William J. DeMeo, ed.

Group Theory in Signal Processing

4 February 2004

TEXTRON Systems Corporation

Hawaii Operations
535 Lipoa Parkway, Suite 149
Kihei, HI 96753

Preface

This document is divided into two parts: Part I is a compendium of mathematics that I find interesting and useful for working in the field know as digital signal processing (DSP). This serves to introduce a powerful, if somewhat abstract, cast of characters from the fields of group theory and linear algebra. Part II then brings these characters down to earth by giving them roles to play in the "real world."

At various places in the book – especially during the more abstract lines of reasoning – the term *digital signal processing* may seem out-of-place, but, as the reader will see, an underlying theme of this work is the importance of interpreting each word in this phrase at whatever level of abstraction is most appropriate to the given context. Typically, a "signal" is a function or process of interest, or data which represent this function or process. The term "processing" refers to understanding, analysis, synthesis, manipulation, etc. of the signal. In this work, the term "digital" merely indicates that the focus in on math tools and methods that are particularly well suited to functions that have discrete representations.

Each application chapter of Part II describes a research area or problem of practical interest from within the field of DSP, and then demonstrates how the mathematical concepts from Part I manifest in elegant and revealing mathematical expressions of the problem. The approach provides powerful means of understanding and analysis which are difficult, if not impossible, with other approaches.

For the reader's part, becoming well acquainted with the theory in Part I can demand significant investments of time, patience, and diligence. However, as noted above, the rewards are substantial.

Kihei, Hawaii

William DeMeo
August 2003

Contents

Part I

Theory

1

Abelian Groups and Univariate DSP

Richard Tolimieri, Myoung An[1], and William DeMeo[2]

[1] psypher@tiac.net
[2] williamdemeo@yahoo.com

Although this chapter assumes some basic group theory, such as that reviewed in Chapter B.2, the following section reviews those prerequisites that arise most frequently in our application of the theory. Further details can be found in [3], and Chapters 1 and 2 of [7].

1.1 Preliminaries

Throughout, unless otherwise stated, A denotes an abelian group and $\mathcal{L}(A)$ is the space of all complex valued functions on A.

1.1.1 Translation and Convolution

Definition 1 (Unitary Transform, Isometry).
A linear mapping $\mathsf{T} : \mathcal{L}(A) \to \mathcal{L}(A)$ is called *unitary* if, for all $f, g \in \mathcal{L}(A)$,

$$(\mathsf{T}f, \mathsf{T}g) = (f, g)$$

More generally, a linear mapping $\tau : \mathcal{L}(A_1) \to \mathcal{L}(A_2)$ is called an *isometry* if, for all $f, g \in \mathcal{L}(A_1)$,

$$(\tau f, \tau g) = (f, g)$$

where the inner product (f, g) is taken in $\mathcal{L}(A_1)$ and the inner product $(\tau f, \tau g)$ is taken in $\mathcal{L}(A_2)$.

Many important linear mappings are isometries.

Theorem 1. *A linear mapping* T *is unitary if and only if* T *maps an orthonormal basis onto an orthonormal basis.*

Definition 2 (Translation).
For $y \in A$, the mapping $\mathsf{T}(y)$ of $\mathcal{L}(A)$ defined by

$$(\mathsf{T}(y)f)(x) = f(y^{-1}x), \qquad f \in \mathcal{L}(A), x \in A$$

is a linear operator of $\mathcal{L}(A)$ called *translation by* y.

In the interest of generality, Definition 2 denotes the inverse of $y \in A$ by y^{-1}. Often translation defined for the special, additive inverse case, $y^{-1}x = x - y$.

Denote the collection of all translations of $\mathcal{L}(A)$ by $\mathsf{T}(A) = \{\mathsf{T}(y) : y \in A\}$. The mapping $\mathsf{T} : A \to \mathsf{T}(A)$ satisfies

$$\mathsf{T}(x + y) = \mathsf{T}(x)\mathsf{T}(y), \qquad x, y \in A,$$

$$\mathsf{T}(y)^{-1} = \mathsf{T}(y^{-1}), \qquad y \in A$$

The collection $\mathsf{T}(A)$ is closed under operator composition and operator inverse. Since A is abelian, and addition commutative, $\mathsf{T}(A)$ is a *commuting family of linear operators*. Since A is finite, $y \in A$ implies $Ny = 0$, and

$$\mathsf{T}(y)^N = \mathsf{T}(Ny) = \mathsf{T}(0) = \mathsf{I}$$

where I is the identity operator of $\mathcal{L}(A)$.

Definition 3 (Convolution).
For $g \in \mathcal{L}(A)$, the mapping $\mathsf{C}(g)$ of $\mathcal{L}(A)$ defined by

$$\mathsf{C}(g) = \sum_{y \in A} g(y)\mathsf{T}(y) \tag{1.1}$$

is a linear operator of $\mathcal{L}(A)$ called *convolution by* g.

The summation (1.1) is a linear combination of translations of $\mathcal{L}(A)$ using operator scalar multiplication and operator addition.

Often the convolution $\mathsf{C}(g)f$ is written as a binary operation, $f * g$, called the *convolution product* on $\mathcal{L}$(A). By definition, the convolution product of f and g, evaluated at $x \in A$, is

$$\begin{aligned}(\mathsf{C}(f)g)(x) = (f * g)(x) &= \sum_{y \in A} f(y)g(y^{-1}x) \\ &= \sum_{y \in A} g(y)f(xy^{-1}) \qquad \text{(change of variables)} \\ &= (\mathsf{C}(g)f)(x) = (g * f)(x) \qquad (\because\ A \text{ is abelian})\end{aligned}$$

Theorem 2. *Convolution is a commutative algebra product on $\mathcal{L}(A)$.*

Definition 4 (Convolution Algebra over A).
The algebra formed by $\mathcal{L}$(A)paired with the convolution product is called the *convolution algebra over* A.

