) \\
&= E_{e_k,e_k} \otimes 0 \\
&= 0.
\end{aligned}$$

Suppose now that $\bar{\gamma} \in \mathcal{Y}$, $\bar{\gamma} \neq \bar{\rho}$, and $\gamma \in \mathrm{IRR}(G_j)$ with $j \neq k$. Then

$$\bar{\gamma}\left(\lfloor p_b \rfloor \left(\sum_{x \in G_k} c_{i,j}(x)\, \lfloor x \rfloor\right) \left\lfloor {p_a}^{-1} \right\rfloor\right) = 0$$

because

$$\begin{aligned}
\bar{\gamma}\left(\sum_{x \in G_k} c_{i,j}(x)\, \lfloor x \rfloor\right) &= \sum_{x \in G_k} c_{i,j}(x)\bar{\gamma}(\lfloor x \rfloor) \\
&= \sum_{x \in G_k} c_{i,j}(x)[0] \\
&= 0.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\bar{\rho}\left(\lfloor p_b \rfloor \left(\sum_{x\in S_k} c_{i,j}(x)\,\lfloor x \rfloor\right)\lfloor {p_a}^{-1}\rfloor\right)\\
=&\bar{\rho}\left((p_{p_bp_b{}^{-1}})^{-1}p_bp_{p_b{}^{-1}p_b}E_{p_bp_b{}^{-1},p_b{}^{-1}p_b}\right)\times\bar{\rho}\left(\sum_{x\in G_k} c_{i,j}(x)\,\lfloor x\rfloor\right)\times\\
&\bar{\rho}\left((p_{p_a{}^{-1}p_a})^{-1}{p_a}^{-1}p_{p_ap_a{}^{-1}}E_{p_a{}^{-1}p_a,p_ap_a{}^{-1}}\right)\\
=&\bar{\rho}\left({p_b}^{-1}p_bp_{e_k}E_{b,e_k}\right)\left(E_{e_k,e_k}\otimes\rho\left(\sum_{x\in G_k} c_{i,j}(x)x\right)\right)\bar{\rho}\left({p_{e_k}}^{-1}{p_a}^{-1}p_aE_{e_k,a}\right)\\
=&\bar{\rho}\left(e_kp_{e_k}E_{b,e_k}\right)\left(E_{e_k,e_k}\otimes X\right)\bar{\rho}\left(p_{e_k}e_kE_{e_k,a}\right)\\
=&\bar{\rho}\left(e_kE_{b,e_k}\right)\left(E_{e_k,e_k}\otimes X\right)\bar{\rho}\left(e_kE_{e_k,a}\right)\\
=&\left(E_{b,e_k}\otimes I_{d_\rho}\right)\left(E_{e_k,e_k}\otimes X\right)\left(E_{e_k,a}\otimes I_{d_\rho}\right)\\
=&E_{b,a}\otimes X.
\end{aligned}$$

□

6.2 Inner Products and Isotypic Subspaces

Let S be a finite inverse semigroup. Since $\mathbb{C}S$ is semisimple, it decomposes into a direct sum of irreducible submodules. This decomposition, however, need not be unique, and different choices of Fourier bases typically realize different decompositions. On the other hand, the *number* of subspaces of a given isomorphism type occurring in the decomposition is unique. Let $\mathcal{Y}$ be a complete set of inequivalent, irreducible representations of $\mathbb{C}S$. Let $\bar{\rho}\in\mathcal{Y}$. By the discussion in Section 3.3, there are $d_{\bar{\rho}}$ subspaces of $\mathbb{C}S$ isomorphic to the representation module for $\bar{\rho}$ in the decomposition of $\mathbb{C}S$ into irreducible subspaces. Let $V_{\bar{\rho}}$ denote the sum of these

subspaces. Then

$$\mathbb{C}S \cong \bigoplus_{\bar{\rho} \in \mathcal{Y}} V_{\bar{\rho}},$$

and the $V_{\bar{\rho}}$ are called the *isotypic subspaces* or *isotypic components* of $\mathbb{C}S$. They are unique, and do not depend on the choice of $\mathcal{Y}$. Calculating the Fourier transform of a function $f : S \to \mathbb{C}$ automatically computes the projection of f onto the $V_{\bar{\rho}}$, and we would like an inner product on $\mathbb{C}S$ (and hence on $\mathbb{C}$-valued functions on S, depending on the basis association used—see Section 3.4) under which this projection is orthogonal, i.e., under which the $V_{\bar{\rho}}$ are mutually orthogonal. In this section, we prove that the inner product induced by declaring the groupoid basis of $\mathbb{C}S$ mutually orthonormal accomplishes this while, in general, the inner product induced by declaring the semigroup basis of $\mathbb{C}S$ mutually orthonormal does not.

Theorem 6.2.1. *Let $\langle \cdot, \cdot \rangle$ be the sesquilinear form on $\mathbb{C}S$ induced by, for $s, t \in S$,*

$$\langle \lfloor s \rfloor , \lfloor t \rfloor \rangle = \begin{cases} 1 & \text{if } s = t, \\ 0 & \text{otherwise.} \end{cases}$$

Then, with respect to this inner product, the isotypic subspaces of $\mathbb{C}S$ are mutually orthogonal.

Proof. By linearity, it suffices to show that $\langle v, v' \rangle = 0$ in the case that v and v' are Fourier basis elements of $\mathbb{C}S$ in distinct isotypic subspaces. Let $D_0, \ldots, D_n$ be the $\mathcal{D}$-classes of S, let r_k denote the number of idempotents in D_k, and choose an idempotent e_k from each $\mathcal{D}$-class D_k. For every idempotent $a \in D_k$, fix an element $p_a \in S$ such that ${p_a}^{-1} p_a = e_k$ and $p_a {p_a}^{-1} = a$ (and take $p_{e_k} = e_k$). Let G_k be the maximal subgroup at e_k. We take $\mathcal{Y}$ to be the complete set of inequivalent, irreducible representations of $\mathbb{C}S$ induced by $\biguplus_{k=0}^{n} \mathrm{IRR}(G_k)$ in the

manner described in Section 5.2, and we assume that v and v' are part of a Fourier basis for $\mathbb{C}S$ according to $\mathcal{Y}$. Let $v \in V_{\bar{\rho}}$ and $v' \in V_{\bar{\rho}'}$, with $V_{\bar{\rho}} \neq V_{\bar{\rho}'}$ (and hence $\bar{\rho} \neq \bar{\rho}'$).

We know that

$$\mathbb{C}G_k = \bigoplus_{\rho \in \mathrm{IRR}(G_k)} W_\rho$$

where W_ρ is the sum of all the irreducible submodules of $\mathbb{C}G_k$ isomorphic to the representation module for ρ. If $w \in W_\rho, w' \in W_{\rho'}, W_\rho \neq W_{\rho'}$, then under the inner product $[\cdot, \cdot]$ on $\mathbb{C}G_k$ defined by

$$[w, w'] = [\sum_{s \in G_k} w(s)s, \sum_{s \in G_k} w'(s)s] = \sum_{s \in G_k} w(s)\overline{w'(s)},$$

it follows from the discussion in Chapter 2 of [35] that we have $[w, w'] = 0$.

Now, suppose that $\bar{\rho}$ was induced by $\rho \in \mathrm{IRR}(G_k)$ and that $\bar{\rho}'$ was induced by $\rho' \in \mathrm{IRR}(G_j)$. By Theorem 6.1.1, when written in terms of the groupoid basis, v contains nonzero coefficients only for the elements $\lfloor s \rfloor$ where s is in D_k, and v' contains nonzero coefficients only for the elements $\lfloor s \rfloor$ where s is in D_j. Thus, if $k \neq j$, we have $\langle v, v' \rangle = 0$. Suppose then that $k = j$. By Theorem 6.1.1, we have

$$v = \lfloor p_b \rfloor \sum_{s \in G_k} v(s) \lfloor s \rfloor \left\lfloor {p_a}^{-1} \right\rfloor,$$
$$v' = \lfloor p_{b'} \rfloor \sum_{s \in G_k} v'(s) \lfloor s \rfloor \left\lfloor {p_{a'}}^{-1} \right\rfloor,$$

for b, a, b', a' some idempotents in in D_k, and

$$\sum_{s \in G_k} v(s)s \in W_\rho, \sum_{s \in G_k} v'(s)s \in W_{\rho'}$$

some Fourier basis elements for $\mathbb{C}G_k$.

If $a \neq a'$ or $b \neq b'$, it is apparent that $\langle v, v' \rangle = 0$, so suppose further that $a = a'$ and $b = b'$.

Now, since $\bar{\rho} \neq \bar{\rho}'$ and $k = j$, we have $\rho \neq \rho'$, and we therefore note that

$$[\sum_{s \in G_k} v(s)s, \sum_{s \in G_k} v'(s)s] = 0.$$

Now, we have

$$\langle v, v' \rangle = \sum_{s \in G_k} \sum_{t \in G_k} v(s)\overline{v'(t)} \left\langle \left\lfloor p_b s {p_a}^{-1} \right\rfloor, \left\lfloor p_b t {p_a}^{-1} \right\rfloor \right\rangle,$$

and, since $s, t \in G_k$, $\left\lfloor p_b s {p_a}^{-1} \right\rfloor = \left\lfloor p_b t {p_a}^{-1} \right\rfloor$ if and only if $s = t$, so

$$\langle v, v' \rangle = \sum_{s \in G_k} v(s)\overline{v'(s)} = [\sum_{s \in G_k} v(s), \sum_{s \in G_k} v'(s)] = 0.$$

□

Remark 6.2.2. Let $\langle \cdot, \cdot \rangle$ be the sesquilinear form on $\mathbb{C}S$ induced by, for $s, t \in S$,

$$\langle s, t \rangle = \begin{cases} 1 & \text{if } s = t, \\ 0 & \text{otherwise.} \end{cases}$$

The isotypic subspaces of $\mathbb{C}S$ need not be orthogonal under this inner product. For example, consider $S = R_1$. $\mathbb{C}R_1$ has two nonisomorphic irreducible representations, each of degree 1, and hence a unique Fourier basis. It decomposes as

$$\mathbb{C}R_1 = (\mathbb{C}\text{-span}(\lfloor \text{Id} \rfloor)) \oplus (\mathbb{C}\text{-span}(\lfloor 0 \rfloor)).$$

Under this inner product, we have

$$\langle \lfloor \mathrm{Id} \rfloor, \lfloor 0 \rfloor \rangle = \langle \mathrm{Id} - (0), (0) \rangle = -1.$$

6.3 The Fourier Inversion Theorem

In this section, we state and prove a general Fourier inversion theorem for finite inverse semigroups. We begin by recalling the Fourier inversion theorem for finite groups.

Theorem 6.3.1. *Let G be a finite group, and $f = \sum_{s \in G} f(s)s \in \mathbb{C}G$. Let $\mathrm{IRR}(G)$ be a complete set of inequivalent, irreducible matrix representations of $\mathbb{C}G$. Then*

$$f(s) = \frac{1}{|G|} \sum_{\rho \in \mathrm{IRR}(G)} d_\rho \mathrm{trace}\left(\hat{f}(\rho)\rho(s^{-1})\right).$$

Proof. See [35], Section 6.2. □

Now, let S be a finite inverse semigroup. As usual, let $D_0, \ldots, D_n$ be the $\mathcal{D}$-classes of S, let r_k denote the number of idempotents in D_k, and choose an idempotent e_k from each $\mathcal{D}$-class D_k. For every idempotent $a \in D_k$, fix an element $p_a \in S$ such that ${p_a}^{-1} p_a = e_k$ and $p_a {p_a}^{-1} = a$ (and take $p_{e_k} = e_k$). Let G_k be the maximal subgroup at e_k.

We now state a preliminary lemma.

Lemma 6.3.2. *Let $f = \sum_{x \in S} f(x) \lfloor x \rfloor \in \mathbb{C}S$, and let $\mathcal{Y}$ be the complete set of inequivalent, irreducible matrix representations of $\mathbb{C}S$ induced by $\biguplus_{k=0}^{n} \mathrm{IRR}(G_k)$ in the manner described in Section 5.2. If $\rho \in \mathrm{IRR}(G_k)$, denote the corresponding*

representation of $\mathbb{C}S$ *by* $\bar{\rho}$*. Suppose* $x \in D_k$*. Let* $y \in G_k$ *be the element defined by*

$$y = {p_{xx^{-1}}}^{-1} x p_{x^{-1}x}.$$

Denote the idempotents xx^{-1} *and* $x^{-1}x$ *by* a *and* b*, respectively. Viewing* $\hat{f}(\bar{\rho})$ *as an* $r_k \times r_k$ *matrix whose entries are themselves* $d_\rho \times d_\rho$ *matrices, and whose rows and columns are indexed by the idempotents in* D_k*, let us denote the* a, b *entry of* $\hat{f}(\bar{\rho})$ *by* $\hat{f}(\bar{\rho})_{a,b}$*. Then we have*

$$f(x) = \frac{1}{|G_k|} \sum_{\rho \in \mathrm{IRR}(G_k)} d_\rho \mathrm{trace}\left(\hat{f}(\bar{\rho})_{a,b} \rho(y^{-1})\right).$$

Proof. For $\rho \in \mathrm{IRR}(G_k)$ we have

$$\hat{f}(\bar{\rho}) = \sum_{s \in S} f(s) \bar{\rho}(\lfloor s \rfloor),$$

with $\bar{\rho}(s) = 0$ if $s \notin D_k$. The a, b entry of $\hat{f}(\bar{\rho})$ is determined by the $f(s)$ for which $ss^{-1} = xx^{-1}$ and $s^{-1}s = x^{-1}x$, and such $f(s)$ do not affect any other entries of $\hat{f}(\bar{\rho})$. Explicitly, the a, b entry of $\hat{f}(\bar{\rho})$ is given by (as in (5.4)):

$$\hat{f}(\bar{\rho})_{a,b} = \sum_{s \in G_k} f(p_a s {p_b}^{-1}) \rho(s).$$

Let us define a function $f_{a,b} : G_k \to \mathbb{C}$ by

$$f_{a,b}(s) = f(p_a s {p_b}^{-1}).$$

Then

$$\sum_{s \in G_k} f(p_a s {p_b}^{-1}) \rho(s) = \sum_{s \in G_k} f_{a,b}(s) \rho(s) = \widehat{f_{a,b}}(\rho).$$

The Fourier inversion theorem for groups then applies, and yields:

$$\begin{aligned} f(x) &= f(p_{xx^{-1}} y {p_{x^{-1}x}}^{-1}) = f(p_a y {p_b}^{-1}) = f_{a,b}(y) \\ &= \frac{1}{|G_k|} \sum_{\rho \in \mathrm{IRR}(G_k)} d_\rho \mathrm{trace}\left(\widehat{f_{a,b}}(\rho) \rho(y^{-1}) \right), \end{aligned}$$

and since

$$\widehat{f_{a,b}}(\rho) = \hat{f}(\bar{\rho})_{a,b},$$

we are done. □

Now, let $\mathcal{X}$ be *any* set of inequivalent, irreducible matrix representations for $\mathbb{C}S$. Let $\mathcal{Y}$ be the complete set of inequivalent, irreducible matrix representations for $\mathbb{C}S$ induced by $\biguplus_{k=0}^{n} \mathrm{IRR}(G_k)$ in the manner described in Section 5.2. If $\rho \in \mathrm{IRR}(G_k)$, we have the corresponding $\bar{\rho} \in \mathcal{Y}$, which is equivalent to some representation in $\mathcal{X}$, which we denote by $\widetilde{\rho}$.

