Chapter 1. Introduction

This monograph delves into the uses of group theory, particularly noncommutative Fourier analysis, in probability and statistics. It presents useful tools for applied problems and develops familiarity with one of the most active areas in modern mathematics.

Groups arise naturally in applied problems. For instance, consider 500 people asked to rank 5 brands of chocolate chip cookies. The rankings can be treated as permutations of 5 objects, leading to a function on the group of permutations S_5 (how many people choose ranking π). Group theorists have developed natural bases for the functions on the permutation group. Data can be analyzed in these bases. The "low order" coefficients have simple interpretations such as "how many people ranked item i in position j." Higher order terms also have interpretations and the benefit of being orthogonal to lower order terms. The theory developed includes the usual spectral analysis of time series and the analysis of variance under one umbrella.

The second half of this monograph develops such techniques and applies them to partially ranked data, and data with values in homogeneous spaces such as the circle and sphere. Three classes of techniques are suggested — techniques based on metrics (Chapter 6), techniques based on direct examination of the coefficients in a Fourier expansion (spectral analysis, Chapter 8), and techniques based on building probability models (Chapter 9).

All of these techniques lean heavily on the tools and language of group representations. These tools are developed from first principles in Chapter 2. Fortunately, there is a lovely accessible little book — Serre's *Linear Representations of Finite Groups* — to lean on. The first third of this may be read while learning the material.

Classically, probability precedes statistics, a path followed here. Chapters 3 and 4 are devoted to concrete probability problems. These serve as motivation for the group theory and as a challenging research area. Many of the problems have the following flavor: how many times must a deck of cards be shuffled to bring it close to random? Repeated shuffling is modeled as repeatedly convolving a fixed probability on the symmetric group. As usual, the Fourier transform turns the analysis of convolutions into the analysis of products. This can lead to very explicit results as described in Chapter 3. Chapter 4 develops some "pure probability" tools - the methods of coupling and stopping times - for random walk problems.

Both card shuffling and data analysis of permutations require detailed knowledge of the representation theory of the symmetric group. This is developed in Chapter 7. Again, a friendly, short book is available: G. D. James' *Representation*