1.1.2 Factor Groups

Many of our examples will focus on the group

$$N\mathbb{Z} = \{\ldots, -2N, -N, 0, N, 2N, \ldots\}, \qquad N \in \mathbb{Z}$$

as well as the *factor group*,

$$\mathbb{Z}/N\mathbb{Z} = \{0 + N\mathbb{Z},\ 1 + N\mathbb{Z}, \ldots,\ (N-1) + N\mathbb{Z}\} \tag{1.2}$$

The notation $k + N\mathbb{Z}$ in (1.2) denotes the group

$$k + N\mathbb{Z} = \{\ldots, k - 2N,\ k - N,\ k,\ k + N,\ k + 2N \ldots\}$$

Thus a factor group, such as $\mathbb{Z}/N\mathbb{Z}$, is really a "group of groups", or a group of *equivalence classes*. By choosing one element from each equivalence class as a *representative* of that class, we arrive at an isomorphism between the factor group and the group of all equivalence class representatives:

$$\mathbb{Z}/N\mathbb{Z} \simeq \{0, 1, 2, \ldots, (N-1)\} \tag{1.3}$$

The right hand side of (1.3) is often denoted by $\mathbb{Z}_N$, and the left hand side by $\mathbb{Z}/N$. In the sequel, we abuse this notation and simply refer to the equivalence as $\mathbb{Z}/N = \{0, 1, 2, \ldots, (N-1)\}$.

We also make heavy use of the group $L\mathbb{Z}/N\mathbb{Z}$, where L is a divisor of N, say $LM = N,\ M \in \mathbb{Z}$. Recall that, for such an L, the factor group $L\mathbb{Z}/N\mathbb{Z}$ is

$$\begin{aligned} L\mathbb{Z}/N\mathbb{Z} &= \{0 + N\mathbb{Z},\ L + N\mathbb{Z},\ 2L + N\mathbb{Z}, \ldots,\ (M-1)L + N\mathbb{Z}\} \\ &\simeq \{0,\ L,\ 2L, \ldots, (M-1)L\}, \\ &\simeq \{0, 1, 2, \ldots, (M-1)\} = \mathbb{Z}_M \end{aligned} \tag{1.4}$$

Again, abusing notation, we will write equivalence (1.4) as $L\mathbb{Z}/N = \{0,\ L,\ 2L, \ldots, (M-1)L\}$.

Finally, recall that

$$(\mathbb{Z}/N)/(L\mathbb{Z}/N) \simeq \mathbb{Z}/L \simeq \{0, 1, \ldots, (L-1)\} \tag{1.5}$$

which we also write $\mathbb{Z}/L = \{0, 1, \ldots, (L-1)\}$, as remarked above.

Example 1. Consider the group

$$L\mathbb{Z}/N \simeq \{0, L, 2L, \ldots, (M-1)L\}, \quad \text{for } LM = N,$$

and suppose $f \in \mathcal{L}(\mathbb{Z}/N)$ is $L\mathbb{Z}/N$-periodic. Then a vector representation of f is

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}^{(0)} \\ \mathbf{f}^{(0)} \\ \vdots \\ \mathbf{f}^{(0)} \end{pmatrix}, \quad \text{where} \quad \mathbf{f}^{(0)} = \begin{pmatrix} f(0) \\ f(1) \\ \vdots \\ f(L-1) \end{pmatrix}$$

That is $\mathbf{f} \in \mathbb{C}^N$ is a vector containing $M = N/L$ sub-vectors, $\mathbf{f}^{(0)} \in \mathbb{C}^L$.

Remark 1. The group $\mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ is useful for image processing applications. Note that the foregoing definitions of translation and convolution are general enough to apply to functions in $\mathcal{L}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)$, so long as we have a suitable definition for addition and additive inverse on $\mathbb{Z}/N_1 \times \mathbb{Z}/N_2$.

Definition 5 (Canonical Basis).
For $y \in \mathbb{Z}/N$, define $e_y \in \mathcal{L}(\mathbb{Z}/N)$ as the function whose value is 1 at y and 0 otherwise. The set $\{e_y : y \in \mathbb{Z}/N\}$ is called the *canonical basis*.

Translations permute the elements of the canonical basis by the equation $\mathsf{T}(y)e_x = e_{y+x}$, where $y+x$ is taken modulo N. In particular, $\mathsf{T}(y)$ maps the evaluation basis onto the evaluation basis. As such, by Theorem 1, translations are unitary operators.

1.2 Fourier Analysis on Finite Abelian Groups

1.2.1 Character Groups

Suppose A is a finite abelian group of order N. Denote by U the multiplicative group of all complex numbers of absolute value 1.

Definition 6 (Character). A mapping $a^* : A \to U$ is called a *character* of A if it is a homomorphism; that is, if the mapping satisfies,

$$a^*(a + b) = a^*(a)a^*(b), \qquad a, b \in A.$$

We often denote $a^*(a)$ by $\langle a, a^* \rangle$. Every character of A maps A into the subgroup U_N of all N-complex roots of unity. Denote by A^* the set of all characters of A. The set A^* is an abelian group, called the *character group* of A, under the composition law,

$$\langle a, a^* + b^* \rangle = \langle a, a^* \rangle \langle a, b^* \rangle, \qquad a \in A,\ a^*, b^* \in A^*$$