Theorem 6.3.3 (Fourier inversion theorem for $\mathbb{C}S$)**.** *Let*

$$f = \sum_{x \in S} f(x) \lfloor x \rfloor \in \mathbb{C}S.$$

Let $x \in D_k$, and let us denote the semigroup inverse of x by x^{-1}. Then

$$f(x) = \frac{1}{|G_k|} \sum_{\rho \in \mathrm{IRR}(G_k)} d_\rho \mathrm{trace}\left(\hat{f}(\widetilde{\rho}) \widetilde{\rho}(\lfloor x^{-1} \rfloor) \right).$$

Proof. Since $\widetilde{\rho}$ is equivalent to $\bar{\rho}$, write

$$\bar{\rho} = A^{-1} \widetilde{\rho} A$$

for some invertible matrix A. We therefore have

$$\hat{f}(\bar{\rho}) = A^{-1}\hat{f}(\widetilde{\rho})A.$$

As in Lemma 6.3.2, let $y \in G_k$ be the element defined by

$$y = {p_{xx^{-1}}}^{-1} x p_{x^{-1}x}.$$

Now, we have

$$\begin{aligned}
\text{trace}\left(\hat{f}(\bar{\rho})_{xx^{-1},x^{-1}x}\rho(y^{-1})\right) =& \text{trace}\left(\left[\hat{f}(\bar{\rho})\right]\left[E_{x^{-1}x,xx^{-1}} \otimes \rho(y^{-1})\right]\right) \\
=& \text{trace}\left(\hat{f}(\bar{\rho})\bar{\rho}(\lfloor x^{-1} \rfloor)\right) \\
=& \text{trace}\left(\left[A^{-1}\hat{f}(\widetilde{\rho})A\right]\left[A^{-1}\widetilde{\rho}(\lfloor x^{-1} \rfloor)A\right]\right) \\
=& \text{trace}\left(A^{-1}\hat{f}(\widetilde{\rho})\widetilde{\rho}(\lfloor x^{-1} \rfloor)A\right) \\
=& \text{trace}\left(\hat{f}(\widetilde{\rho})\widetilde{\rho}(\lfloor x^{-1} \rfloor)\right),
\end{aligned}$$

the last equality arising from the similarity-invariance of trace. The theorem now follows from Lemma 6.3.2. □

We can also state the Fourier inversion theorem without referencing $\text{IRR}(G_k)$.

Theorem 6.3.4 (Fourier inversion theorem for $\mathbb{C}S$, alternate version)**.** *Let*

$$f = \sum_{x \in S} f(x)\lfloor x \rfloor \in \mathbb{C}S.$$

Let $x \in D_k$, and let us denote the semigroup inverse of x by x^{-1}. Let $\mathcal{X}$ be any

complete set of inequivalent, irreducible representations of $\mathbb{C}S$. *Then*

$$f(x) = \frac{1}{|G_k|} \sum_{\widetilde{\rho} \in \mathcal{X}} \frac{d_{\widetilde{\rho}}}{r_k} \text{ trace} \left(\hat{f}(\widetilde{\rho}) \widetilde{\rho}(\lfloor x^{-1} \rfloor) \right).$$

Proof. Notice that, since $x \in D_k$, we also have $x^{-1} \in D_k$, and thus $\widetilde{\rho}(\lfloor x^{-1} \rfloor)$ is 0 unless $\rho \in \text{IRR}(G_k)$. If $\rho \in \text{IRR}(G_k)$, then we have $d_{\widetilde{\rho}} = d_{\bar{\rho}} = r_k d_\rho$. The theorem now follows from Theorem 6.3.3. □

Chapter 7

An FFT for the Rook Monoid

In this chapter, we develop a fast Fourier transform for the rook monoid R_n. Our approach in this chapter is based on Theorems 5.2.6 and 5.2.7. In particular, if we are given an arbitrary element $f \in \mathbb{C}R_n$ expressed with respect to the semigroup basis, we compute the Fourier transform of f by re-expressing it in terms of the groupoid basis and then calculating the Fourier transform of the resulting element. We deal with the latter operation first.

7.1 From the Groupoid Basis to a Fourier Basis

Let $f \in \mathbb{C}R_n$ be an arbitrary element, given with respect to the groupoid basis:

$$f = \sum_{s \in R_n} f(s) \lfloor s \rfloor .$$

We have already seen (Corollary 5.1.3) that

$$\mathbb{C}R_n \cong \bigoplus_{k=0}^{n} M_{\binom{n}{k}}(\mathbb{C}S_k).$$

Let $\mathcal{Y}_k$ denote Young's seminormal (or orthogonal) matrix representations for the symmetric group S_k. For our complete set of inequivalent, irreducible representations for $\mathbb{C}R_n$, we take the set $\mathcal{Y}$ induced by $\biguplus_{k\in\{0,\ldots,n\}} \mathcal{Y}_k$ in the manner described in Section 5.2. Theorem 5.2.6 implies that

$$\mathcal{C}^{\text{groupoid}}(R_n) \leq \sum_{k=0}^{n} \binom{n}{k}^2 \mathcal{C}(S_k).$$

Since we chose $\mathcal{Y}_k$ to be in Young's seminormal or orthogonal form, we can use Maslen's algorithm (see [21] or Theorem 4.3.5) for the Fourier transform on S_k. Doing so yields:

Theorem 7.1.1. *$\mathcal{C}^{\text{groupoid}}(R_n) \leq \frac{3}{4}n(n-1)|R_n|$, and hence*

$$\mathcal{C}^{\text{groupoid}}(R_n) = O(|R_n| \log^2 |R_n|).$$

Proof.

$$\begin{aligned}
\mathcal{C}^{\text{groupoid}}(R_n) &\leq \sum_{k=0}^{n} \binom{n}{k}^2 \frac{3}{4}k(k-1)|S_k| \\
&\leq \frac{3}{4}n(n-1)\sum_{k=0}^{n} \binom{n}{k}^2 |S_k| \\
&= \frac{3}{4}n(n-1)\sum_{k=0}^{n} \binom{n}{k}^2 k! \\
&= \frac{3}{4}n(n-1)|R_n|.
\end{aligned}$$

Since $|R_n| \geq n!$ and $n = O(\log(n!))$, we are done. □

7.2 From the Semigroup Basis to the Groupoid Basis

Let $f \in \mathbb{C}R_n$ be an arbitrary element, expressed with respect to the semigroup basis:

$$f = \sum_{s \in R_n} f(s)s.$$

We may express f with respect to the groupoid basis:

$$f = \sum_{s \in R_n} g(s) \lfloor s \rfloor ,$$

where, by (3.1), the coefficients $g(s)$ are given by

$$g(s) = \sum_{\substack{t \in R_n: \\ t \geq s}} f(t).$$

Our goal is to compute the coefficients $g(s)$ in an efficient manner, and in this section we give an algorithm for doing so. We note that the savings in time afforded by this algorithm come at the expense of a modest additional storage requirement over the naive algorithm—the algorithm presented here requires the storage of potentially $n|R_n|$ complex numbers in memory during runtime, as opposed to the naive algorithm, which requires at most $2|R_n|$.

We now present the proof of Theorem 2.4.4, as the algorithm presented in this section is based (at least in part) on the ideas involved in the proof.

Theorem (Theorem 2.4.4)**.** *For $n \geq 3$,*

$$|R_n| = 2n|R_{n-1}| - (n-1)^2|R_{n-2}|.$$

Proof. Viewing the elements of R_n as rook matrices, R_n consists of those elements having all 0's in column 1 and row 1 (of which there are $|R_{n-1}|$), together with, for each $\alpha \in \{1, \ldots, n\}$, those having a 1 in position $(\alpha, 1)$ (of which there are $n|R_{n-1}|$ total), together with, for each $\alpha \in \{2, \ldots, n\}$, those having a 1 in position $(1, \alpha)$ (of which there are $(n-1)|R_{n-1}|$ total). Counting the number of elements of R_n in this way overcounts. For each pair α, β with $2 \leq \alpha, \beta \leq n$, every element with ones in positions $(\alpha, 1)$ and $(1, \beta)$ (of which there are $(n-1)^2|R_{n-2}|$ total) gets counted twice. □

We now explain the fast zeta transform. Let us denote $\sum_{t \geq s} f(t)$ by $\zeta_f(s)$. The basic idea is to "work from the top down." Since we are trying to compute $\zeta_f(s)$ for all $s \in R_n$, it makes sense to begin with the elements of rank n. If $\mathrm{rk}(s) = n$, then there is no element t such that $t > s$, so $\zeta_f(s) = f(s)$, and this requires no operations. Next, if $\mathrm{rk}(s) = n-1$, then there is only one element $t \in R_n$ such that $t > s$, so

$$\zeta_f(s) = f(s) + f(t) = f(s) + \zeta_f(t).$$

Next, if $\mathrm{rk}(s) = n-2$, consider the poset consisting of the elements $t \in R_n$ for which $t \geq s$. This poset is isomorphic to the poset for R_2, with s in the place of the 0 element. We proceed down in rank in this manner, and the aim of this fast zeta transform is, in general, to find a way to re-use the $\zeta_f(t)$ we have already computed in order to efficiently compute $\zeta_f(s)$. In fact, instead of computing just $\zeta_f(s)$ for the elements s of rank k, we compute $\zeta_f(s)$ along with $n-k$ other numbers for each element s of rank k. These other numbers are needed for the efficient computation of the zeta transform at elements of lower rank, and can be discarded once all calculations are complete. We introduce some notation.

Let $s \in R_n$. Then s is a partial permutation of $\{1, 2, \ldots, n\}$.

- Let $d_i(s)$ be the i^{th} element of $\{1, 2, \ldots, n\}$ *not* in $\mathrm{dom}(s)$.
- Let $r_i(s)$ be the i^{th} element of $\{1, 2, \ldots, n\}$ *not* in $\mathrm{ran}(s)$.

That is, $d_i(s)$ is simply the i^{th} element of the complement of the domain of s, and similarly for $r_i(s)$. Define "partial" zeta transforms at s as follows:

$$\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_m(s)\}, \{r_1(s), r_2(s), \ldots, r_m(s)\}) = \sum_{\substack{t \geq s: \\ d_1(s), \ldots, d_m(s) \notin \mathrm{dom}(t) \\ r_1(s), \ldots, r_m(s) \notin \mathrm{ran}(t)}} f(t)$$

Our zeta transform proceeds as follows, with steps $0, 1, \ldots, n$:

- Step 0: For all $s \in R_n$ with $\mathrm{rk}(s) = n$, compute all $\zeta_f(s, \{\}, \{\}) = \zeta_f(s)$ (0 operations).
- Step 1: For all $s \in R_n$ with $\mathrm{rk}(s) = n - 1$, compute $\zeta_f(s, \{\}, \{\}) = \zeta_f(s)$ and $\zeta_f(s, \{d_1(s), r_1(s)\})$ (1 operation for each element s).

$$\vdots$$

- Step $n - k$: For all $s \in R_n$ with $\mathrm{rk}(s) = k$, compute all

$$\begin{aligned}
&\zeta_f(s, \{\}, \{\}) = \zeta_f(s), \\
&\zeta_f(s, \{d_1(s)\}, \{r_1(s)\}), \\
&\zeta_f(s, \{d_1(s), d_2(s)\}, \{r_1(s), r_2(s)\}), \\
&\vdots \\
&\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\}).
\end{aligned}$$

$$\vdots$$

Theorem 7.2.1. *Step $n-k$ requires at most*

$$\left((n-k)^2 + \frac{(n-k-1)(n-k)(2n-2k-1)}{6}\right)\binom{n}{k}^2 k!$$

operations in total.

Proof. We will show that, for an element $s \in R_n$ with $\mathrm{rk}(s) = k$, computing all

$$\begin{aligned}
&\zeta_f(s, \{\}, \{\}) = \zeta_f(s),\\
&\zeta_f(s, \{d_1(s)\}, \{r_1(s)\}),\\
&\zeta_f(s, \{d_1(s), d_2(s)\}, \{r_1(s), r_2(s)\}),\\
&\vdots\\
&\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\})
\end{aligned}$$

requires at most

$$(n-k)^2 + \frac{(n-k-1)(n-k)(2n-2k-1)}{6}$$

additions, assuming that steps $0, 1, \ldots, n-k-1$ have already been completed.

Let $s * (d_i(s) \to r_j(s))$ denote the element of R_n that is obtained by adding $d_i(s)$ to the domain of s and sending it to $r_j(s)$. For example, if

$$s = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & - & 5 & - & - & - & 3 \end{pmatrix},$$

then

$$s * (d_2(s) \to r_3(s)) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & - & 5 & 6 & - & - & 3 \end{pmatrix}.$$

Now, consider the poset of elements $t \in R_n$ with $t \geq s$. This poset is isomorphic to the poset for R_{n-k}, with s in the place of the 0 element. Following the idea in the proof of Theorem 2.4.4, we have:

$$\begin{aligned}
\zeta_f(s, \{\}, \{\}) = \zeta_f(s * (d_1(s) \to r_1(s)), \{\}, \{\}) \\
+ \zeta_f(s * (d_2(s) \to r_1(s)), \{\}, \{\}) + \cdots \\
+ \zeta_f(s * (d_{n-k}(s) \to r_1(s)), \{\}, \{\}) \\
+ \zeta_f(s * (d_1(s) \to r_2(s)), \{\}, \{\}) + \cdots \\
+ \zeta_f(s * (d_1(s) \to r_{n-k}(s)), \{\}, \{\}) \\
- \sum_{i,j \in \{2,\ldots,n-k\}} \zeta_f(s * (d_i(s) \to r_1(s)) * (d_1(s) \to r_j(s)), \{\}, \{\}) \\
+ \zeta_f(s, \{d_1(s)\}, \{r_1(s)\}).
\end{aligned}$$

Notice that every term in this sum, with the exception of $\zeta_f(s, \{d_1(s)\}, \{r_1(s)\})$, was computed in an earlier step. After all, $\mathrm{rk}(s * (d_i(s) \to r_j(s))) > \mathrm{rk}(s)$ for all i, j. Thus, once we have $\zeta_f(s, \{d_1(s)\}, \{r_1(s)\})$, all we have to do to compute $\zeta_f(s)$ is add these terms up. To compute $\zeta_f(s, \{d_1(s)\}, \{r_1(s)\})$, we again follow the idea

in the proof of Theorem 2.4.4, and we write:

$$\begin{aligned}\zeta_f(s,\{d_1(s)\},\{r_1(s)\}) &= \zeta_f(s*(d_2(s)\to r_2(s)),\{d_1(s)\},\{r_1(s)\})\\ &+\zeta_f(s*(d_3(s)\to r_2(s)),\{d_1(s)\},\{r_1(s)\})+\cdots\\ &+\zeta_f(s*(d_{n-k}(s)\to r_2(s)),\{d_1(s)\},\{r_1(s)\})\\ &+\zeta_f(s*(d_2(s)\to r_3(s)),\{d_1(s)\},\{r_1(s)\})+\cdots\\ &+\zeta_f(s*(d_2(s)\to r_{n-k}(s)),\{d_1(s)\},\{r_1(s)\})\\ &-\sum_{i,j\in\{3,\ldots,n-k\}}\zeta_f(s*(d_i(s)\to r_2(s))*\\ &\qquad(d_2(s)\to r_j(s)),\{d_1(s)\},\{r_1(s)\})\\ &+\zeta_f(s,\{d_1(s),d_2(s)\},\{r_1(s),r_2(s)\}).\end{aligned}$$

Notice that $i,j\geq 2$ implies

$$\begin{aligned}&\zeta_f(s*(d_i(s)\to r_j(s)),\{d_1(s)\},\{r_1(s)\})\\ &=\zeta_f(s*(d_i(s)\to r_j(s)),\{d_1(s*(d_i(s)\to r_j(s)))\},\{r_1(s*(d_i(s)\to r_j(s)))\}),\end{aligned}$$

so every term in this sum, with the exception of $\zeta_f(s,\{d_1(s),d_2(s)\},\{r_1(s),r_2(s)\})$, was computed in an earlier step.

In general, suppose $D=\{d_1(s),d_2(s),\ldots,d_m(s)\}$, $R=\{r_1(s),r_2(s),\ldots,r_m(s)\}$. If $m=n-k$, then we have

$$\zeta_f(s,D,R)=f(s).$$

If $m = n - k - 1$, then we have

$$\begin{aligned}\zeta_f(s, D, R) = \zeta_f(s * (d_{n-k}(s) \to r_{n-k}(s), D, R) \\ + \zeta_f(s, D \cup \{d_{n-k}(s)\}, R \cup \{r_{n-k}(s)\}).\end{aligned}$$

Otherwise, $m < n - k - 1$, and we have

$$\begin{aligned}\zeta_f(s, D, R) = \zeta_f(s * (d_{m+1}(s) \to r_{m+1}(s)), D, R) \\ + \zeta_f(s * (d_{m+2}(s) \to r_{m+1}(s)), D, R) + \cdots \\ + \zeta_f(s * (d_{n-k}(s) \to r_{m+1}(s)), D, R) \\ + \zeta_f(s * (d_{m+1}(s) \to r_{m+2}(s)), D, R) + \cdots \\ + \zeta_f(s * (d_{m+1}(s) \to r_{n-k}(s)), D, R) \\ - \sum_{i,j \in \{m+2,\ldots,n-k\}} \zeta_f(s * (d_i(s) \to r_{m+1}(s)) * \\ (d_{m+1}(s) \to r_j(s)), D, R) \\ + \zeta_f(s, D \cup \{d_{m+1}(s)\}, R \cup \{r_{m+1}(s)\}),\end{aligned}$$

where every term in the sum, with the exception of

$$\zeta_f(s, D \cup \{d_{m+1}(s)\}, R \cup \{r_{m+1}(s)\}),$$

was computed in an earlier step of the algorithm. Once we have computed $\zeta_f(s, D \cup \{d_{m+1}(s)\}, R \cup \{r_{m+1}(s)\})$, the number of operations required to compute $\zeta_f(s, D, R)$ is thus no more than

$$(n - k - m) + (n - k - m - 1) + (n - k - m - 1)^2.$$

We do this for m from $n-k$ to 0 to compute, in order:

$$\begin{aligned}
&\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\}),\\
&\vdots\\
&\zeta_f(s, \{d_1(s), d_2(s)\}, \{r_1(s), r_2(s)\}),\\
&\zeta_f(s, \{d_1(s)\}, \{r_1(s)\}),\\
&\zeta_f(s, \{\}, \{\}) = \zeta_f(s),
\end{aligned}$$

which is what we want. The total number of operations required is thus no more than

$$\begin{aligned}
&\sum_{m=0}^{n-k} (n-k-m) + (n-k-m-1) + (n-k-m-1)^2\\
&= (n-k)^2 + \frac{(n-k-1)(n-k)(2n-2k-1)}{6}.
\end{aligned}$$

□

This algorithm yields the bound:

Theorem 7.2.2. $\mathcal{C}(\zeta_{R_n}) = O(|R_n| \log^3 |R_n|)$.