The mapping 0^* of A^* defined by

$$0^*(a) = \langle a, 0^* \rangle = 1, \qquad a \in A,$$

is the *zero* element in the group A^*. Since, for any $a^* \in A^*$,

$$\langle a, a^* \rangle \overline{\langle a, a^* \rangle} = 1, \qquad a \in A,$$

the additive inverse of a^* in A^* is given by

$$\langle a, -a^* \rangle = \overline{\langle a, a^* \rangle}, \qquad a \in A.$$

Example 2. Consider the pairing

$$\langle a, c \rangle = \exp(\mathrm{i}2\pi \frac{ca}{N}), \qquad a, c \in \mathbb{Z}/N$$

For $c \in \mathbb{Z}/N$, we define the mapping $\chi(c) : \mathbb{Z}/N \to U_N$ by the following equivalent expressions

$$\chi(c)(a) = \langle a, \chi(c) \rangle = \langle a, c \rangle = \exp(\mathrm{i}2\pi \frac{ca}{N}) \tag{1.6}$$

Thus, for all $c \in \mathbb{Z}/N$, the mapping $\chi(c)$ is a character of $\mathbb{Z}/N$, and the mapping $\chi : \mathbb{Z}/N \to (\mathbb{Z}/N)^*$ is an isomorphism of $\mathbb{Z}/N$ onto its character group $(\mathbb{Z}/N)^*$.

Any isomorphism from a finite abelian group A onto its character group A^* is called a *presentation* of A. The presentation defined in example 2 is called the *standard presentation* of $\mathbb{Z}/N$.

1.2.2 Character Formulas

Example 3. Suppose $A = \mathbb{Z}/N$ and N is even. Then the N-roots of unity contain -1 and are written as

$$1, v, \ldots, v^{M-1}, -1, -v, \ldots, -v^{M-1}, \qquad N = 2M,\ v = \exp(\frac{\mathrm{i}2\pi}{N})$$

implying that the sum of the N-roots of unity vanishes for N even. For N odd, this fact is still true but harder to show.

Theorem 3. *For $a^* \in A^*$,*

$$\sum_{a \in A} \langle a, a^* \rangle = \begin{cases} N, & a^* = 0^*, \\ 0, & \textit{otherwise.} \end{cases}$$

Proof. If $a^* = 0^*$, then $\langle a, a^* \rangle = 1$ for all $a \in A$, so the result is obvious. Suppose $a^* \neq 0^*$. Then there exists an $a_0 \in A$ such that $\langle a_0, a^* \rangle \neq 1$. Also, since A is a group, summing over all $a \in A$ is the same as summing over all $a + a_0 \in A$. Therefore,

$$\begin{aligned} \sum_{a \in A} \langle a, a^* \rangle &= \sum_{a \in A} \langle a + a_0, a^* \rangle \\ &= \langle a_0, a^* \rangle \sum_{a \in A} \langle a, a^* \rangle \end{aligned}$$

Since $\langle a_0, a^* \rangle \neq 1$, the sum must be 0. □

Theorem 3 implies, for example, that the N-complex roots of unity sum to 0.

Theorem 4. *For $a \in A$,*

$$\sum_{a^* \in A^*} \langle a, a^* \rangle = \begin{cases} N, & a = 0, \\ 0, & \textit{otherwise.} \end{cases}$$

The proof of Theorem 4 is the same as that for Theorem 3, *mutatis mutandis*. A simple but important corollary is

Corollary 1. For $a \in A$, if for all $a^* \in A^*$,

$$\langle a, a^* \rangle = 1$$

then $a = 0$.

1.2.3 Duality Theory

Denote by A^{**} the character group of A^*. The groups A and A^{**} are canonically isomorphic, and this permits A^{**} to be replaced by A in all discussions and results. We call this replacement *duality*. For $a \in A$, the mapping

$$\Theta(a) : A^* \to U$$

defined by

$$\Theta(a)(a^*) = \overline{\langle a, a^* \rangle}, \qquad a^* \in A^*,$$

is a character of A^*.

Theorem 5. *The mapping*

$$\Theta : A \to A^{**}$$

is an isomorphism of A onto A^{**}.

Proof. If $\Theta(a)$ is the trivial character of A^*, then

$$\langle a, a^* \rangle = 1, \text{ for all } a^* \in A^*,$$

So, by Corollary 1, $a = 0$. □

We identify A with A^{**} by the canonical isomorphism Θ and denote by a the character $\Theta(a)$ of A^*, for $a \in A$. It will be clear from context whether a is being viewed as an element of A or an element of A^{**}.

Definition 7 (Dual).
For a subgroup B of A, the set

$$B_* = \{a^* \in A^* : \langle b, a^* \rangle = 1, \text{ for all } b \in B\}$$

is a subgroup of A^* called the *dual* of B.

The subgroup B_* of A^* is the set of all characters of A that act trivially on B. For $y^* \in B_*, a \in A$, and $b \in B$,

$$y^*(a + b) = \langle a + b, y^* \rangle = \langle a, y^* \rangle \langle b, y^* \rangle = \langle a, y^* \rangle$$

Therefore, y^* defines a character of $A/B = \{a + B : a \in A\}$. We denote this character by $\theta(y^*) \in (A/B)^*$. It is defined by the formula,

$$\theta(y^*)(a + B) = \langle a + B, \theta(y^*) \rangle = \langle a, y^* \rangle, \qquad a \in A$$

Exercise 1. Consider the subgroup $L\mathbb{Z}/N$, of $\mathbb{Z}/N$, for $N = LM$. Identifying $\mathbb{Z}/N$ with $(\mathbb{Z}/N)^*$ by the standard presentation, show that[3]

$$(L\mathbb{Z}/N)_* = (M\mathbb{Z}/N)^*$$

That is, show that the *dual* of $L\mathbb{Z}/N$ is $(M\mathbb{Z}/N)^*$. Or, to put it another way, show that (the standard presentation of) the set of characters that act trivially on $L\mathbb{Z}/N$ is

$$M\mathbb{Z}/N = \{0,\ M,\ 2M, \ldots, (L-1)M\}$$

Solution 1. Recall from (1.4),

$$L\mathbb{Z}/N = \{0,\ L,\ 2L, \ldots, (M-1)L\} \tag{1.7}$$

Also, recall the definition of the dual; i.e., the set of characters that act trivially on $L\mathbb{Z}/N$:

$$(L\mathbb{Z}/N)_* = \{a^* \in (\mathbb{Z}/N)^* : \langle b, a^* \rangle = 1, \text{ for all } b \in L\mathbb{Z}/N\}$$

Expression (1.7) makes clear that any element $b \in L\mathbb{Z}/N$ can be written yL for some $y \in \{0, 1, 2, \ldots, (M-1)\}$. Thus,

[3] cf. typo on page 32 of [7].

$$\langle b, a^* \rangle = \exp(\mathrm{i}2\pi \frac{ayL}{N}) = \exp(\mathrm{i}2\pi \frac{ay}{M}), \qquad b \in L\mathbb{Z}/N$$

Now suppose $a^* \in (M\mathbb{Z}/N)^*$, so that $a^* = xM$, for some $x \in \{0, 1, 2, \ldots, (L-1)\}$. Then

$$\langle b, a^* \rangle = \exp(\mathrm{i}2\pi xy) = 1, \qquad b \in L\mathbb{Z}/N$$

Therefore, the elements of $(M\mathbb{Z}/N)^*$ act trivially on $L\mathbb{Z}/N$.

Next suppose $a^* \notin (M\mathbb{Z}/N)^*$, so that a^* is not a multiple of M, and let $b = L$. Then,

$$\langle b, a^* \rangle = \exp(\mathrm{i}2\pi \frac{a}{M}) \neq 1$$

We have thus shown that $a^* \in (L\mathbb{Z}/N)_*$ if and only if $a^* \in (M\mathbb{Z}/N)^*$, *q.e.d.*

Theorem 6. *The mapping*

$$\theta : B_* \to (A/B)^*$$

is a canonical isomorphism of B_* onto $(A/B)^*$.

For proof, see [7], page 32.

The identification of B_* with $(A/B)^*$, and B with $(A^*/B_*)^*$ results in the following refinements of the character formulas.

Corollary 2. For $x \in A$,

$$\sum_{x^* \in B_*} \langle x, x^* \rangle = \begin{cases} o(B_*), & x \in B, \\ 0, & \text{otherwise} \end{cases}$$

Corollary 3. For $x^* \in A^*$,

$$\sum_{x \in B} \langle x, x^* \rangle = \begin{cases} o(B), & x^* \in B_*, \\ 0, & \text{otherwise} \end{cases}$$

1.2.4 Character Group Basis

Consider the character group A^* as a subset of $\mathcal{L}(A)$, the space of all complex valued functions on A. In this section we show that A^* is an orthogonal basis of $\mathcal{L}(A)$.

Example 4. Consider the group $\mathbb{Z}/N$ and the standard presentation χ. Recalling the defining equation (1.6), the inner product in $\mathcal{L}(\mathbb{Z}/N)$ of two characters $\chi(a)$ and $\chi(b)$ is given by

$$(\chi(b), \chi(c)) = \sum_{a \in A} \langle a, \chi(b) \rangle \overline{\langle a, \chi(c) \rangle} = \sum_{a \in A} \exp(2\pi \mathrm{i} \frac{a(b-c)}{N}), \qquad b,\, c \in \mathbb{Z}/N$$

The sum vanishes unless $b = c$, in which case the sum is equal to N.

In general, we have the following result:

Theorem 7. *A^* is an orthogonal basis of $\mathcal{L}(A)$.*

Proof. Orthogonality follows from Theorem 3 and

$$(a^*, c^*) = \sum_{a \in A} \langle a, a^* \rangle \overline{\langle a, c^* \rangle} = \sum_{a \in A} \langle a, a^* - c^* \rangle, \qquad a^*,\, c^* \in A^*$$

Arguing by dimension completes the proof. □

Corollary 4.

$$\frac{1}{\sqrt{N}} A^*$$

is an orthonormal basis of $\mathcal{L}(A)$.

Technically, A^* is a basis only if some ordering is placed on A^*. In Fourier theory the distinction is significant since there is no canonical ordering on A^*. For the development of the theory, we use summation notation and order is no problem. However, in order to represent algorithms in terms of matrix factorization, some order must be taken, usually by realization. The form of the realization can affect the size and dimension of the algorithm.

1.2.5 Fourier Transform

Definition 8 (Fourier Expansion, Fourier Coefficient Set).
Suppose $f \in \mathcal{L}(A)$. The expansion of f over the character group basis A^* given by

$$f = \sum_{a^* \in A^*} \alpha(a^*) a^* \tag{1.8}$$

is called the *Fourier expansion* of f and the coefficient set

$$\alpha \in \mathcal{L}(A^*)$$

is called the *Fourier coefficient set*.

Note that expression (1.8), evaluated at a point $a \in A$, is

$$f(a) = \sum_{a^* \in A^*} \alpha(a^*) a^*(a) = \sum_{a^* \in A^*} \alpha(a^*) \langle a, a^* \rangle \tag{1.9}$$

Theorem 8. *If $f \in \mathcal{L}(A)$ has Fourier coefficient set $\alpha \in \mathcal{L}(A^*)$, then for all $a^* \in A^*$,*

$$\alpha(a^*) = \frac{1}{N}(f, a^*) \tag{1.10}$$

Proof. For $c^* \in A^*$,

$$\begin{aligned}
(f, c^*) &= \sum_{a \in A} f(a) \overline{\langle a, c^* \rangle} \\
&= \sum_{a \in A} \sum_{a^* \in A^*} \alpha(a^*) \langle a, a^* \rangle \overline{\langle a, c^* \rangle}, \qquad \text{by equation (1.9)} \\
&= \sum_{a^* \in A^*} \alpha(a^*) \sum_{a \in A} \langle a, a^* - c^* \rangle \\
&= \alpha(c^*) N
\end{aligned}$$