Proof. For $n \geq 3$,

$$\begin{aligned}
(n-k)^2 + \frac{(n-k-1)(n-k)(2n-2k-1)}{6} &\leq n^2 + \frac{2n^3}{6}\\
&\leq \frac{2}{3} n^3.
\end{aligned}$$

Hence

$$\mathcal{C}(\zeta_{R_n}) \leq \frac{2}{3} n^3 |R_n|,$$

and since $|R_n| \geq n!$ and $n = O(\log(n!))$, we are done. □

Combining this algorithm with the algorithm in Section 7.1, we obtain

Theorem 7.2.3. $\mathcal{C}^{\text{semigroup}}(R_n) = O(|R_n| \log^3 |R_n|)$.

Proof. We have

$$\begin{aligned}\mathcal{C}^{\text{semigroup}}(R_n) &\leq \mathcal{C}^{\text{groupoid}}(R_n) + \mathcal{C}(\zeta_{R_n}) \\ &\leq \frac{3}{4}n^2|R_n| + \frac{2}{3}n^3|R_n|.\end{aligned}$$

□

Chapter 8

FFTs for Rook Wreath Products

In this chapter, we study the wreath products of R_n by arbitrary finite groups G, and we exhibit fast algorithms for their Fourier transforms.

8.1 Properties of Wreath Products

Let G be a finite group.

Definition. The *rook wreath product* $G \wr R_n$ is the semigroup of all $n \times n$ matrices with entries in $\{0\} \cup G$ having at most one non-zero entry per row and column. The operation on $G \wr R_n$ is matrix multiplication.

Clearly, we recover the rook monoid as $Z_1 \wr R_n$. Let us write 1 for the identity of G. It is easy to see that the idempotents of $G \wr R_n$ are precisely the idempotents of R_n.

Theorem 8.1.1. *$G \wr R_n$ is an inverse semigroup.*

Proof. Let $s \in G \wr R_n$. Let $i_1, \ldots, i_k$ be the row indices of the nonzero entries of s, and let $j_1, \ldots, j_k$ be the column indices of the nonzero entries of s. For

$x \in \{1, \ldots, k\}$, then, s_{i_x,j_x} is an element of G. Define $t \in G \wr R_n$, for $x \in \{1, \ldots, k\}$, by

$$t_{i_x,j_x} = {s_{j_x,i_x}}^{-1}.$$

That is, t is given by transposing s and replacing the resulting nonzero entries with their G-inverses. Then t is a semigroup inverse of s and hence $G \wr R_n$ is regular. To prove that $G \wr R_n$ is indeed an inverse semigroup, we appeal to the alternate definition given in Section 2.3, which states that an inverse semigroup is a regular semigroup with commuting idempotents. Since the idempotents of $G \wr R_n$ are just the idempotents of R_n, they commute, and hence $G \wr R_n$ is inverse. □

We generalize the notion of rank to $G \wr R_n$:

Definition. For an element $s \in G \wr R_n$, define the *rank* of s, denoted $\mathrm{rk}(s)$, to be the number of rows of s which contain nonzero entries.

By definition of $G \wr R_n$, $\mathrm{rk}(s)$ is the same as the number of columns of s which contain nonzero entries.

Definition. The *symmetric group wreath product* $G \wr S_n$ is the group of all $n \times n$ matrices with entries in $\{0\} \cup G$ having exactly one non-zero entry per row and column. The operation on $G \wr S_n$ is matrix multiplication.

Thus, $G \wr S_n$ is contained in $G \wr R_n$ as the rank-n elements.

We also generalize the notion of domain and range from R_n to $G \wr R_n$:

Definition. Let $s \in G \wr R_n$. Define $\mathrm{dom}(s)$ to be the set of indices of the columns of s which contain nonzero entries, and $\mathrm{ran}(s)$ to be the set of indices of the rows of s which contain nonzero entries.

This definition agrees with our previous definitions of inverse semigroup domain and range, that is,

$$\mathrm{dom}(s) = s^{-1}s, \quad \mathrm{ran}(s) = ss^{-1},$$

provided that we once again abuse the distinction between the domain and range of a map and the corresponding partial identities (as elements of R_n).

In order to apply Theorem 5.2.7 to create an FFT for $G \wr R_n$, we must understand the maximal subgroups of $G \wr R_n$. Let $e \in G \wr R_n$ be idempotent, with $\mathrm{rk}(e) = k$.

Theorem 8.1.2. *The maximal subgroup of $G \wr R_n$ at e is isomorphic to $G \wr S_k$.*

Proof. Denote this subgroup by G_e. We have

$$G_e = \{s \in G \wr R_n : ss^{-1} = s^{-1}s = e\}.$$

Suppose that $\mathrm{dom}(e) = \{i_1, \ldots, i_k\}$ (and hence $\mathrm{ran}(e) = \{i_1, \ldots, i_k\}$, because e is idempotent). Clearly, then,

$$\{x \in G \wr R_n : \mathrm{dom}(x) = \mathrm{ran}(x) = \{i_1, \ldots, i_k\}\} \subseteq G_e.$$

If $x \in G_e$, then $\mathrm{dom}(x) = \mathrm{ran}(x)$. Furthermore, if $\mathrm{rk}(x) \neq k$, then $x \notin G_e$. If $j \in \mathrm{dom}(x)$ with $j \notin \{i_1, \ldots, i_k\}$, then $x^{-1}x \neq e$, and so $x \notin G_e$. Thus

$$G_e = \{x \in G \wr R_n : \mathrm{dom}(x) = \mathrm{ran}(x) = \{i_1, \ldots, i_k\}\},$$

which is isomoprhic to $G \wr S_k$ in the obvious way (i.e., for $x \in G_e$, delete the rows and columns of x which contain only zeroes). □

We will also need to understand the poset structure of $G \wr R_n$.

Theorem 8.1.3. *Let $s, t \in G \wr R_n$. Then $s \leq t$ if and only if s may be obtained by replacing entries in t with 0.*

Proof. This follows directly from our definition

$$s \leq t \iff s = et \text{ for some idempotent } e \in G \wr R_n,$$

together with the fact that the idempotents of $G \wr R_n$ are the idempotents of R_n (i.e., the restrictions of the identity matrix). □

Finally, we record the size of $G \wr R_n$:

Theorem 8.1.4.

$$|G \wr R_n| = \sum_{k=0}^{n} \binom{n}{k}^2 k! |G|^k.$$

Proof. There are $\binom{n}{k}^2 k!$ rook matrices of rank k, and for a given rook matrix X of rank k, there are $|G|$ options to replace each of the 1's in X with elements of G. □

8.2 From the Groupoid Basis to a Fourier Basis

Let $f \in \mathbb{C}G \wr R_n$ be given with respect to the groupoid basis:

$$f = \sum_{s \in G \wr R_n} f(s) \lfloor s \rfloor .$$

Theorem 5.1.1 and the discussion in Section 8.1 imply that we have

$$\mathbb{C}G \wr R_n \cong \bigoplus_{k=0}^{n} M_{\binom{n}{k}}(\mathbb{C}G \wr S_k),$$

and Theorem 5.2.6 yields

$$\mathcal{C}^{\text{groupoid}}(G \wr R_n) \leq \sum_{k=0}^{n} \binom{n}{k}^2 \mathcal{C}(G \wr S_k).$$

In [32], D. Rockmore constructs a computationally advantageous and complete set of inequivalent, irreducible representations $\mathcal{Y}_k$ for $G \wr S_k$ which we tensor up to be our set of representations $\mathcal{Y}$ for $\mathbb{C}G \wr R_n$, and proves the following (Corollary 2 of [32]).

Theorem 8.2.1. *Let h denote the number of inequivalent, irreducible representations of G. Then*

$$\mathcal{T}_{\mathcal{Y}_k}(G \wr S_k) \leq k!|G|^k \cdot \left[\frac{\mathcal{C}(G)}{|G|} \cdot \frac{k(k+1)}{2} + 2^h \frac{k^2(k+1)^2}{4} + 1\right].$$

In particular, $\mathcal{C}(G \wr S_k)$ is bounded by the same amount. Thus:

Theorem 8.2.2. *We have*

$$\mathcal{C}^{\text{groupoid}}(G \wr R_n) = O(|G \wr R_n| \log^4 |G \wr R_n|).$$

Proof. We have

$$\begin{aligned}
\mathcal{T}_{\mathcal{Y}}^{\text{groupoid}}(G \wr R_n) &\leq \sum_{k=0}^{n} \binom{n}{k}^2 k!|G|^k \cdot \left[\frac{\mathcal{C}(G)}{|G|} \cdot \frac{k(k+1)}{2} + 2^h \frac{k^2(k+1)^2}{4} + 1\right] \\
&\leq \left[\frac{\mathcal{C}(G)}{|G|} \cdot \frac{n(n+1)}{2} + 2^h \frac{n^2(n+1)^2}{4} + 1\right] \cdot \sum_{k=0}^{n} \binom{n}{k}^2 k!|G|^k \\
&\leq \left[\frac{\mathcal{C}(G)}{|G|} \cdot \frac{n(n+1)}{2} + 2^h \frac{n^2(n+1)^2}{4} + 1\right] \cdot |G \wr R_n|.
\end{aligned}$$

Now, $|G|$, $\mathcal{C}(G)$, and 2^h are constants with respect to n, and $n = O(\log |G \wr R_n|)$. The theorem follows. □

8.3 From the Semigroup Basis to the Groupoid Basis

Let G be a finite group. Let $f \in \mathbb{C}G \wr R_n$ be an arbitrary element, expressed with respect to the semigroup basis:

$$f = \sum_{s \in G \wr R_n} f(s)s.$$

We may express f with respect to the groupoid basis:

$$f = \sum_{s \in G \wr R_n} g(s) \lfloor s \rfloor,$$

where, by (3.1), the coefficients $g(s)$ are given by

$$g(s) = \sum_{\substack{t \in G \wr R_n: \\ t \geq s}} f(t).$$

In this section, we give a fast algorithm for computing the coefficients $g(s)$. The algorithm presented in this section reduces to the algorithm for the rook monoid given in Section 7.2 when $G = \mathbb{Z}_1$. As with that algorithm, the time savings afforded by this algorithm come at the expense of an additional storage requirement over the naive algorithm. The algorithm presented here requires storage of potentially $n|G \wr R_n|$ complex numbers in memory during runtime, as opposed to the naive algorithm, which requires at most $2|G \wr R_n|$.

As in Section 7.2, we begin with a recursive formula for the size of $G \wr R_n$.

Theorem 8.3.1. *For $n \geq 3$,*

$$|G \wr R_n| = (2n-1)|G||G \wr R_{n-1}| + |G \wr R_{n-1}| - (n-1)^2|G|^2|G \wr R_{n-2}|.$$

Proof. $G \wr R_n$ consists of those elements having all 0's in column 1 and row 1 (of which there are $|G \wr R_{n-1}|$), together with, for each $x \in G$ and $\alpha \in \{1, \ldots, n\}$, those having an x in position $(\alpha, 1)$ (of which there are $n|G||G \wr R_{n-1}|$ total), together with, for each $x \in G$ and $\alpha \in \{2, \ldots, n\}$, those having an x in position $(1, \alpha)$ (of which there are $(n-1)|G||G \wr R_{n-1}|$ total). Counting the number of elements of $G \wr R_n$ in this way overcounts. For each pair α, β with $2 \leq \alpha, \beta \leq n$ and for each pair of elements $x, y \in G$, every element with x position $(\alpha, 1)$ and y in position $(1, \beta)$ (of which there are $(n-1)^2|G|^2|G \wr R_{n-2}|$ total) gets counted twice. □

We now explain the fast zeta transform. As in Section 7.2, let us denote $\sum_{t \geq s} f(t)$ by $\zeta_f(s)$. Let $s \in G \wr R_n$. The rows and columns of s are indexed by $\{1, 2, \ldots, n\}$.

- Let $d_i(s)$ be the index of the i^{th} column of s which contains only zeroes.
- Let $r_i(s)$ be the index of the i^{th} row of s which contains only zeroes.

Define "partial" zeta transforms at s as follows:

$$\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_m(s)\}, \{r_1(s), r_2(s), \ldots, r_m(s)\}) = \sum_{\substack{t \geq s: \\ \text{columns } d_1(s),\ldots,d_m(s) \text{ of } t \text{ contain only zeroes and} \\ \text{rows } r_1(s),\ldots,r_m(s) \text{ of } t \text{ contain only zeroes}}} f(t).$$

As with R_n, we work from the "top" down, and our zeta transform proceeds as follows, with steps $0, 1, \ldots, n$:

- Step 0: For all $s \in G \wr R_n$ with $\mathrm{rk}(s) = n$, compute all $\zeta_f(s, \{\}, \{\}) = \zeta_f(s)$ (0 operations).

- Step 1: For all $s \in G \wr R_n$ with $\mathrm{rk}(s) = n - 1$, compute $\zeta_f(s, \{\}, \{\}) = \zeta_f(s)$ and $\zeta_f(s, \{d_1(s), r_1(s)\})$ ($|G|$ operations for each element s).

$$\vdots$$

- Step $n - k$: For all $s \in G \wr R_n$ with $\mathrm{rk}(s) = k$, compute all

$$\begin{aligned}
&\zeta_f(s, \{\}, \{\}) = \zeta_f(s),\\
&\zeta_f(s, \{d_1(s)\}, \{r_1(s)\}),\\
&\zeta_f(s, \{d_1(s), d_2(s)\}, \{r_1(s), r_2(s)\}),\\
&\vdots\\
&\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\}).
\end{aligned}$$

$$\vdots$$

Thus, instead of computing just $\zeta_f(s)$ for the elements s of rank k, we compute $\zeta_f(s)$ along with $n - k$ other numbers for each element s of rank k. These other numbers are needed for the efficient computation of the zeta transform at elements of lower rank, and can be discarded once all calculations are complete.

Theorem 8.3.2. *Step $n - k$ requires at most*

$$\left(|G|(n-k)^2 + |G|^2 \frac{(n-k-1)(n-k)(2n-2k-1)}{6}\right) \binom{n}{k}^2 k! |G|^k$$

operations in total.

Proof. We will show that, for an element $s \in G \wr R_n$ with $\mathrm{rk}(s) = k$, computing all

$$\begin{aligned}
&\zeta_f(s, \{\}, \{\}) = \zeta_f(s),\\
&\zeta_f(s, \{d_1(s)\}, \{r_1(s)\}),\\
&\zeta_f(s, \{d_1(s), d_2(s)\}, \{r_1(s), r_2(s)\}),\\
&\vdots\\
&\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\})
\end{aligned}$$

requires at most

$$|G|(n-k)^2 + |G|^2 \frac{(n-k-1)(n-k)(2n-2k-1)}{6}$$

additions, assuming that steps $0, 1, \ldots, n-k-1$ have already been completed.

Suppose $G = \{g_1, g_2, \ldots, g_{|G|}\}$. Let $s*(g_y E_{r_j(s), d_i(s)})$ denote the element of $G \wr R_n$ that is obtained by inserting g_y into the $r_j(s), d_i(s)$ position of s. For example, if

$$s = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0\\
g_{y_1} & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & g_{y_3}\\
0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & g_{y_2} & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

then

$$d_1(s) = 2, \quad r_1(s) = 1,$$
$$d_2(s) = 4, \quad r_2(s) = 4,$$
$$d_3(s) = 5, \quad r_3(s) = 6,$$
$$d_4(s) = 6, \quad r_4(s) = 7,$$

and

$$s * (g_{y_4} E_{r_3(s),d_2(s)}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ g_{y_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & g_{y_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & g_{y_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g_{y_4} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now, consider the poset of elements $t \in G \wr R_n$ with $t \geq s$. This poset is isomorphic to the poset for $G \wr R_{n-k}$, with s in the place of the 0 matrix.