The final equality holds because, recall,

$$\sum_{a \in A} \langle a, a^* - c^* \rangle = \begin{cases} N, & \text{for } a^* - c^* = 0^* \\ 0, & \text{otherwise} \end{cases}$$

□

Note that expression (1.10) is equivalent to

$$\alpha(a^*) = \frac{1}{N} \sum_{a \in A} f(a) \overline{\langle a, a^* \rangle},$$

Example 5. Consider the group $\mathbb{Z}/N$ identified with its character group by the standard presentation. If $f \in \mathcal{L}(\mathbb{Z}/N)$ has Fourier coefficient set $\alpha \in \mathcal{L}(\mathbb{Z}/N)$, then

$$\alpha(c) = \frac{1}{N} \sum_{a=0}^{N-1} f(a) \exp(-2\pi \mathrm{i} \frac{ca}{N}), \qquad c \in \mathbb{Z}/N$$

Denoting by $\mathbf{f} \in \mathbb{C}^N$ the vector representation of f and by $F(N)$ the N-point Fourier transform matrix,

$$F(N) = \left\{ \exp(-2\pi \mathrm{i} \frac{ca}{N}) \right\}_{0 \le c, a < N}$$

the Fourier coefficient set $\alpha(c), c \in \mathbb{Z}/N$, is given by the matrix product

$$\alpha = \frac{1}{N} F(N) \mathbf{f}$$

Example 6. Consider the group $\mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ identified with its character group by the standard presentation. If

$$f \in \mathcal{L}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2), \qquad N = N_1 N_2,$$

has Fourier coefficient set $\alpha \in \mathcal{L}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)$, then for all $c = (c_1, c_2) \in \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$,

$$\alpha(c) = \frac{1}{N} \sum_{a_1=0}^{N_1-1} \sum_{a_2=0}^{N_2-1} f(a_1, a_2) \exp(-2\pi \mathrm{i} \frac{c_1 a_1}{N_1}) \exp(-2\pi \mathrm{i} \frac{c_2 a_2}{N_2})$$

Denote the matrices

$$[f(a_1, a_2)]_{\substack{0 \le a_1 < N_1 \\ 0 \le a_2 < N_2}} \qquad \text{and} \qquad [\alpha(c_1, c_2)]_{\substack{0 \le c_1 < N_1 \\ 0 \le c_2 < N_2}}$$

by $Mat(f)$ and $Mat(\alpha)$, respectively. Then the relationship between f and α can be written as

$$Mat(\alpha) = \frac{1}{N} F(N_1) Mat(f) F(N_2)$$

Definition 9 (Fourier Transform).
The *Fourier transform* F_A over A is the linear mapping $F_A : \mathcal{L}(A) \to \mathcal{L}(A^*)$ defined by

$$F_A f(a^*) = (f, a^*) = \sum_{a \in A} f(a) \overline{\langle a, a^* \rangle}, \qquad f \in \mathcal{L}(A),\ a^* \in A^*$$

The mapping F_A, up to a scalar factor $1/N$, maps f onto the coefficient set of its Fourier expansion over A.

By duality, the Fourier transform F_{A^*} over A^* is the linear mapping $F_{A^*} : \mathcal{L}(A^*) \to \mathcal{L}(A)$ defined by

$$F_{A^*} \alpha(a) = \sum_{a^* \in A^*} \alpha(a^*) \langle a, a^* \rangle, \qquad \alpha \in \mathcal{L}(A^*),\ a \in A$$

The composition $F_{A^*} F_A$ is a linear mapping from $\mathcal{L}(A)$ to $\mathcal{L}(A)$.

Theorem 9.

$$F_{A^*} F_A = N I_A$$

where I_A is the identity mapping on $\mathcal{L}(A)$.

Proof. The line of argument is the same as that of Theorem 8, *mutatis mutandis*. For $a \in A$,

$$\begin{aligned} F_{A^*} F_A f(a) &= \sum_{a^* \in A^*} F_A f(a^*) \langle a, a^* \rangle \\ &= \sum_{a^* \in A^*} \sum_{c \in A} f(c) \overline{\langle c, a^* \rangle} \langle a, a^* \rangle \\ &= \sum_{c \in A} f(c) \sum_{a^* \in A^*} \overline{\langle c - a, a^* \rangle} \\ &= f(a) N \end{aligned}$$

□

1.3 Poisson Summation Formula

The Poisson summation (PS) formula describes the fundamental duality between periodization and decimation operators under the Fourier transform. In this section, the finite abelian group version of the PS formula is derived as a simple application of the character formulas of Section 1.2.2.

In sampling applications, the main use for the PS formula is to provide a procedure for preprocessing a signal for the Fourier transform and other computations by establishing a Nyquist sampling rate. In FFT algorithm design, it is the first step in a divide-and-conquer strategy.

The text [7] – the source for these notes – uses the PS formula to unify basic algorithmic procedures in time-frequency processing and to derive closed form formulas that play a significant role in *Weyl-Heisenberg duality theory*, which is the subject of Chapter 13 of [7].

1.3.1 Statement and proof

Assume B is a subgroup of a finite abelian group A and $f \in \mathcal{L}(A)$. Denote the orders of A and B by N and M, respectively, with $N = LM$. Consider the Fourier expansion of f,

$$f = \sum_{a^* \in A^*} \alpha(a^*) a^*$$

with coefficient set

$$\alpha \in \mathcal{L}(A^*)$$

Definition 10 (Periodic).
The function $f \in \mathcal{L}(A)$ is called *B-periodic* if, for all $a \in A$ and $b \in B$,

$$f(a+b) = f(a)$$

and $\alpha \in \mathcal{L}(A^*)$ is called B_*-*decimated* if α vanishes off of B_*.

Example 7. Suppose $f \in \mathcal{L}(\mathbb{Z}/N)$ is $L\mathbb{Z}/N$-periodic. Recall,

$$L\mathbb{Z}/N \simeq \{0, L, 2L, \ldots, (M-1)L\}, \qquad \text{for } LM = N$$

and the vector representation of f can be written $\mathbf{f} = (\mathbf{f}^{(0)}, \mathbf{f}^{(0)}, \ldots, \mathbf{f}^{(0)})^t$, That is $\mathbf{f} \in \mathbb{C}^N$ is a vector containing $M = N/L$ sub-vectors, $\mathbf{f}^{(0)} \in \mathbb{C}^L$. Identifying $\mathbb{Z}/N$ with its character group by the standard presentation, the Fourier coefficient set $\alpha \in \mathcal{L}(\mathbb{Z}/N)$ of f is given by

$$\alpha = \frac{1}{N} F(N) \mathbf{f}$$

Consider the value of α at $a^* \in (\mathbb{Z}/N)^*$.

$$\begin{aligned}
\alpha(a^*) &= \frac{1}{N} \sum_{n=0}^{N-1} f(n) \overline{\langle n, a^* \rangle} \\
&= \frac{1}{N} \sum_{a \in \mathbb{Z}/L} \sum_{b \in L\mathbb{Z}/N} f(a+b) \overline{\langle a+b, a^* \rangle} \\
&= \frac{1}{N} \sum_{a \in \mathbb{Z}/L} f(a) \overline{\langle a, a^* \rangle} \sum_{b \in L\mathbb{Z}/N} \overline{\langle b, a^* \rangle}
\end{aligned} \tag{1.11}$$

Corollary 3 implies, for $a^* \in (\mathbb{Z}/N)^*$,

$$\sum_{b \in L\mathbb{Z}/N} \overline{\langle b, a^* \rangle} = \begin{cases} o(L\mathbb{Z}/N), & a^* \in (L\mathbb{Z}/N)_*, \\ 0, & \text{otherwise} \end{cases}$$

Recall that $M\mathbb{Z}/N = (L\mathbb{Z}/N)_*$, the dual of $L\mathbb{Z}/N$. Therefore, since $o(L\mathbb{Z}/N) = M$, (1.11) becomes

$$\alpha(a^*) = \begin{cases} \frac{1}{L} \sum_{a \in \mathbb{Z}/L} f(a) \overline{\langle a, a^* \rangle}, & a^* \in M\mathbb{Z}/N, \\ 0, & \text{otherwise} \end{cases}$$

Thus we have shown that, for any $a^* \notin M\mathbb{Z}/N$, the value of $\alpha(a^*)$ vanishes. So we say that α is $(L\mathbb{Z}/N)_*$-decimated.

Before stating this result more generally as Theorem 10, let's consider a more concrete example. The value of α at 1 is

$$\begin{aligned}\alpha(1) &= \frac{1}{N}\sum_{n=0}^{N-1} f(n)\overline{\langle n,1\rangle} \\ &= \frac{1}{N}\sum_{a=0}^{L-1}\sum_{y=0}^{M-1} f(a+yL)\overline{\langle a+yL,1\rangle} \\ &= \frac{1}{N}\sum_{a=0}^{L-1} f(a)\overline{\langle a,1\rangle}\sum_{y=0}^{M-1}\exp(-\mathrm{i}2\pi\frac{yL}{N})\end{aligned}$$

which vanishes since

$$\sum_{y=0}^{M-1}\exp(-\mathrm{i}2\pi\frac{yL}{N}) = \sum_{y=0}^{M-1}\exp(-\mathrm{i}2\pi\frac{y}{M}) = 0$$

The foregoing is a special case of the following result:

Theorem 10. *If $f \in \mathcal{L}(A)$ has Fourier coefficient set α, then f is B-periodic if and only if α is B_*-decimated.*

Proof. Recall that

$$f(a) = \sum_{x^*\in A^*}\alpha(x^*)\langle a,x^*\rangle \tag{1.12}$$

and

$$f(a+b) = \sum_{x^*\in A^*}\alpha(x^*)\langle b,x^*\rangle\langle a,x^*\rangle \tag{1.13}$$

So f is B-periodic if and only if equations (1.12) and (1.13) are equal; that is, if and only if,

$$\alpha(x^*) = \alpha(x^*)\langle b,x^*\rangle$$

This occurs if and only if $\alpha(x^*) = 0$ whenever $\langle b,x^*\rangle \neq 1$ for all $b \in B$; that is, whenever $x^* \notin B_*$. □

Definition 11 (Evaluation Function).
For any subset $X \subset A$, the *evaluation function* of X is the function $e_X \in \mathcal{L}(A)$ which is equal to 1 on X and vanishes elsewhere.

Corollary 5. As usual, assume A is a group of order N, and B is a subgroup of A having order $M = N/L$; e.g., $A = \mathbb{Z}/N$ and $B = L\mathbb{Z}/N$. If $f \in \mathcal{L}(A)$ is the evaluation function on B and $\alpha \in \mathcal{L}(A^*)$ is the Fourier coefficient set of f, then $L\alpha$ is the evaluation function on the dual subgroup B_*.

Exercise 2. Suppose $f \in \mathcal{L}(\mathbb{Z}/6)$ is the evaluation function on the subgroup $2\mathbb{Z}/6$ with Fourier coefficient set α. Show that 2α is the evaluation function on the dual $3\mathbb{Z}/6$, as implied by corollary 5.