As in Section 7.2, we compute our $n-k$ partial zeta transforms at s in the following order. First, we have

$$\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\}) = f(s),$$

which requires no operations. Next, let $D = \{d_1(s), d_2(s), \ldots, d_{n-k-1}(s)\}$ and $R = \{r_1(s), r_2(s), \ldots, r_{n-k-1}(s)\}$. We have

$$\begin{aligned} \zeta_f(s, D, R) = &\sum_{i=1}^{|G|} \zeta_f(s * (g_i E_{r_{n-k}(s),d_{n-k}(s)}), D, R) \\ &+ \zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\}), \end{aligned}$$

which requires $|G|$ operations.

Now, suppose $D = \{d_1(s), d_2(s), \ldots, d_m(s)\}$ and $R = \{r_1(s), r_2(s), \ldots, r_m(s)\}$, with $m < n - k - 1$. Following the proof of Theorem 8.3.1, we have

$$\begin{aligned}
\zeta_f(s, D, R) = &\sum_{i=1}^{|G|} \zeta_f(s * (g_i E_{r_{m+1}(s), d_{m+1}(s)}), D, R) \\
&+ \sum_{i=1}^{|G|} \zeta_f(s * (g_i E_{r_{m+1}(s), d_{m+2}(s)}), D, R) \\
&+ \cdots \\
&+ \sum_{i=1}^{|G|} \zeta_f(s * (g_i E_{r_{m+1}(s), d_{n-k}(s)}), D, R) \\
&+ \sum_{i=1}^{|G|} \zeta_f(s * (g_i E_{r_{m+2}(s), d_{m+1}(s)}), D, R) \\
&+ \cdots \\
&+ \sum_{i=1}^{|G|} \zeta_f(s * (g_i E_{r_{n-k}(s), d_{m+1}(s)}), D, R) \\
&- \sum_{\substack{i,j \in \{m+2,\ldots,n-k\} \\ k,l \in \{1,2,\ldots,|G|\}}} \zeta_f(s * (g_k E_{r_{m+1}(s), d_i(s)}) * (g_l E_{r_j(s), d_{m+1}(s)}), D, R) \\
&+ \zeta_f(s, D \cup \{d_{m+1}(s)\}, R \cup \{r_{m+1}(s)\}).
\end{aligned}$$

Computing $\zeta_f(s, D, R)$ therefore requires no more than

$$|G|(2n - 2k - 2m - 1) + |G|^2(n - k - m - 1)^2$$

operations. We do this for $m = n - k$ to 0 to compute, in order,

$$\begin{aligned}
&\zeta_f(s, \{d_1(s), d_2(s), \ldots, d_{n-k}(s)\}, \{r_1(s), r_2(s), \ldots, r_{n-k}(s)\}),\\
&\vdots\\
&\zeta_f(s, \{d_1(s), d_2(s)\}, \{r_1(s), r_2(s)\}),\\
&\zeta_f(s, \{d_1(s)\}, \{r_1(s)\}),\\
&\zeta_f(s, \{\}, \{\}) = \zeta_f(s),
\end{aligned}$$

which is what we want. The total number of operations required to compute these is thus no more than

$$\begin{aligned}
&|G| + \sum_{m=0}^{n-k-2} |G|(2n - 2k - 2m - 1) + |G|^2(n - k - m - 1)^2\\
&= |G|(n-k)^2 + |G|^2 \frac{(n-k-1)(n-k)(2n-2k-1)}{6}.
\end{aligned}$$

□

This result should be compared to the result for the zeta transform on R_n in Section 7.2. This algorithm yields the bound:

Theorem 8.3.3. $\mathcal{C}(\zeta_{G \wr R_n}) = O(|G \wr R_n| \log^3 |G \wr R_n|)$.

Proof. For $n \geq 3$,

$$\begin{aligned}
|G|(n-k)^2 + |G|^2 \frac{(n-k-1)(n-k)(2n-2k-1)}{6} &\leq |G|n^2 + |G|^2 \frac{2n^3}{6}\\
&\leq \frac{2}{3}|G|^2 n^3.
\end{aligned}$$

Hence

$$\mathcal{C}(\zeta_{G \wr R_n}) \leq \frac{2}{3}|G|^2 n^3 |G \wr R_n|.$$

Since $|G|$ is a constant with respect to n and $n = O(\log |G \wr R_n|)$, we are done. □

Combining this algorithm with the algorithm in Section 8.2, we obtain:

Theorem 8.3.4. $\mathcal{C}^{\text{semigroup}}(G \wr R_n) = O(|G \wr R_n| \log^4 |G \wr R_n|)$.

Proof. We have

$$\begin{aligned}\mathcal{C}^{\text{semigroup}}(G \wr R_n) &\leq \mathcal{C}^{\text{groupoid}}(G \wr R_n) + \mathcal{C}(\zeta_{G \wr R_n}) \\ &= O(|G \wr R_n| \log^4 |G \wr R_n|) + O(|G \wr R_n| \log^3 |G \wr R_n|) \\ &= O(|G \wr R_n| \log^4 |G \wr R_n|).\end{aligned}$$

□

Chapter 9

Another FFT for the Rook Monoid

In this chapter, we generalize Clausen's approach to computing the Fourier transform on the symmetric group (explained in Section 4.3.2) to create an FFT for the rook monoid. The resulting algorithm is an efficient change of basis from the semigroup basis of the rook monoid algebra $\mathbb{C}R_n$ to a Fourier basis. While the algorithm presented here is not as efficient as the one presented in Chapter 7, it is interesting in that it demonstrates that the notion of coset-based factorizations, which are so important in the development of group FFTs, can be generalized to create inverse semigroup FFTs. We do not consider the additional change of basis necessary to run this algorithm on an element of $\mathbb{C}R_n$ expressed with respect to the groupoid basis.

9.1 Description of the Algorithm

Let $f \in \mathbb{C}R_n$ be an arbitrary element, given with respect to the semigroup basis:

$$f = \sum_{s \in R_n} f(s)s.$$

As with the symmetric group (Section 4.3.2), let t_j be the transposition $(j-1, j)$. We use the following sets of "coset representatives." Note that we use quotation marks here because, while these elements play the same role in this FFT as coset representatives did in the symmetric group FFT, these elements do not naturally partition R_n into equally sized sets:

$$\{T_i : 1 \leq i \leq n, T_i = t_{i+1}t_{i+2}\cdots t_n\} \text{ (where } T_n = \text{Id}), \text{ and}$$

$$\{T^i : 1 \leq i \leq n-1, T^i = t_n t_{n-1}\cdots t_{i+1}\}.$$

We use the semigroup chain $R_n > R_{n-1} > \cdots > R_1$, where

$$R_k = \{\sigma \in R_n : \sigma(j) = j \text{ if } j > k\}.$$

Halverson has found a complete set of inequivalent, irreducible matrix representations for R_n adapted to this chain [16]. We recall this description in Section 9.3. Call this set $\mathcal{Y}_n$. For each $\rho \in \mathcal{Y}_n$, we must compute

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

A subsemigroup does not necessarily partition its parent semigroup into equally sized cosets, so we cannot directly factor through a subsemigroup as in (4.4). Instead, we use an approach for R_n that is based on the recursive formula given

in Theorem 2.4.4. With this, we have the following factorization theorem.

Theorem 9.1.1 (Factorization theorem for R_n)**.** *For any representation ρ of R_n, if $n \geq 3$, we have the following factorization.*

$$\begin{aligned}\hat{f}(\rho) &= \sum_{i=1}^{n} \rho(T_i) \sum_{s \in R_{n-1}} f_{T_i}(s)\rho(s) + \rho([n]) \sum_{s \in R_{n-1}} f_{[n]}(s)\rho(s) \\ &+ \sum_{i=1}^{n-1} \left[\sum_{s \in R_{n-1}} f^{T^i}(s)\rho(s)\right] \rho(T^i),\end{aligned} \tag{9.1}$$

where $[n]$ is the link $(1)(2)\cdots(n-1)[n]$, $f_A(s) = f(As)$, and

$$f^{T^i}(s) = \begin{cases} 0 & \text{if } n-1 \in \mathrm{dom}(s), \\ f(sT^i) & \text{otherwise.} \end{cases}$$

Before we prove this theorem, we explain how it leads to an algorithm for the Fourier transform. As with the symmetric group, we use the above breakdown to compute $\hat{f}(\rho)$ recursively. If we knew $\widehat{f_{[n]}}(\gamma)$ for all $\gamma \in \mathcal{Y}_{n-1}$, $\widehat{f_{T_i}}(\gamma)$ for all $\gamma \in \mathcal{Y}_{n-1}$ and $1 \leq i \leq n$, and $\widehat{f^{T^i}}(\gamma)$ for all $\gamma \in \mathcal{Y}_{n-1}$ and $1 \leq i \leq n-1$, we could assemble them based on how ρ splits when restricted to R_{n-1} (see Theorem 9.3.3) to compute the inner sums in (9.1), and this assembly can be done for free since we are using chain-adapted representations. To finish calculating $\hat{f}(\rho)$ for any $\rho \in \mathcal{Y}_n$, we take these inner sums, multiply them by the $\rho(T_i)$, $\rho([n])$, and the $\rho(T^i)$, and add up the results. Therefore, we have:

Lemma 9.1.2. *For $n \geq 3$,*

$$\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_n}(R_n) \leq 2n\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_{n-1}}(R_{n-1}) + M_{R_n},$$

where M_{R_n} is the total number of operations required to compute the sum (9.1) for all $\rho \in \mathcal{Y}_n$, given knowledge of the $\widehat{f_{[n]}}$, all the $\widehat{f_{T_i}}$, and all the $\widehat{f^{T^i}}$ on R_{n-1}.

Before proving Theorem 9.1.1, we prove Theorem 2.4.4, as the ideas in the proof will be used to prove Theorem 9.1.1.

Theorem (Theorem 2.4.4)**.** *For $n \geq 3$,*

$$|R_n| = 2n|R_{n-1}| - (n-1)^2|R_{n-2}|.$$

Proof. Viewing the elements of R_n as rook matrices, R_n consists of those elements having all 0's in column n and row n (of which there are $|R_{n-1}|$), together with, for each $\alpha \in \{1, \ldots, n\}$, those having a 1 in position (α, n) (of which there are $n|R_{n-1}|$ total), together with, for each $\alpha \in \{1, \ldots, n-1\}$, those having a 1 in position (n, α) (of which there are $(n-1)|R_{n-1}|$ total). Counting the number of elements of R_n in this way overcounts. For each pair α, β with $1 \leq \alpha, \beta \leq n-1$, every element with 1's in positions (α, n) and (n, β) (of which there are $(n-1)^2|R_{n-2}|$ total) gets counted twice. □

We now prove Theorem 9.1.1.

Proof of Theorem 9.1.1. Let $n \geq 3$. We have 3 types of elements s of R_n.

- Type 1: Those for which $s(n) = i$ for some $1 \leq i \leq n$.
- Type 2: Those for which both $s(i) = n$ for some $1 \leq i \leq n-1$ and $n \notin \mathrm{dom}(s)$.
- Type 3: Those for which both $s(i) \neq n$ for all $1 \leq i \leq n$ and $n \notin \mathrm{dom}(s)$.

By the argument given in the above proof of Theorem 2.4.4, this counts all elements of R_n precisely once.

Now, let $1 \leq i \leq n$. View the "coset representative" T_i as a permutation matrix, and view the elements $s \in R_n$ as rook matrices. Multiplying any matrix

X on the left by T_i simply moves row j of X to row $j+1$ (for all j such that $i \leq j \leq n-1$) and moves row n of X to row i. Thus, as s varies over R_{n-1}, $T_i s$ varies bijectively over $\{s \in R_n : s(n) = i\}$. Therefore, we have

$$\begin{aligned}\sum_{s \in R_n \text{ of Type 1}} f(s)\rho(s) &= \sum_{i=1}^{n} \sum_{s \in R_{n-1}} f(T_i s)\rho(T_i s) \\ &= \sum_{i=1}^{n} \sum_{s \in R_{n-1}} f_{T_i}(s)\rho(T_i)\rho(s) \\ &= \sum_{i=1}^{n} \rho(T_i) \sum_{s \in R_{n-1}} f_{T_i}(s)\rho(s),\end{aligned}$$

where $f_{T_i}(s) = f(T_i s)$.

Similarly, multiplying any matrix X on the right by T^i moves column j of X to column $j+1$ ($i \leq j \leq n-1$) and moves column n of X to column i. Thus, as s varies over R_{n-1}, sT^i varies bijectively over $\{s \in R_n : s(i) = n\}$. So

$$\begin{aligned}\sum_{s \in R_n : s(i) = n} f(s)\rho(s) &= \sum_{s \in R_{n-1}} f(sT^i)\rho(sT^i) \\ &= \sum_{s \in R_{n-1}} f(sT^i)\rho(s)\rho(T^i).\end{aligned}$$

To ensure that we only count the elements of Type 2, we restrict our attention to $1 \leq i \leq n-1$, and we set the function values of the elements of Type 1 appearing in the above sum to 0:

$$\begin{aligned}\sum_{s \in R_n \text{ of Type 2}} f(s)\rho(s) &= \sum_{i=1}^{n-1} \sum_{s \in R_{n-1}} f^{T^i}(s)\rho(sT^i) \\ &= \sum_{i=1}^{n-1} \left[\sum_{s \in R_{n-1}} f^{T^i}(s)\rho(s) \right] \rho(T^i),\end{aligned}$$

where

$$f^{T^i}(s) = \begin{cases} 0 & \text{if } n-1 \in \mathrm{dom}(s) \text{ (i.e., } n \in \mathrm{dom}(sT^i)), \\ f(sT^i) & \text{otherwise.} \end{cases}$$

Finally,

$$\begin{aligned} \sum_{s \in R_n \text{ of Type 3}} f(s)\rho(s) &= \sum_{s \in R_{n-1}} f([n]s)\rho([n]s) \\ &= \rho([n]) \sum_{s \in R_{n-1}} f_{[n]}(s)\rho(s). \end{aligned}$$

Putting this all together, then, we find that for any representation ρ of R_n and any $n \geq 3$,

$$\begin{aligned} \hat{f}(\rho) &= \sum_{s \in R_n \text{ of Type 1}} f(s)\rho(s) + \sum_{s \in R_n \text{ of Type 2}} f(s)\rho(s) + \sum_{s \in R_n \text{ of Type 3}} f(s)\rho(s) \\ &= \sum_{i=1}^{n} \rho(T_i) \sum_{s \in R_{n-1}} f_{T_i}(s)\rho(s) + \rho([n]) \sum_{s \in R_{n-1}} f_{[n]}(s)\rho(s) \\ &+ \sum_{i=1}^{n-1} \left[\sum_{s \in R_{n-1}} f^{T^i}(s)\rho(s) \right] \rho(T^i). \end{aligned}$$

□

9.2 Analysis of the Algorithm

We now analyze the number of operations necessary to run the algorithm presented in Section 9.1. We begin by analyzing the M_{R_n} term in Lemma 9.1.2 to obtain:

Theorem 9.2.1. For $n \geq 3$,

$$\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_n}(R_n) \leq 2n\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_{n-1}}(R_{n-1}) + 2n^2|R_n|. \tag{9.2}$$

Proof. To analyze M_{R_n}, let:

- M_1 = The maximum number of operations necessary to calculate the matrix product $\rho(T_i)A_{T_i}(\rho)$ for arbitrary matrices $A_{T_i}(\rho)$, for all $\rho \in \mathcal{Y}_n$ and for all T_i $(1 \leq i \leq n)$.

- M_2 = The maximum number of operations necessary to calculate the matrix product $A^{T^i}(\rho)\rho(T^i)$ for arbitrary matrices $A^{T^i}(\rho)$, for all $\rho \in \mathcal{Y}_n$ and for all T^i $(1 \leq i \leq n-1)$.

- M_3 = The maximum number of operations necessary to calculate the matrix product $\rho([n])A_{[n]}(\rho)$ for arbitrary matrices $A_{[n]}(\rho)$, for all $\rho \in \mathcal{Y}_n$.

- M_4 = The maximum number of operations necessary to add together $2n$ $d_\rho \times d_\rho$ arbitrary matrices, for all $\rho \in \mathcal{Y}_n$.

Then $M_{R_n} \leq \sum_{i=1}^{4} M_i$.

Analysis of M_1: For each $\rho \in \mathcal{Y}_n$ and each T_i, we must perform the multiplication $\rho(T_i)A_{T_i}(\rho) = \rho(t_{i+1})\rho(t_{i+2})\cdots\rho(t_n)A_{T_i}(\rho)$ for an arbitrary matrix $A_{T_i}(\rho)$. As was the case with S_n, for $j > 2$, $t_j \in R_j$, t_j commutes with R_{j-2}, and $\mathcal{M}(R_j, R_{j-2}) = 2$. By Schur's Lemma, $\rho(t_j)$ $(j > 2)$ contains at most 2 non-zero entries per row and column. For $j = 2$, $t_2 \in R_2$, and the maximum dimension of an irreducible representation of R_2 is 2. Therefore, $\rho(t_2)$ contains at most 2 non-zero entries per row and column. Alternatively, it is obvious from the description of $\mathcal{Y}_n$ (see Section 9.3) that $\rho(t_j)$ $(j \geq 2)$ contains at most 2 nonzero entries per row and column. Therefore, multiplying an arbitrary matrix by $\rho(t_j)$ on the left requires at most $2d_\rho^2$ operations, and so performing the multiplication $\rho(T_i)A_{T_i}(\rho)$ requires at most

$2(n-i)d_\rho^2$ operations. Therefore, we have

$$M_1 \leq \sum_{\rho \in \mathcal{Y}_n} \sum_{i=1}^{n} 2(n-i)d_\rho^2 = (n)(n-1)|R_n|,$$

where the final equality comes from (2.2).

Analysis of M_2: The only difference between M_1 and M_2 is that M_2 involves multiplying arbitrary matrices by $\rho(T^i)$ on the right rather than by $\rho(T_i)$ on the left, so M_2 is the same as M_1 in the complexity analysis. Thus

$$M_2 \leq (n)(n-1)|R_n|.$$

Analysis of M_3: Since $[n] \in R_n$, $[n]$ commutes with R_{n-1}, and $\mathcal{M}(R_n, R_{n-1}) = 1$, we have that $\rho([n])$ contains at most 1 non-zero entry per row. Thus

$$M_3 \leq \sum_{\rho \in \mathcal{Y}_n} d_\rho^2 = |R_n|.$$

Analysis of M_4: For a particular ρ, the matrix additions can be accomplished with $(2n-1)d_\rho^2$ operations. Thus

$$M_4 \leq \sum_{\rho \in \mathcal{Y}_n} (2n-1)d_\rho^2 = (2n-1)|R_n|.$$

Putting this all together, we obtain

$$M_{R_n} \leq (2(n)(n-1) + 1 + 2n - 1)|R_n| = 2n^2|R_n|.$$

□

We now prove that this algorithm gives the complexity result:

Theorem 9.2.2. *For $n \geq 5$, $\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_n}(R_n) \leq 2^n n|R_n|$.*

Proof. We proceed by induction. For the base case, we note that $|R_2| = 7$, so a naive implementation of the FFT on R_2 gives $\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_2}(R_2) \leq 49$. Applying (9.2) repeatedly, we have

$$\begin{aligned}
\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_3}(R_3) &\leq 2(3)\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_2}(R_2) + 2(3)^2|R_3| \leq 6(49) + 18(34) = 906, \text{ so} \\
\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_4}(R_4) &\leq 2(4)\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_3}(R_3) + 2(4)^2|R_4| \leq 8(906) + 32(209) = 13936, \text{ so} \\
\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_5}(R_5) &\leq 2(5)\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_4}(R_4) + 2(5)^2|R_5| \leq 10(13936) + 50(1546) \\
&= 216660, \text{ and } 216660 < 2^5(5)|R_5| = 247360.
\end{aligned}$$

This proves the base case.

Similarly, for $n = 6$, we have

$$\begin{aligned}
\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_6}(R_6) &\leq 2(6)\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_5}(R_5) + 2(6)^2|R_6| \leq 12(216660) + 72(13327) \\
&= 3559464, \text{ and } 3559464 < 2^6(6)|R_6| = 5117568.
\end{aligned}$$

Now, let $n \geq 7$. Observe that, for $\alpha = 1$ to n, the sets $\{\sigma \in R_n : \sigma(\alpha) = n\}$ are disjoint, and each are of size $|R_{n-1}|$. Thus $n|R_{n-1}| \leq |R_n|$. Therefore, we have:

$$\begin{aligned}
\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_n}(R_n) &\leq 2n\mathcal{T}^{\text{semigroup}}_{\mathcal{Y}_{n-1}}(R_{n-1}) + 2n^2|R_n| \\
&\leq 2n(2^{n-1}(n-1)|R_{n-1}|) + 2n^2|R_n| \\
&\leq 2^n(n-1)|R_n| + 2n^2|R_n| \\
&= 2^n n|R_n| + (2n^2 - 2^n)|R_n| \\
&\leq 2^n n|R_n|,
\end{aligned}$$

where the last inequality arises from the fact that $2n^2 \leq 2^n$ for $n \geq 7$. $\square$

We end this section by proving that this algorithm compares favorably to the naive algorithm.

Theorem 9.2.3. *Let $\epsilon > 0$. Then, based on the algorithm presented in this chapter, we have $\mathcal{T}_{\mathcal{Y}_n}^{\text{semigroup}}(R_n) = O(|R_n|^{1+\epsilon})$.*

Proof. If $n \geq 5$, then the algorithm presented in this chapter requires no more than $2^n n|R_n|$ operations to run. Since $|R_n| \geq n!$ and $n2^n = O(n!^\epsilon)$ for any $\epsilon > 0$, we have $n2^n = O(|R_n|^\epsilon)$. The theorem follows. $\square$

9.3 Seminormal Representations of the Rook Monoid

In this section, we give a description of a complete set of irreducible, inequivalent, chain-adapted matrix representations for the rook monoid R_n relative to the chain $R_n > R_{n-1} > \cdots > R_1$, and we describe the Branching Theorem for R_n. These results are needed for the FFT for R_n described in Sections 9.1 and 9.2. The results in this section are a special case of the results in [16] (with the possible exception of the generalized last-letter ordering described here, which was at least implicit in [16]).

As in Section 4.3.1, we define a *partition* λ of a nonnegative integer k (written $\lambda \vdash k$) to be a weakly decreasing sequence of nonnegative integers whose sum is k. We consider two partitions to be equal if and only if they only differ by the number of 0's they contain, and we identify a partition λ with its Young diagram.

It is well-known that a complete set of inequivalent, irreducible representations for R_n is indexed by the set of all partitions of the integers $\{0, 1, \ldots n\}$ (see, for

example, [15] or [36]). In fact, this also follows directly from the discussion in Section 5.2. Therefore, for integers $n \geq 0$, let

$$\Lambda_n = \{\lambda \vdash k : 0 \leq k \leq n\}.$$

Definition (n-tableau, n-standard tableau)**.** For $\lambda \in \Lambda_n$, define L to be an *n-tableau of shape* λ if it is a filling of the diagram for λ with numbers from $\{1, 2, \ldots n\}$ such that each number in L appears at most once. L is an *n-standard tableau* if, furthermore, the entries in each column of L increase from top to bottom and the entries in each row of L increase from left to right.

Fix λ. Let T_n^λ denote the set of n-standard tableaux of shape λ. The symmetric group acts on tableaux by permuting their entries. If L is an n-tableau, then $(i-1, i)L$ is the tableau that is obtained from L by replacing $i-1$ (if $i-1 \in L$) with i, and by replacing i (if $i \in L$) with $i-1$. Note that $L \in T_n^\lambda$ need not imply $(i-1, i)L \in T_n^\lambda$.

Let $\{v_L : L \in T_n^\lambda\}$ be a set of independent vectors. We form

$$V^\lambda = \mathbb{C}\text{-span}\{v_L : L \in T_n^\lambda\}.$$

As such, the symbols v_L, for $L \in T_n^\lambda$, are a basis for the vector space V^λ. Halverson defines an action of R_n on V^λ in such a way that (extending by linearity) V^λ is an irreducible $\mathbb{C}R_n$-module and such that, as λ ranges over Λ_n, the V^λ constitute a complete set of inequivalent, irreducible representation modules for R_n. We first describe this action, and we then describe an ordering of the bases for the V^λ so that the resulting matrix representations are chain-adapted to $R_n > R_{n-1} > \cdots > R_1$.

Definition (content)**.** If b is a box of λ in position (i, j), then the *content* of b is defined to be

$$ct(b) = j - i.$$

Let $L \in T_n^\lambda$. If $i-1, i \in L$, then let $L(i-1)$ and $L(i)$ denote the box in L containing $i-1$ and i, respectively.

To define the action of R_n on V^λ, it is sufficient to define the action of a set of generators of R_n on V^λ.

Definition (action of R_n on V^λ)**.** Define the action of the transpositions $t_i = (i-1, i)$, for $2 \leq i \leq n$, as follows:

$$t_i v_L = \begin{cases} \frac{1}{ct(L(i))-ct(L(i-1))} v_L + (1 + \frac{1}{ct(L(i))-ct(L(i-1))}) v_{L'} & \text{if } i-1, i \in L, \\ v_{t_i L} & \text{if exactly one of } i-1, i \in L, \\ v_L & \text{if } i-1, i \notin L, \end{cases}$$

where

$$v_{L'} = \begin{cases} v_{t_i L} & \text{if } t_i L \text{ is } n\text{-standard}, \\ 0 & \text{otherwise.} \end{cases}$$

Define the action of the link $(1)(2)\cdots(n-1)[n] = [n]$ on V^λ by

$$[n] v_L = \begin{cases} v_L & \text{if } n \notin L, \\ 0 & \text{if } n \in L. \end{cases}$$

Remark: If $\lambda = (0)$, then V^λ is 1-dimensional, and the action of R_n on V^λ is the trivial action given by $xv = v$ for all $x \in R_n$ and all $v \in V^\lambda$.

Theorem 9.3.1 (Halverson, [16]). *As λ varies over all partitions of all nonnegative integers less than or equal to n, the V^λ constitute a complete set of irreducible, pairwise non-isomorphic representation modules for R_n.*

Definition (corner of a partition). A corner is a box c of λ for which λ contains no box to the right or below c. In other words, the corners are the possible positions of n in an n-standard tableau of shape λ.

We now record the Branching Theorem for R_n.

Theorem 9.3.2 (Branching Theorem, Halverson [16]). *As a $\mathbb{C}R_{n-1}$ module,*

$$V^\lambda \cong \oplus_{\mu \in \lambda^{-,=}} V^\mu,$$

where $\lambda^{-,=}$ is the set of all partitions $\mu \in \Lambda_{n-1}$ such that either $\mu = \lambda$ (if $\lambda \nvdash n$) or μ is obtained by removing a corner from λ.

Now, for purposes of chain-adaptation, we order the basis $\{v_L\}$ for V^λ using the following generalized last-letter ordering. We begin by partitioning the v_L into subsets based on the corners $c_1, \ldots, c_l$ of λ. Number the corners from top to bottom. Now, form the sets

$$V^\lambda(0) = \{v_L : L \in T_n^\lambda \text{ and } n \notin L\},$$

$$V^\lambda(i) = \{v_L : L \in T_n^\lambda \text{ and } n \in c_i \text{ of } L\}, \quad 1 \leq i \leq l,$$

and declare elements of $V^\lambda(j)$ to be earlier in the ordering than elements of $V^\lambda(k)$ whenever $j < k$. To order the subset $V^\lambda(k)$, delete the corner c_k (do nothing if $k = 0$) and repeat the same ordering process (starting by identifying the corners of

the resulting partition and partitioning the v_L into subsets based on those corners) with $n-1$ in place of n, etc.

As an example, consider $\lambda = (2,1,1)$, which has two corners, and R_5. Our ordered basis for the 15-dimensional V^λ is

$$v_{\begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1}3\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}4\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}4\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&5\\\hline 2\\\cline{1-1}3\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&5\\\hline 2\\\cline{1-1}4\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&5\\\hline 3\\\cline{1-1}4\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 2&5\\\hline 3\\\cline{1-1}4\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&3\\\hline 2\\\cline{1-1}5\\\cline{1-1}\end{array}}$$

$$< v_{\begin{array}{|c|c|}\hline 1&2\\\hline 3\\\cline{1-1}5\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&4\\\hline 2\\\cline{1-1}5\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&4\\\hline 3\\\cline{1-1}5\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 2&4\\\hline 3\\\cline{1-1}5\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&2\\\hline 4\\\cline{1-1}5\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 1&3\\\hline 4\\\cline{1-1}5\\\cline{1-1}\end{array}} < v_{\begin{array}{|c|c|}\hline 2&3\\\hline 4\\\cline{1-1}5\\\cline{1-1}\end{array}}.$$

Remark: If $\lambda \vdash n$, then the generalized last-letter ordering scheme given above reduces to the usual last-letter ordering scheme used for Young's orthogonal and seminormal representations of the symmetric group.

It is now easy to see, under this ordering of the bases for the V^λ, that the matrix representations described in this section are chain-adapted to the chain $R_n > R_{n-1} > \cdots > R_1$. We know, by the Branching Theorem for R_n, how V^λ decomposes as a R_{n-1} module (it decomposes into the $\mathbb{C}$-span of the $V^\lambda(k)$ for $0 \leq k \leq l$), and it is obvious by the action of R_{n-1} on V^λ that

$$R_{n-1}\mathbb{C}\text{-span}(V^\lambda(k)) \subseteq \mathbb{C}\text{-span}(V^\lambda(k))$$

for all k. The same argument is used to perform induction down the chain $R_n > R_{n-1} > \cdots > R_1$, now, and is trivial because we ordered our basis for V^λ inductively according to the same rule. We now restate the Branching Theorem for R_n under our ordering of the bases for the V^λ.

Theorem 9.3.3 (Branching Theorem)**.** *Let ρ^λ be the matrix representation asso-*

ciated to V^λ with respect to the basis $\{v_L\}$, with the basis ordered according to the generalized last-letter ordering. Then

$$\rho^\lambda|_{R_{n-1}} = \oplus_{\mu \in \lambda^{-,=}} \rho^\mu,$$

where $\lambda^{-,=}$ is the set of all partitions $\mu \in \Lambda_{n-1}$ such that either $\mu = \lambda$ or μ is obtained by removing a corner from λ. The first $\mu \in \lambda^{-,=}$ is the one that removes no corners (if $\lambda \not\vdash n$), the next μ is the one that removes the highest corner, the next μ is the one that removes the second highest corner, etc.

Chapter 10

An Application to Partially Ranked Data

In [11], P. Diaconis gives a method for analyzing partially ranked voting data with Fourier transforms on the symmetric group. In this chapter, we explore the natural generalization of his method to the rook monoid R_n. We will see how the Fourier transform on R_n (i.e., spectral analysis for R_n) can be used to analyze a collection of partially ranked data and how our approach differs from his.

The analysis in [11] considers the results of the 1980 American Psychological Association election, in which voters were asked to rank five candidates in order of preference. 15449 people voted, of which 5738 fully ranked all five candidates. The votes are tallied in Tables 10.1 through 10.4.