Solution 2. Recall,

$$\mathbb{Z}/6 \simeq \{0,1,2,3,4,5\}, \quad 2\mathbb{Z}/6 \simeq \{0,2,4\}, \quad \text{and} \quad (2\mathbb{Z}/6)_* = 3\mathbb{Z}/6 \simeq \{0,3\}$$

In terms of the foregoing discussion, this is the special case in which $N = 6$, $L = 2$, and $M = 3$. If f is the evaluation function on $2\mathbb{Z}/6$, then

$$f(k) = \begin{cases} 1, & k \in \{0,2,4\} \\ 0, & k \in \{1,3,5\} \end{cases}$$

Therefore, f is $2\mathbb{Z}/6$-periodic and, by Theorem 10, α is $(2\mathbb{Z}/6)_*$-decimated; that is, it is zero off of $3\mathbb{Z}/6$. More precisely,

$$\begin{aligned}\alpha(a) &= \frac{1}{6}(f,a) = \frac{1}{6}\sum_{j=0}^{5} f(j)\overline{\langle j,a\rangle} \\ &= \frac{1}{6}\sum_{j\in\{0,2,4\}} \mathrm{e}^{-\mathrm{i}2\pi\frac{ja}{6}} = \frac{1}{6}\sum_{j=0}^{2} v^{ja} \\ &= \frac{1}{6}\left(v^{0a} + v^{a} + v^{2a}\right), \qquad \text{where } v = \mathrm{e}^{-\mathrm{i}2\pi/3}\end{aligned}$$

For example, $\alpha(1) = 0$, since $v^0 = 1$ and $v^1 = v^2 = -1/2$. We still must show that 2α is 1 on $3\mathbb{Z}/6 = \{0, 3\}$. Obviously, $\alpha(0) = 1/2$. Also, $\alpha(3) = (1 + \mathrm{e}^{-\mathrm{i}2\pi} + \mathrm{e}^{-\mathrm{i}4\pi})/6 = 1/2$, *q.e.d.*

Definition 12 (Periodization).
Define $\mathrm{Per}_B f \in \mathcal{L}(A)$ by the formula

$$\mathrm{Per}_B f(a) = \sum_{x \in B} f(a + x), \qquad a \in A$$

and call $\mathrm{Per}_B f$ the *periodization of f over B*. The function $\mathrm{Per}_B f$ is B-periodic.

Example 8. For $f \in \mathcal{L}(\mathbb{Z}/6)$, the periodization of f over $3\mathbb{Z}/6$ is given by

$$\begin{aligned}\mathrm{Per}_{3\mathbb{Z}/6} f(0) &= f(0)+f(3)\\ \mathrm{Per}_{3\mathbb{Z}/6} f(1) &= f(1)+f(4)\\ \mathrm{Per}_{3\mathbb{Z}/6} f(2) &= f(2)+f(5)\end{aligned}$$

with the remaining values given by periodicity.

In matrix form, the essential values of the periodization of f over $3\mathbb{Z}/6$ can be given by

$$\begin{pmatrix} 1\,0\,0\,1\,0\,0 \\ 0\,1\,0\,0\,1\,0 \\ 0\,0\,1\,0\,0\,1 \end{pmatrix} \mathbf{f}$$

Example 9. For $f \in \mathcal{L}(\mathbb{Z}/N)$, the periodization of f over the subgroup $L\mathbb{Z}/N$, for $N = LM$, is given by

$$\mathrm{Per}_{L\mathbb{Z}/N} f(a) = \sum_{m=0}^{M-1} f(a + mL), \qquad 0 \le l < N$$

In matrix form the periodization of f over $L\mathbb{Z}/N$ is given by

$$(I_L\ I_L\ \cdots I_L)\,\mathbf{f}$$

The matrix pre-multiplying $\mathbf{f}$ consists of M copies of the $L \times L$ identity matrix I_L.

The *Poisson summation formula* describes the Fourier expansion of $\mathrm{Per}_B f$.

Theorem 11 (Poisson Summation Formula). *Suppose A is a finite abelian group of order N, and B is a subgroup of order M, where $N = LM$. If $f \in \mathcal{L}(A)$ has Fourier expansion*

$$f = \sum_{x^* \in A^*} \alpha(a^*) x^*$$

then $\mathrm{Per}_B f$ *has Fourier expansion*

$$\mathrm{Per}_B f = M \sum_{x^* \in B^*} \alpha(a^*) x^*$$

Proof. For $a \in A$,

$$\begin{aligned}\mathrm{Per}_B f(x) &= \sum_{b \in B} \sum_{a^* \in A^*} \alpha(a^*)\langle x, a^*\rangle\langle b, a^*\rangle \\ &= \sum_{a^* \in A^*} \alpha(a^*)\langle x, a^*\rangle \sum_{b \in B} \langle b, a^*\rangle\end{aligned}$$

Recall that $\sum_{b \in B}\langle b, a^*\rangle$ vanishes unless $a^* \in B_*$, in which case the sum equals M. □

Example 10. For $0 \le m < M, 0 \le l < L$,

$$\begin{aligned} f(l+mL) &= \sum_{b^* \in B^*} \alpha(b^*)\langle l, x^*\rangle\langle mL, x^*\rangle \\ &= \sum_{b=0}^{N-1} \alpha(b)\exp(2\pi\mathrm{i}\frac{lb}{N})\exp(2\pi\mathrm{i}\frac{mb}{M}) \end{aligned}$$

Therefore,

$$\mathrm{Per}_{L\mathbb{Z}/N}\, f(l) = \sum_{b=0}^{N-1} \alpha(b)\exp(2\pi\mathrm{i}\frac{lb}{N}) \sum_{m=0}^{M-1} \exp(2\pi\mathrm{i}\frac{mb}{M})$$