Table 10.1: Fully ranked ballots

Vote	Tally	Vote	Tally	Vote	Tally	Vote	Tally
[5, 4, 3, 2, 1]	29	[4, 3, 5, 2, 1]	91	[3, 2, 5, 4, 1]	41	[2, 1, 5, 4, 3]	36
[5, 4, 3, 1, 2]	67	[4, 3, 5, 1, 2]	84	[3, 2, 5, 1, 4]	64	[2, 1, 5, 3, 4]	42
[5, 4, 2, 3, 1]	37	[4, 3, 2, 5, 1]	30	[3, 2, 4, 5, 1]	34	[2, 1, 4, 5, 3]	24
[5, 4, 2, 1, 3]	24	[4, 3, 2, 1, 5]	35	[3, 2, 4, 1, 5]	75	[2, 1, 4, 3, 5]	26
[5, 4, 1, 3, 2]	43	[4, 3, 1, 5, 2]	38	[3, 2, 1, 5, 4]	82	[2, 1, 3, 5, 4]	30
[5, 4, 1, 2, 3]	28	[4, 3, 1, 2, 5]	35	[3, 2, 1, 4, 5]	74	[2, 1, 3, 4, 5]	40
[5, 3, 4, 2, 1]	57	[4, 2, 5, 3, 1]	58	[3, 1, 5, 4, 2]	30	[1, 5, 4, 3, 2]	40
[5, 3, 4, 1, 2]	49	[4, 2, 5, 1, 3]	66	[3, 1, 5, 2, 4]	34	[1, 5, 4, 2, 3]	35
[5, 3, 2, 4, 1]	22	[4, 2, 3, 5, 1]	24	[3, 1, 4, 5, 2]	40	[1, 5, 3, 4, 2]	36
[5, 3, 2, 1, 4]	22	[4, 2, 3, 1, 5]	51	[3, 1, 4, 2, 5]	42	[1, 5, 3, 2, 4]	17
[5, 3, 1, 4, 2]	34	[4, 2, 1, 5, 3]	52	[3, 1, 2, 5, 4]	30	[1, 5, 2, 4, 3]	70
[5, 3, 1, 2, 4]	26	[4, 2, 1, 3, 5]	40	[3, 1, 2, 4, 5]	34	[1, 5, 2, 3, 4]	50
[5, 2, 4, 3, 1]	54	[4, 1, 5, 3, 2]	50	[2, 5, 4, 3, 1]	35	[1, 4, 5, 3, 2]	52
[5, 2, 4, 1, 3]	44	[4, 1, 5, 2, 3]	45	[2, 5, 4, 1, 3]	34	[1, 4, 5, 2, 3]	48
[5, 2, 3, 4, 1]	26	[4, 1, 3, 5, 2]	31	[2, 5, 3, 4, 1]	40	[1, 4, 3, 5, 2]	51
[5, 2, 3, 1, 4]	24	[4, 1, 3, 2, 5]	23	[2, 5, 3, 1, 4]	21	[1, 4, 3, 2, 5]	24
[5, 2, 1, 4, 3]	35	[4, 1, 2, 5, 3]	22	[2, 5, 1, 4, 3]	106	[1, 4, 2, 5, 3]	70
[5, 2, 1, 3, 4]	50	[4, 1, 2, 3, 5]	16	[2, 5, 1, 3, 4]	79	[1, 4, 2, 3, 5]	45
[5, 1, 4, 3, 2]	50	[3, 5, 4, 2, 1]	71	[2, 4, 5, 3, 1]	63	[1, 3, 5, 4, 2]	35
[5, 1, 4, 2, 3]	46	[3, 5, 4, 1, 2]	61	[2, 4, 5, 1, 3]	53	[1, 3, 5, 2, 4]	28
[5, 1, 3, 4, 2]	25	[3, 5, 2, 4, 1]	41	[2, 4, 3, 5, 1]	44	[1, 3, 4, 5, 2]	37
[5, 1, 3, 2, 4]	19	[3, 5, 2, 1, 4]	27	[2, 4, 3, 1, 5]	28	[1, 3, 4, 2, 5]	35
[5, 1, 2, 4, 3]	11	[3, 5, 1, 4, 2]	45	[2, 4, 1, 5, 3]	162	[1, 3, 2, 5, 4]	95
[5, 1, 2, 3, 4]	29	[3, 5, 1, 2, 4]	36	[2, 4, 1, 3, 5]	96	[1, 3, 2, 4, 5]	102
[4, 5, 3, 2, 1]	31	[3, 4, 5, 2, 1]	107	[2, 3, 5, 4, 1]	45	[1, 2, 5, 4, 3]	34
[4, 5, 3, 1, 2]	54	[3, 4, 5, 1, 2]	133	[2, 3, 5, 1, 4]	52	[1, 2, 5, 3, 4]	35
[4, 5, 2, 3, 1]	34	[3, 4, 2, 5, 1]	62	[2, 3, 4, 5, 1]	53	[1, 2, 4, 5, 3]	29
[4, 5, 2, 1, 3]	24	[3, 4, 2, 1, 5]	28	[2, 3, 4, 1, 5]	52	[1, 2, 4, 3, 5]	27
[4, 5, 1, 3, 2]	38	[3, 4, 1, 5, 2]	87	[2, 3, 1, 5, 4]	186	[1, 2, 3, 5, 4]	28
[4, 5, 1, 2, 3]	30	[3, 4, 1, 2, 5]	35	[2, 3, 1, 4, 5]	172	[1, 2, 3, 4, 5]	30

Table 10.2: Rank-3 ballots

Vote	Tally	Vote	Tally	Vote	Tally	Vote	Tally
[1, 2, 3, 0, 0]	27	[3, 1, 0, 0, 2]	38	[1, 0, 0, 2, 3]	44	[0, 3, 1, 0, 2]	16
[1, 3, 2, 0, 0]	79	[2, 3, 0, 0, 1]	35	[1, 0, 0, 3, 2]	35	[0, 2, 3, 0, 1]	14
[2, 1, 3, 0, 0]	31	[3, 2, 0, 0, 1]	41	[2, 0, 0, 1, 3]	46	[0, 3, 2, 0, 1]	15
[3, 1, 2, 0, 0]	32	[1, 0, 2, 3, 0]	30	[2, 0, 0, 3, 1]	62	[0, 1, 0, 2, 3]	55
[2, 3, 1, 0, 0]	83	[1, 0, 3, 2, 0]	21	[3, 0, 0, 1, 2]	90	[0, 1, 0, 3, 2]	45
[3, 2, 1, 0, 0]	57	[2, 0, 1, 3, 0]	39	[3, 0, 0, 2, 1]	75	[0, 2, 0, 1, 3]	54
[1, 2, 0, 3, 0]	19	[3, 0, 1, 2, 0]	15	[0, 1, 2, 3, 0]	9	[0, 3, 0, 1, 2]	62
[1, 3, 0, 2, 0]	22	[2, 0, 3, 1, 0]	15	[0, 1, 3, 2, 0]	17	[0, 2, 0, 3, 1]	50
[2, 1, 0, 3, 0]	31	[3, 0, 2, 1, 0]	13	[0, 3, 1, 2, 0]	26	[0, 3, 0, 2, 1]	59
[3, 1, 0, 2, 0]	45	[1, 0, 3, 0, 2]	41	[0, 2, 1, 3, 0]	17	[0, 0, 1, 2, 3]	15
[2, 3, 0, 1, 0]	28	[1, 0, 2, 0, 3]	49	[0, 2, 3, 1, 0]	21	[0, 0, 1, 3, 2]	19
[3, 2, 0, 1, 0]	51	[2, 0, 1, 0, 3]	74	[0, 3, 2, 1, 0]	18	[0, 0, 2, 1, 3]	16
[1, 2, 0, 0, 3]	26	[3, 0, 1, 0, 2]	47	[0, 1, 2, 0, 3]	8	[0, 0, 3, 1, 2]	46
[1, 3, 0, 0, 2]	31	[2, 0, 3, 0, 1]	37	[0, 1, 3, 0, 2]	15	[0, 0, 2, 3, 1]	17
[2, 1, 0, 0, 3]	17	[3, 0, 2, 0, 1]	32	[0, 2, 1, 0, 3]	16	[0, 0, 3, 2, 1]	20

Table 10.3: Rank-2 ballots

Vote	Tally	Vote	Tally
[1, 2, 0, 0, 0]	83	[0, 1, 0, 2, 0]	87
[2, 1, 0, 0, 0]	72	[0, 2, 0, 1, 0]	114
[1, 0, 2, 0, 0]	302	[0, 1, 0, 0, 2]	80
[2, 0, 1, 0, 0]	547	[0, 2, 0, 0, 1]	70
[1, 0, 0, 2, 0]	74	[0, 0, 1, 2, 0]	56
[2, 0, 0, 1, 0]	104	[0, 0, 2, 1, 0]	48
[1, 0, 0, 0, 2]	72	[0, 0, 1, 0, 2]	93
[2, 0, 0, 0, 1]	117	[0, 0, 2, 0, 1]	64
[0, 1, 2, 0, 0]	51	[0, 0, 0, 1, 2]	196
[0, 2, 1, 0, 0]	89	[0, 0, 0, 2, 1]	143

Table 10.4: Rank-1 ballots

Vote	Tally
[1, 0, 0, 0, 0]	895
[0, 1, 0, 0, 0]	881
[0, 0, 1, 0, 0]	1198
[0, 0, 0, 1, 0]	1145
[0, 0, 0, 0, 1]	1022

Note that this collection of tables defines a $\mathbb{C}$-valued (actually, a $\mathbb{Z}$-valued) function f on R_5, where $f(\sigma)$ is the number of voters casting a ballot of type σ. The σ here are written in standard list-form, with the top row removed. For example, we have

$$f(\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix}) = 172$$

and

$$f(\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & - & - & 2 & 1 \end{pmatrix}) = 75.$$

We begin by reviewing Diaconis's method for analyzing this dataset.

10.1 The Symmetric Group Approach

To illustrate Diaconis's method, let us begin by examining the fully-ranked votes. Full details may be found in [11]. The collection of fully ranked votes defines a $\mathbb{C}$-valued function f on the symmetric group S_5. In other words, it defines a vector in $\mathbb{C}S_5$ given by

$$\sum_{\sigma \in S_5} f(\sigma)\sigma.$$

$\mathbb{C}S_5$ decomposes uniquely into its isotypic subspaces. A complete set of inequivalent, irreducible representations of S_5 is indexed by the partitions of 5 (see Section 4.3.1), so let us write

$$\mathbb{C}S_5 = V^{(5)} \oplus V^{(4,1)} \oplus V^{(3,2)} \oplus V^{(3,1,1)} \oplus V^{(2,2,1)} \oplus V^{(2,1,1,1)} \oplus V^{(1,1,1,1,1)},$$

where V^λ is the isotypic subspace corresponding to the irreducible representation for λ. There is a natural inner product on $\mathbb{C}S_n$ given by

$$\langle f, g \rangle = \left\langle \sum_{\sigma \in S_n} f(\sigma)\sigma, \sum_{\sigma \in S_n} g(\sigma)\sigma \right\rangle = \sum_{\sigma \in S_n} f(\sigma)\overline{g(\sigma)}.$$

Under this inner product, the isotypic subspaces of $\mathbb{C}S_n$ are mutually orthogonal (see Chapter 2 of [35]). Furthermore, each isotypic subspace contains certain natural statistical information which we are about to describe. Let us project $f \in \mathbb{C}S_5$ onto each subspace. That is, let us write

$$f = \sum_{\lambda \vdash 5} f^\lambda,$$

where $f^\lambda \in V^\lambda$. These projections f^λ may be computed by running an FFT on S_5. Next, we examine the projections f^λ. This is the analog of examining the component frequencies of a function in the $\mathbb{Z}/n\mathbb{Z}$ case, and is called *spectral analysis* [11]. However, in our case, many of the V^λ are multidimensional, and we need concrete descriptions of these spaces together with an additional tool (which Diaconis attributes to C. Mallows) to understand the projections onto these spaces.

$V^{(5)}$ is the one-dimensional space of constant functions on the fully ranked votes, so no additional tools are needed to understand $f^{(5)}$. We have

$$f^{(5)} = 5738 \left(\frac{1}{120} \sum_{x \in S_5} x \right).$$

5738 is simply the number of rank-5 votes cast.

$V^{(4,1)}$ is a sixteen-dimensional space. There are twenty-five "easily inter-

pretable" first-order functions. They are of the form

$$\delta_{i\mapsto j} = \sum_{x\in S_5} \delta_{i\mapsto j}(x)x,$$

where

$$\delta_{i\mapsto j}(x) = \begin{cases} 1 & \text{if } x(i) = j, \\ 0 & \text{otherwise.} \end{cases}$$

A general element of $V^{(4,1)}$ has the form

$$\sum_{i,j} a_{i,j}\delta_{i\mapsto j},$$

where, since $V^{(4,1)}$ is orthogonal to $V^{(5)}$, the $a_{i,j}$ sum to 0. $V^{(4,1)}$ therefore carries "pure" first-order statistics for fully ranked votes. Mallows's idea, then, is to project each of the $\delta_{i\mapsto j}$ onto $V^{(4,1)}$, compute inner products of these projections with the original function f, and examine the resulting scalars. For computational purposes, we note that this is the same as, for a given easily interpretable first-order function δ, taking the projection of f onto $V^{(4,1)}$ and taking the inner product of that with δ. The results of this analysis of $f^{(4,1)}$ are summarized in Table 10.5. The i, j entry is $\left\langle f^{(4,1)}, \delta_{i\mapsto j}\right\rangle$, and entries have been rounded to integers.

Table 10.5: First-order analysis, rank-5 data

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-95	371	165	-146	-297
2	-373	-71	267	268	-93
3	461	-188	-355	-98	178
4	24	-176	-59	16	193
5	-19	62	-20	-42	17

These entries can also be categorized, for general n, as

$$\text{Entry}_{i,j} = \sum_{x \in S_n} f(x)w(x),$$

$$w(x) = \begin{cases} \frac{n-1}{n} & \text{if } x(i) = j, \\ -\frac{1}{n} & \text{otherwise.} \end{cases}$$

These entries indicate that, among the rank-5 votes, there is a strong effect for ranking candidate 3 in position 1, candidate 1 in position 2, and candidate 2 in positions 3 or 4. Candidates 3 and 4 also received some "hate vote," as indicated by the 3^{rd} and 4^{th} values in the last column of the table. There wasn't an overwhelming effect for ranking candidate 5 in any particular position.

At this point, it would be useful to pause and ask how much of the "energy" of f we have analyzed. Just as the frequencies of highest amplitude carry the most information about the structure of a continuous waveform, here (and more generally) the projections of longest length carry the most information about the structure of the dataset. By orthogonality of isotypic subspaces, we have

$$||f||^2 = \langle f, f \rangle = \sum_{\lambda \vdash 5} \left\langle f^\lambda, f^\lambda \right\rangle = 378756.$$

The squared lengths of the projections f^λ are summarized in Table 10.6, whose entries have been rounded to integers.

The third column of this table indicates the squared lengths of the projections divided by the dimensions of the subspaces in which they reside. It is common in statistics to divide the squared length of a vector by the dimension of the subspace in which it resides in order to account for the distribution of the vector among a set of basis vectors for the subspace. This is reasonable if the vector might be

Table 10.6: APA election: Rank-5 squared projection lengths

λ	$\langle f^\lambda, f^\lambda \rangle$	$\frac{\langle f^\lambda, f^\lambda \rangle}{\dim V^\lambda}$
(5)	274372	274372
(4, 1)	35790	2237
(3, 2)	55098	2204
(3, 1, 1)	9385	261
(2, 2, 1)	3264	131
(2, 1, 1, 1)	820	51
(1, 1, 1, 1, 1)	26	26

reasonably distributed throughout an orthonormal basis for that subspace, so that the result is a measure of signal to noise. However, if the original data vector was highly structured (as it might well be for voting data!), then the projections are more likely to lie close to a small number of easily interpretable vectors, and dividing by the dimensions of the subspaces might therefore be misleading. In either case, it is evident that the projection $f^{(3,2)}$ warrants our attention. Also, note that

$$||f^{(5)} + f^{(4,1)} + f^{(3,2)}|| > .98||f||,$$

so $f^{(3,2)}$ will be the last projection that we consider.

Just as there are easily interpretable first-order functions, there are also easily interpretable second-order (ordered and unordered) functions. The easily interpretable second-order unordered functions are the

$$\delta_{\{i_1,i_2\}\to\{j_1,j_2\}} = \sum_{x\in S_5} \delta_{\{i_1,i_2\}\to\{j_1,j_2\}}(x)x,$$

where

$$\delta_{\{i_1,i_2\}\to\{j_1,j_2\}}(x) = \begin{cases} 1 & \text{if } \{x(i_1), x(i_2)\} = \{j_1, j_2\}, \\ 0 & \text{otherwise,} \end{cases}$$

and elements of $V^{(3,2)}$ are linear combinations of the $\delta_{\{i_1,i_2\}\to\{j_1,j_2\}}$ which are orthogonal to the other isotypic subspaces. The easily interpretable second-order ordered functions for fully ranked votes are the

$$\delta_{i_1\mapsto j_1 \& i_2\mapsto j_2}$$

(defined analogously), and elements of $V^{(3,1,1)}$ are linear combinations of such. Likewise, there are easily interpretable third-order functions and so on. For a nice intuitive explanation for which isotypic subspaces correspond to which easily interpretable functions, see [11]. We compute the inner products of $f^{(3,2)}$ with the $\delta_{\{i_1,i_2\}\to\{j_1,j_2\}}$ to obtain Table 10.7.

Table 10.7: Second-order unordered analysis, rank-5 data

	Rank									
Candidates	**1,2**	**1,3**	**1,4**	**1,5**	**2,3**	**2,4**	**2,5**	**3,4**	**3,5**	**4,5**
1,2	-137	-20	18	140	111	22	4	6	-97	-46
1,3	476	-88	-179	-209	-147	-169	-160	107	128	241
1,4	-189	51	113	24	-9	98	99	-65	23	-146
1,5	-150	57	47	45	44	49	56	-48	-53	-48
2,3	-42	84	19	-61	30	-16	27	-76	-39	73
2,4	157	-20	-43	-93	-25	-76	-56	8	38	112
2,5	22	-44	7	15	-117	69	25	62	99	-138
3,4	-265	-7	72	119	39	140	85	20	-52	-233
3,5	-169	10	88	70	78	44	47	-52	-36	-81
4,5	296	-24	-142	-130	-5	-163	-128	38	-9	267

With this table, the structure of the data becomes clear. Diaconis explains the historical background for this election as follows—the APA is a professional association that divides primarily into academicians and clinicians who are on uneasy terms with one another. Indeed, its presidential elections are actively contested, and the association nearly split in two after this election. Nearly a

third of the APA membership voted in this election. In this election, candidates 1 and 3 were on one side, candidates 4 and 5 on the other, and candidate 2 was somewhere in the middle, a bit closer to candidates 4 and 5. Voters primarily tended to support one of these sets of candidates (either $\{1,3\}$ or $\{4,5\}$), and then chose between them. We can read this structure from this table as follows.

Recall that the entries in this table have already been adjusted for individual candidate popularity (as $f^{(3,2)}$ is orthogonal to $f^{(4,1)}$). The large positive numbers in the $(\{1,3\},\{1,2\})$, $(\{1,3\},\{4,5\})$, $(\{4,5\},\{1,2\})$, and $(\{4,5\},\{4,5\})$ positions indicate strong effects for ranking candidates 1 and 3 either 1st and 2nd or 4th and 5th, and the same for candidates 4 and 5. Among "opposite" groups such as $\{1,4\}$ and $\{3,5\}$, voters were split, tending not to rank the group in either the first two positions or the last two positions.

Remark: We may generalize this technique to work for partially ranked votes as well. Say we are interested in analyzing the rank-k votes in an election with n candidates. The tally of the rank-k votes defines a function f on the rank-k elements of R_n. For each rank-k element x of R_n, form the following element of $\mathbb{C}S_n$:

$$x' = \sum_{t \in S_n : t \geq x} \frac{1}{E(x)} f(x) t,$$

where $t \geq x$ simply means that t extends x, and $E(x)$ is the number of elements $t \in S_n$ that extend x. Next, form the following element of $\mathbb{C}S_n$:

$$F = \sum_{x \in R_n : \mathrm{rk}(x) = k} x'.$$

Finally, apply the technique described in this section to F. This technique is equivalent to the technique given by Diaconis in [11] for analyzing partially ranked

data. The results of Diaconis's partially ranked data analysis for the APA election are given in Tables 10.8 through 10.11.

Table 10.8: Diaconis's first-order analysis, rank-3 data

	Rank		
Candidate	**1**	**2**	**3**
1	2	76	114
2	-78	-28	52
3	2	-103	-116
4	38	-7	-48
5	35	63	-1

Table 10.9: Diaconis's second-order unordered analysis, rank-3 data

	Rank		
Candidates	**1,2**	**1,3**	**2,3**
1,2	-50	6	12
1,3	150	-3	-41
1,4	-71	-8	11
1,5	-28	5	16
2,3	-2	24	28
2,4	57	-5	-7
2,5	-5	-24	-34
3,4	-84	-12	-4
3,5	-63	-8	17
4,5	97	26	0

Table 10.10: Diaconis's first-order analysis, rank-2 data

	Rank	
Candidate	**1**	**2**
1	38	347
2	-202	-136
3	292	-27
4	-30	-132
5	-98	-51

Table 10.11: Diaconis's second-order unordered analysis, rank-2 data

	Rank
Candidates	**1,2**
1,2	-107
1,3	385
1,4	-142
1,5	-136
2,3	-81
2,4	122
2,5	66
3,4	-176
3,5	-127
4,5	197

We now give a generalization of this method using the rook monoid, and we compare the resulting spectral analysis to Diaconis's approach.

10.2 The Rook Monoid Approach

We illustrate our approach to spectral analysis using the same APA dataset, and we follow Diaconis's lead. That is, we use the vote tallies to define a function f on R_5, we associate f to a vector in the semigroup algebra $\mathbb{C}R_5$, we project f onto the isotypic subspaces of $\mathbb{C}R_5$, we use inner products with "easily interpretable functions" to extract statistical information from these projections, and we examine the resulting coefficients. We have two important considerations in generalizing this approach to the rook monoid:

- We have two natural bases of $\mathbb{C}R_n$, the semigroup basis and the groupoid basis. Declaring a basis association (see Section 3.4) is necessary for performing spectral analysis.

- We require an inner product on $\mathbb{C}R_n$ under which its isotypic subspaces are mutually orthogonal. Of the two "natural" inner products arising from the natural bases of $\mathbb{C}R_n$ (see Section 6.2), only the one arising from the groupoid basis guarantees this.

Therefore, we take our inner product on $\mathbb{C}R_n$ to be the one induced by declaring the groupoid basis of $\mathbb{C}R_n$ mutually orthonormal. We first illustrate our approach using the groupoid basis association.

10.2.1 The Groupoid Basis Association

Choosing the groupoid basis association amounts to passing directly to the groupoid algebra to analyze f. We have $f \in \mathbb{C}R_5$ by

$$f = \sum_{s \in R_5} f(s) \lfloor s \rfloor .$$

First of all, note that this is a special type of dataset. The APA uses the Hare election system, which is also known as proportional voting. A winner is chosen as follows. If a candidate is ranked in first place by more than half of the voters, then he or she is declared the winner. If not, then the candidate with the fewest first-place votes is eliminated, the votes are re-tallied and the candidates re-ranked in the new relative order, and the process repeats. In the Hare system, it only makes sense for voters to rank their top k candidates in order. It would not make sense for a voter to cast a vote such as

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ - & - & 1 & 5 & 4 \end{pmatrix}.$$

However, our technique is perfectly general in that it can handle any $\mathbb{C}$-valued function on R_n.

We begin our analysis by noting how $\mathbb{C}R_5$ decomposes into isotypic subspaces. A complete set of inequivalent, irreducible representations for R_5 is indexed by all

partitions of all integers from 0 to 5 (see Section 9.3), so we write

$$\begin{aligned}\mathbb{C}R_5 =& V^{(5)} \oplus V^{(4,1)} \oplus V^{(3,2)} \oplus V^{(3,1,1)} \oplus V^{(2,2,1)} \oplus V^{(2,1,1,1)} \oplus V^{(1,1,1,1,1)} \oplus \\ & V^{(4)} \oplus V^{(3,1)} \oplus V^{(2,2)} \oplus V^{(2,1,1)} \oplus V^{(1,1,1,1)} \oplus \\ & V^{(3)} \oplus V^{(2,1)} \oplus V^{(1,1,1)} \oplus \\ & V^{(2)} \oplus V^{(1,1)} \oplus \\ & V^{(1)} \oplus \\ & V^{(0)},\end{aligned}$$

where V^λ is the isotypic subspace of $\mathbb{C}R_5$ corresponding to the irreducible representation for λ. We begin our analysis by projecting f onto the isotypic subspaces, that is, by writing

$$f = \sum_{k=0}^{5} \sum_{\lambda \vdash k} f^\lambda$$

for unique elements $f^\lambda \in V^\lambda$. These projections can be computed by running a single FFT on R_n.

Under this basis association, the rank-k data projects onto the V^λ where $\lambda \vdash k$, and we may therefore carry out our analysis rank-by-rank. Our rank-5 (that is, fully ranked) analysis is exactly the same as Diaconis's given in Section 10.1, as the rank-5 portion of $\mathbb{C}R_5$ is exactly $\mathbb{C}S_5$ and the rank-5 portion of f is exactly the function analyzed in Section 10.1. It is in the partially-ranked data that our analysis is different.

The projections f^λ for $\lambda \vdash 4$ are all zero, as there is no rank-4 data. After all, ranking $n-1$ candidates naturally ranks the n^{th} as well. Note, however, that this is only necessarily true for voting data. A more general dataset on R_n could differentiate between a rank-4 element and the corresponding rank-5 element.

Also, note that this is a dataset containing missing data (in the rank-3 and lower-rank votes), and that the data missing here is not missing at random (NMAR). That is, the rank-3 and lower-rank voters intentionally declined to state their preferences about certain candidates. There are a variety of methods for analyzing NMAR data, one of which is to average over the possible extensions of the missing data (which amounts to Diaconis's technique in [11]). Another method is to analyze the subsets of non-missing data separately, which, as we are about to see, amounts to spectral analysis for the rook monoid.

To illustrate, let us now consider the projections f^λ for $\lambda \vdash 3$. Every element $s \in R_n$ has a domain and a range, and $V^{(3)}$ is the sum of the spaces of constant functions for each of the rank-3 choices of domain and range. That is, $V^{(3)}$ is spanned by the elements

$$\delta^{D,R} = \sum_{x \in R_5} \delta^{D,R}(x) \lfloor x \rfloor ,$$

where

$$\delta^{D,R}(x) = \begin{cases} 1 & \text{if } \mathrm{dom}(x) = D \text{ and } \mathrm{ran}(x) = R, \\ 0 & \text{otherwise,} \end{cases}$$

and D and R range across all 3-subsets of $\{1, 2, 3, 4, 5\}$. Following Diaconis's and Mallows's lead, we take the projection $f^{(3)}$ and compute the inner products of it with these $\delta^{D,R}$ to obtain Table 10.12. Notice that the D, R entry is simply the number of rank-3 voters ranking the candidates in D in the positions in R.

More interesting is $f^{(2,1)}$, which in this case contains both the pure first-order and second-order unordered information. To explain, we have the following easily

Table 10.12: Zeroth-order groupoid analysis, rank-3 data

	Range				
Domain	**1,2,3**	**1,2,4**	**1,2,5**	$\cdots$	**3,4,5**
1,2,3	309	0	0	$\cdots$	0
1,2,4	196	0	0	$\cdots$	0
1,2,5	188	0	0	$\cdots$	0
1,3,4	133	0	0	$\cdots$	0
1,3,5	280	0	0	$\cdots$	0
1,4,5	352	0	0	$\cdots$	0
2,3,4	108	0	0	$\cdots$	0
2,3,5	84	0	0	$\cdots$	0
2,4,5	325	0	0	$\cdots$	0
3,4,5	133	0	0	$\cdots$	0

interpretable first-order rank-3 functions

$$\delta_{i\mapsto j}^{D,R} = \sum_{x\in R_5} \delta_{i\mapsto j}^{D,R}(x)\, \lfloor x \rfloor\,,$$

where D, R are 3-subsets of $\{1,2,3,4,5\}$, $i \in D, j \in R$, and

$$\delta_{i\mapsto j}^{D,R}(x) = \begin{cases} 1 & \text{if } \mathrm{dom}(x) = D,\ \mathrm{ran}(x) = R,\ \text{and } x(i) = j, \\ 0 & \text{otherwise.} \end{cases}$$

Every element of $V^{(2,1)}$ is of the form

$$\sum_{D,R} \sum_{i,j} a_{i,j}^{D,R} \delta_{i\mapsto j}^{D,R},$$

where, for every choice of D, R,

$$\sum_{i,j} a_{i,j}^{D,R} = 0.$$

When ranking 3 candidates, choosing a domain, a range, and the ranking of one of the candidates automatically defines the unordered set of rankings for the other two candidates. Thus, for the analogous second-order unordered rank-3 functions, we have (for $\{i_1, i_2\} \subset D, \{j_i, j_2\} \subset R$),

$$\delta^{D,R}_{\{i_1,i_2\}\to\{j_1,j_2\}} = \delta^{D,R}_{D\setminus\{i_1,i_2\}\to R\setminus\{j_1,j_2\}}.$$

$V^{(2,1)}$ therefore carries pure second-order unordered statistics as well.

We may also define the easily interpretable first-order rank-3 functions

$$\delta^3_{i\mapsto j} = \sum_{x\in R_5} \delta^3_{i\mapsto j}(x)\, \lfloor x \rfloor\,,$$

where

$$\delta^3_{i\mapsto j}(x) = \begin{cases} 1 & \text{if } \mathrm{rk}(x) = 3, \text{ and } x(i) = j, \\ 0 & \text{otherwise,} \end{cases}$$

and the analogous easily interpretable rank-3 second-order unordered functions $\delta^3_{\{i_1,i_2\}\to\{j_1,j_2\}}$.

Inner products of $f^{(2,1)}$ with the $\delta^3_{i\mapsto j}$ are given in Table 10.13, and inner products of $f^{(2,1)}$ with the $\delta^3_{\{i_1,i_2\}\to\{j_1,j_2\}}$ are given in Table 10.14. These inner products are derived from sums of the inner products of $f^{(2,1)}$ with the $\delta^{D,R}_{i\mapsto j}$ (given in Tables 10.15 through 10.24). Since the numbers occurring in this section are smaller, entries have been rounded to two decimal places. Also, if we denote the rank-3 portion of f by f_3, that is,

$$f_3 = \sum_{s\in R_n : \mathrm{rk}(s)=3} f(s)\, \lfloor s \rfloor\,,$$

Table 10.13: First-order derived groupoid analysis, rank-3 data

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-62	12	50	0	0
2	-60.33	-10.33	70.67	0	0
3	75	-31	-44	0	0
4	44.33	-1.67	-42.67	0	0
5	3	31	-34	0	0

Table 10.14: Second-order unordered derived groupoid analysis, rank-3 data

	Rank		
Candidates	**1,2**	**1,3**	**2,3**
1,2	-80	16	64
1,3	113.33	-32.67	-80.67
1,4	-51	11	40
1,5	-32.33	-6.33	38.67
2,3	-28	21	7
2,4	33.33	-16.67	-16.67
2,5	4	-10	6
3,4	-21.67	32.33	-10.67
3,5	-19.67	10.33	9.33
4,5	82	-25	-57

then we have

$$||f^{(3)} + f^{(2,1)}|| > .996||f_3||,$$

so we discard the projection $f^{(1,1,1)}$ from our analysis.

Diaconis's technique creates the numbers in Tables 10.8 and 10.9 for the rank-3 data. It is clear from the differences in these tables that our analysis is different from Diaconis's. We find the same pattern in the second-order data—a strong effect for choosing candidates 1,3 or 4,5, with 1,3 dominating 4,5. In the first-order data, Diaconis points out that candidates 4 and 5 are preferred among all rank-3 voters. This is true, though at first glance the numbers in Table 10.13

seem to paint a slightly different picture. To explain, note that the numbers in Table 10.13 are just the sums of the numbers in Tables 10.15 through 10.24, and if we examine these ten tables, then a new and interesting structure in the data emerges. We see that there is a positive first-order effect (and often a strong one) for ranking candidate 3 in position 1 whenever candidate 3 is ranked, *except* when the other two candidates ranked are (naturally) candidates 4 and 5. The rook monoid approach therefore offers a more "local," more granular inspection of the data than the symmetric group approach does, in that it allows us to see how the natural subsets of the rank-k voters vote amongst themselves. The disadvantage with the rook monoid approach is that the summary of this more granular data (as in Table 10.13) might be deceptive. In this case, the fact that nearly half of the rank-3 voters simply didn't rank candidate 3 (and thus, in this particular dataset, implicitly ranked him somewhere in the last two places—this can be read from Table 10.12) does not matter (and therefore does not count against him) in the creation of Table 10.13.

Table 10.15: First-order raw groupoid analysis, $D = \{1, 2, 3\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	3	11	-14	0	0
2	-40	-19	59	0	0
3	37	8	-45	0	0
4	0	0	0	0	0
5	0	0	0	0	0

Table 10.16: First-order raw groupoid analysis, $D = \{1, 2, 4\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-24.33	-6.33	30.67	0	0
2	10.67	4.67	-15.33	0	0
3	0	0	0	0	0
4	13.67	1.67	-15.33	0	0
5	0	0	0	0	0

Table 10.17: First-order raw groupoid analysis, $D = \{1, 2, 5\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-5.67	-10.67	16.33	0	0
2	-7.67	4.33	3.33	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	13.33	6.33	-19.67	0	0

Table 10.18: First-order raw groupoid analysis, $D = \{1, 3, 4\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	6.67	9.67	-16.33	0	0
2	0	0	0	0	0
3	9.67	-1.33	-8.33	0	0
4	-16.33	-8.33	24.67	0	0
5	0	0	0	0	0

Table 10.19: First-order raw groupoid analysis, $D = \{1, 3, 5\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-3.33	17.67	-14.33	0	0
2	0	0	0	0	0
3	27.67	-12.33	-15.33	0	0
4	0	0	0	0	0
5	-24.33	-5.33	29.67	0	0

Table 10.20: First-order raw groupoid analysis, $D = \{1, 4, 5\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-38.33	-9.33	47.67	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	18.67	1.67	-20.33	0	0
5	19.67	7.67	-27.33	0	0

Table 10.21: First-order raw groupoid analysis, $D = \{2, 3, 4\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	0	0	0	0	0
2	-10	2	8	0	0
3	7	-9	2	0	0
4	3	7	-10	0	0
5	0	0	0	0	0

Table 10.22: First-order raw groupoid analysis, $D = \{2, 3, 5\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	0	0	0	0	0
2	-5	2	3	0	0
3	4	-5	1	0	0
4	0	0	0	0	0
5	1	3	-4	0	0

Table 10.23: First-order raw groupoid analysis, $D = \{2, 4, 5\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	0	0	0	0	0
2	-8.33	-4.33	12.67	0	0
3	0	0	0	0	0
4	7.67	5.67	-13.33	0	0
5	0.67	-1.33	0.67	0	0

Table 10.24: First-order raw groupoid analysis, $D = \{3, 4, 5\}, R = \{1, 2, 3\}$

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	0	0	0	0	0
2	0	0	0	0	0
3	-10.33	-11.33	21.67	0	0
4	17.67	-9.33	-8.33	0	0
5	-7.33	20.67	-13.33	0	0

Let us now consider the projections f^{λ} for $\lambda \vdash 2$. $V^{(2)}$ contains both the constant and second-order unordered information for rank-2 votes, and $V^{(1,1)}$ contains both the first-order and the second-order ordered information for rank-2 votes.

Table 10.25 is both the zeroth-order and the second-order unordered analysis. It contains the inner products of $f^{(2)}$ with the $\delta^{D,R}$ and the $\delta^{D,R}_{\{i_1,i_2\}\to\{j_1,j_2\}}$ (which are equal), where D, R are 2-subsets of $\{1,2,3,4,5\}$.

Table 10.25: Zeroth-order and second-order unordered groupoid analysis, rank-2 data

	Rank				
Candidates	**1,2**	**1,3**	**1,4**	$\cdots$	**4,5**
1,2	155	0	0	$\cdots$	0
1,3	849	0	0	$\cdots$	0
1,4	178	0	0	$\cdots$	0
1,5	189	0	0	$\cdots$	0
2,3	140	0	0	$\cdots$	0
2,4	201	0	0	$\cdots$	0
2,5	150	0	0	$\cdots$	0
3,4	104	0	0	$\cdots$	0
3,5	157	0	0	$\cdots$	0
4,5	339	0	0	$\cdots$	0

Table 10.26 contains the inner products of $f^{(1,1)}$ with the $\delta^2_{i\mapsto j}$. These values are the sums of the inner products of $f^{(1,1)}$ with the $\delta^{D,R}_{i\mapsto j}$ (where D and R are 2-subsets of $\{1,2,3,4,5\}$), which are omitted.

Again, Diaconis's method produces different numbers from ours, although both approaches reveal the same general pattern in the data. Diaconis's method gives the results in Tables 10.10 and 10.11.

Finally, in our case, there is nothing to analyze for the rank-1 data, since we have only one isotypic subspace for rank-1 data, and the result of the rank-1 analysis would be the number of rank-1 voters ranking candidate i in position j,

Table 10.26: First-order derived groupoid analysis, rank-2 data

	Rank				
Candidates	**1**	**2**	**3**	**4**	**5**
1	-154.5	154.5	0	0	0
2	-33	33	0	0	0
3	160	-160	0	0	0
4	51	-51	0	0	0
5	-23.5	23.5	0	0	0

and this is already tallied in the dataset. Diaconis's method applied to the rank-1 data yields the average vote along with, for each candidate, the deviation from the average vote.

10.2.2 The Semigroup Basis Association

Let us now associate the semigroup basis of $\mathbb{C}R_5$ to the basis of characteristic functions of R_5. We have $f \in \mathbb{C}R_5$ by

$$f = \sum_{s \in R_5} f(s)s.$$

Note that f, when expressed with respect to the groupoid basis, is

$$\sum_{x \in R_5} g(x) \lfloor x \rfloor$$

where

$$g(x) = \sum_{t \geq x} f(t).$$

In other words, $g(x)$ is just the number of votes which *extend* the partial ranking x.

As with the groupoid basis association approach, we will project f onto the

isotypic subspaces of $\mathbb{C}R_5$ and inner product with the appropriate easily interpretable functions. Let E be an easily interpretable function for rank-k data. That is, $E : R_5 \to \{0, 1\}$ and $E(x) = 0$ if $\mathrm{rk}(x) \neq k$. Let us denote by E_G and E_S denote the image of E in $\mathbb{C}R_5$ under the groupoid basis association and the semigroup basis association, respectively. We claim that it does not matter whether we compute inner products of the projections of f with E_G or E_S. To see this, let $\lambda \vdash k$. We have

$$E_G = \sum_{x \in R_5 : \mathrm{rk}(x) = k} E(x) \lfloor x \rfloor, \quad E_S = \sum_{x \in R_5 : \mathrm{rk}(x) = k} E(x) x.$$

Now, E_S, when expressed with respect to the $\lfloor x \rfloor$ basis, is of the form

$$\sum_{x \in R_5 : \mathrm{rk}(x) = k} E(x) \lfloor x \rfloor + \sum_{x \in R_5 : \mathrm{rk}(x) < k} z(x) \lfloor x \rfloor .$$

for some function $z(x)$ on R_5. Also, the projection f^λ, when expressed in terms of the $\lfloor x \rfloor$ basis, contains nonzero coefficients only for elements $\lfloor x \rfloor$ for which $\mathrm{rk}(x) = k$. Therefore, $\langle f^\lambda, E_G \rangle = \langle f^\lambda, E_S \rangle$.

Spectral analysis of f using the semigroup basis association, then, amounts to the same analysis that the groupoid basis association does on the function g, where

$$g(x) = \sum_{t \geq x} f(t).$$

The rank-5 spectral analysis results from the semigroup basis association for the APA election are thus equal to the results in Section 10.1, and the results of the semigroup basis association analysis for the other ranks of this dataset are summarized in Tables 10.27 through 10.32.

Table 10.27: Zeroth-order semigroup analysis, rank-3 data

	Range				
Domain	**1,2,3**	**1,2,4**	**1,2,5**	$\cdots$	**3,4,5**
1,2,3	1212	0	0	$\cdots$	0
1,2,4	708	0	0	$\cdots$	0
1,2,5	626	0	0	$\cdots$	0
1,3,4	619	0	0	$\cdots$	0
1,3,5	1094	0	0	$\cdots$	0
1,4,5	1084	0	0	$\cdots$	0
2,3,4	478	0	0	$\cdots$	0
2,3,5	434	0	0	$\cdots$	0
2,4,5	1019	0	0	$\cdots$	0
3,4,5	572	0	0	$\cdots$	0

Table 10.28: First-order derived semigroup analysis, rank-3 data

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-304	236	68	0	0
2	-374.33	-22.33	396.67	0	0
3	563.33	-191.66	-371.67	0	0
4	138.66	-107.33	-31.33	0	0
5	-23.67	85.33	-61.67	0	0

Table 10.29: Second-order unordered derived semigroup analysis, rank-3 data

	Rank		
Candidates	**1,2**	**1,3**	**2,3**
1,2	-316.67	-59.67	376.33
1,3	612	-222	-390
1,4	-200.67	71.33	129.33
1,5	-162.67	-25.67	188.33
2,3	-94	139	-45
2,4	41	-35	-6
2,5	-27	-22	49
3,4	-103.33	191.67	-88.33
3,5	-43	83	-40
4,5	294.33	-120.67	-173.67

Table 10.30: Zeroth-order and second-order unordered semigroup analysis, rank-2 data

	Rank				
Candidates	**1,2**	**1,3**	**1,4**	$\cdots$	**4,5**
1,2	687	0	0	$\cdots$	0
1,3	2436	0	0	$\cdots$	0
1,4	781	0	0	$\cdots$	0
1,5	961	0	0	$\cdots$	0
2,3	754	0	0	$\cdots$	0
2,4	977	0	0	$\cdots$	0
2,5	816	0	0	$\cdots$	0
3,4	557	0	0	$\cdots$	0
3,5	814	0	0	$\cdots$	0
4,5	1525	0	0	$\cdots$	0

Table 10.31: First-order derived semigroup analysis, rank-2 data

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	-424.5	424.5	0	0	0
2	-209	209	0	0	0
3	537.5	-537.5	0	0	0
4	174	-174	0	0	0
5	-78	78	0	0	0

Table 10.32: Zeroth-order semigroup analysis, rank-1 data

	Rank				
Candidate	**1**	**2**	**3**	**4**	**5**
1	2903	0	0	0	0
2	2289	0	0	0	0
3	4016	0	0	0	0
4	3239	0	0	0	0
5	3002	0	0	0	0

Chapter 11

Further Directions

The generalization of the theory of Fourier transforms to inverse semigroups and beyond presents a new set of interesting challenges. Theorem 5.2.7 opens the door for the development of further inverse-semigroup FFTs, as it reduces the problem of calculating their Fourier transforms to the problems of calculating Fourier transforms on their maximal subgroups and calculating zeta transforms on their poset structures. While the theory of group FFTs is well-developed, the theory of fast zeta transforms is not. An interesting line of research, then, would be to create a theory of fast zeta transforms for inverse semigroup posets. On the other hand, the poset structure of an inverse semigroup can be about as bad as one wants—any meet semilattice is possible. It remains to be seen whether there are any guiding principles one might employ when creating fast zeta transforms.

Theorem 5.1.1 says that the representation theory of inverse semigroups is completely understood, at least in principle. In particular, if one understands the representations of the maximal subgroups of an inverse semigroup S, then one can tensor up those representations to create all of the representations of $\mathbb{C}S$, and hence of S as well by restricting to the semigroup basis. The representations of

$\mathbb{C}S$ defined in this way are naturally defined on the groupoid basis of $\mathbb{C}S$, and some interesting combinatorics might arise in these representations when changing back to the semigroup basis. For instance, it might be interesting to work out an explicit description of the representations of the signed rook monoid from the representations of the signed symmetric group. More interestingly, it appears that the Möbius transform involved in defining the groupoid basis of $\mathbb{C}S$ might somehow be applicable to more general algebras—in particular, to certain Iwahori-Hecke algebras—and might therefore be useful in understanding their representation theory as well.

We would like to develop further applications of the Fourier transforms presented in this thesis. In Chapter 10, we gave an application of the rook monoid FFT to the statistical analysis of partially ranked data. Rook wreath products arise as the partial automorphisms of nested designs, so their FFTs should have statistical implications as well. We can also use representations to study the underlying processes that generate these sorts of data. For instance, instead of using the representation theory of the symmetric group to statistically analyze voting datasets, in [10], the authors use the representation theory of the symmetric group to study voting systems and paradoxes. The representation theory of the rook monoid and its wreath products should have similar applications.

Finally, the question of how we should generalize the notion of the Fourier transform to semigroups whose algebras are not semisimple (such as the full transformation semigroup on n elements) remains open.

Bibliography

[1] U. Baum, *Existence and efficient construction of fast Fourier transforms for supersolvable groups*, Comput. Complex. **1/3** (1992), 235–256.

[2] L. I. Bluestein, *A linear filtering approach to the computation of the discrete Fourier transform*, IEEE Trans. Electroacoustics **18** (1970), 451–455.

[3] E. O. Brigham, *The Fast Fourier Transform and Applications*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[4] M. Clausen and U. Baum, *Fast Fourier transforms for symmetric groups: Theory and implementation*, Math. Comput. **61** (1993), no. 204, 833–847.

[5] A. H. Clifford, *Matrix representations of completely simple semigroups*, Amer. J. Math. **64** (1942), no. 1, 327–342.

[6] A. H. Clifford and G. B. Preston, *The Algebraic Theory of Semigroups*, vol. 1, Mathematical Surveys No. 7, AMS, Providence, RI, 1961.

[7] ———, *The Algebraic Theory of Semigroups*, vol. 2, Mathematical Surveys No. 7, AMS, Providence, RI, 1961.

[8] J. W. Cooley and J. W. Tukey, *An algorithm for machine calculation of complex Fourier series*, Math. Comput. **19** (1965), 297–301.

[9] C. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*, John Wiley and Sons, New York, 1962.

[10] Z. Daugherty, A. Eustis, G. Minton, and M. Orrison, *Voting, the symmetric group, and representation theory*, Preprint (Dec. 2007).

[11] P. Diaconis, *A generalization of spectral analysis with application to ranked data*, Ann. Statist. **17** (1989), no. 3, 949–979.

[12] P. Diaconis and D. Rockmore, *Efficient computation of the Fourier transform on finite groups*, J. Amer. Math. Soc. **3** (1990), no. 2, 297–332.

[13] B. Farb and R. K. Dennis, *Noncommutative Algebra*, Graduate Texts in Mathematics, vol. 144, Springer-Verlag, New York-Heidelberg, 1993.

[14] J. A. Green, *On the structure of semigroups*, Annals Math. **54** (1951), 163–172.

[15] C. Grood, *A Specht module analog for the rook monoid*, Electron. J. Combin. **9** (2002).

[16] T. Halverson, *Representations of the q-rook monoid*, J. Algebra **273** (2004), 227–251.

[17] R. B. Holmes, *Mathematical foundation of signal processing II. The role of group theory*, Massachusetts Institute of Technology Technical Report 781 (October 13, 1987).

[18] G. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Encyclopedia of Mathematics and its Applications, vol. 16, Cambridge University Press, 1984.

[19] S. Janson and V. Mazorchuk, *Some remarks on the combinatorics of IS_n*, Semigroup Forum **70** (2005), no. 3, 391–405.

[20] M. V. Lawson, *Inverse Semigroups: The Theory of Partial Symmetries*, World Scientific, Singapore, 1998.

[21] D. K. Maslen, *The efficient computation of Fourier transforms on the symmetric group*, Math. Comput. **67** (1998), no. 223, 1121–1147.

[22] D. K. Maslen and D. N. Rockmore, *Adapted Diameters and FFTs on Groups*, Proc. 6th ACM-SIAM SODA, 253–262.

[23] ———, *Generalized FFTs—A survey of some recent results*, Proc. 1995 DIMACS Workshop on Groups and Computation **28** (1997), 183–238.

[24] ———, *Separation of variables and the computation of Fourier transforms on finite groups, I*, J. Amer. Math. Soc. **10** (1997), no. 1, 169–214.

[25] ———, *The Cooley-Tukey FFT and group theory*, Notices of the AMS **48** (2001), no. 10, 1151–1161.

[26] J. D. P. Meldrum, *Wreath Products of Groups and Semigroups*, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 74, Longman Group Limited, Essex, England, 1995.

[27] W. D. Munn, *On semigroup algebras*, Proc. Cambridge Philos. Soc. **51** (1955), 1–15.

[28] ———, *The characters of the symmetric inverse semigroup*, Proc. Cambridge Philos. Soc. **53** (1957), 13–18.

[29] ———, *Matrix representations of semigroups*, Proc. Cambridge Philos. Soc. **53** (1957), 5–12.

[30] A. Ram, *Seminormal representations of Weyl groups and Iwahori-Hecke algebras*, Proc. London Math. Soc. **75** (1997), no. 1, 99–133.

[31] J. Rhodes and Y. Zalcstein, *Elementary representation and character theory of finite semigroups and its application*, Monoids and Semigroups with Applications, pp. 334–367, World Sci. Publishing, River Edge, NJ, 1991.

[32] D. N. Rockmore, *Fast Fourier transforms for wreath products*, Appl. Comput. Harmon. Anal. **2** (1995), 279–292.

[33] ———, *Recent progress and applications in group FFTs*, NATO Science Series, Computational Noncommutative Algebra and Applications, vol. 136, pp. 227–254, Springer Netherlands, 2005.

[34] ———, *Some applications of generalized FFTs*, Proc. 1995 DIMACS Workshop on Groups and Computation **28** (1997), 329-369.

[35] J. P. Serre, *Linear Representations of Finite Groups*, Graduate Texts in Mathematics, vol. 42, Springer-Verlag, New York-Heidelberg, 1977.

[36] L. Solomon, *Representations of the rook monoid*, J. Algebra **256** (2002), 309–342.

[37] R. Stanley, *Enumerative Combinatorics. Vol. 1*, Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, 1997.

[38] B. Steinberg, *Möbius functions and semigroup representation theory*, J. Comb. Theor. Ser. A. **113** (2006), 866–881.

[39] ———, *Möbius functions and semigroup representation theory II: Character formulas and multiplicities*, Adv. In Math. **217** (2008), 1521–1557.

[40] F. Yates, *The design and analysis of factorial experiments*, Imp. Bur. Soil Sci. Tech. Comm. **35** (1937).

[41] A. Young, *The Collected Papers of Alfred Young, 1873-1940*, University of Toronto Press, Toronto, 1977.

Index