Let

$$b = b_1 + b_2 M, \qquad 0 \le b_1 < M, 0 \le b_2 < L$$

so that each b for which $b_1 = 0$ is an element of $M\mathbb{Z}/N = (L\mathbb{Z}/N)_*$. Then,

$$\begin{aligned} \mathrm{Per}_{L\mathbb{Z}/N}\, f(l) &= \sum_{b_1=0}^{M-1}\sum_{b_2=0}^{L-1} \alpha(b_1+b_2M)\exp\{2\pi\mathrm{i}\frac{l(b_1+b_2M)}{N}\} \sum_{m=0}^{M-1} \exp\{2\pi\mathrm{i}\frac{m(b_1+b_2M)}{M}\} \\ &= M\sum_{b_2=0}^{L-1} \alpha(b_2M)\exp(2\pi\mathrm{i}\frac{lb_2}{L}) \end{aligned}$$

The last equality follows from the argument given in the proof of Theorem 11.

By duality, we have identified A with the character group basis of $\mathcal{L}(A^*)$. An element $a \in A$ determines the character of A^*.

$$\Theta(a) : a^* \to \overline{\langle a, a^*\rangle}, \qquad a^* \in A^*$$

Suppose $f \in \mathcal{L}(A)$ has Fourier coefficient set α. For $a^* \in A^*$,

$$\alpha(a^*) = \frac{1}{N}(f, a^*)$$

implying that the expansion of α over the character group basis A of $\mathcal{L}(A^*)$ is given by

$$\alpha = \frac{1}{N}\sum_{a\in A} f(a)a \tag{1.14}$$

(This is Tolimieri's notation, but the point is clearer if we write (1.14) as follows:

$$\alpha = \frac{1}{N}\sum_{a\in A} f(a)\Theta(a)$$

This emphasizes that the role of a here is that of a character.)

For $\alpha \in \mathcal{L}(A^*)$, define $\mathrm{Per}_{B_*}(\alpha) \in \mathcal{L}(A^*)$ by the formula

$$\mathrm{Per}_{B_*}\,\alpha(a^*) = \sum_{x^*\in B_*} \alpha(a^*+x^*), \qquad a^* \in A^*$$

Theorem 12. *If $f \in \mathcal{L}(A)$ has Fourier coefficient set $\alpha \in \mathcal{L}(A^*)$, then*

$$\mathrm{Per}_{B_*}\,\alpha(a^*) = \frac{1}{M}\sum_{x\in B} f(x)x \tag{1.15}$$

(Again, we might express the summand in (1.15) as $f(x)\Theta(x)$.)

Proof. For $a^* \in A^*$,

$$\begin{aligned}\mathrm{Per}_{B_*}\,\alpha(a^*) &= \frac{1}{N}\sum_{x^*\in B_*}\sum_{x\in A} f(x)\Theta(x)(a^*+x^*)\\ &= \frac{1}{N}\sum_{x\in A} f(x)\overline{\langle x,a^*\rangle}\sum_{x^*\in B_*}\overline{\langle x,x^*\rangle}\end{aligned}$$

By Corollary 2, the sum

$$\sum_{x^*\in B_*}\overline{\langle x,x^*\rangle}$$

vanishes unless $x \in B$, in which case the sum is equal to $o(B_*) = N/M$. □

2

Nonabelian Groups and Multivariate DSP

Myoung An[1], Richard Tolimieri, and William DeMeo[2]

[1] psypher@tiac.net
[2] williamdemeo@yahoo.com

2.1 Preliminaries

We first recall some useful notations and definitions. See [1] for elaboration. Throughout, $\mathbb{C}$ denotes complex numbers, G is an arbitrary (nonabelian) group, and $\mathcal{L}(G)$ denotes the collection of complex valued functions on G.

2.1.1 Translations

For $y \in G$, the mapping $\mathsf{T}(y)$ of $\mathcal{L}(G)$ defined by

$$(\mathsf{T}(y)f)(x) = f(y^{-1}x), \qquad x \in G$$

is a linear operator of $\mathcal{L}(G)$ called *left translation by* y. The collection of all left translations of $\mathcal{L}(G)$ is

$$\mathcal{T}(G) = \{\mathsf{T}(y) : y \in G\}.$$

The mapping $\mathsf{T} : G \to \mathcal{T}(G)$ satisfies

$$\mathsf{T}(yz) = \mathsf{T}(y)\mathsf{T}(z), \qquad y, z \in G,$$

$$\mathsf{T}(y^{-1}) = \mathsf{T}(y)^{-1}, \qquad y \in G.$$

These formulae imply that $\mathcal{T}(G)$ is a group under operator composition and T is a group isomorphism of G onto $\mathcal{T}(G)$.

If G has order N, then $y^N = 1$ and

$$\mathsf{T}(y)^N = \mathsf{T}(y^N) = \mathsf{T}(1) = I, \qquad y \in G.$$

For $y \in G$, define $e_y \in \mathcal{L}(G)$ as the function whose value is 1 at y and 0 otherwise. The set $\{e_y : y \in G\}$ is a basis called the *canonical basis* of $\mathcal{L}(G)$. Left translations permute the elements in the canonical basis.

$$\mathsf{T}(y)e_z = e_{yz}, \qquad y, z \in G$$

For $f \in \mathcal{L}(G)$, the mapping $\mathsf{C}(f)$ of $\mathcal{L}(G)$ defined by

$$\mathsf{C}(f) = \sum_{y \in G} f(y)\mathsf{T}(y)$$

is a linear operator of $\mathcal{L}(G)$ called *left convolution by* f. By definition,

$$(\mathsf{C}(f)g)(x) = \sum_{y\in G} f(y)g(y^{-1}x), \qquad g \in \mathcal{L}(G),\ x \in G$$

The collection of all left convolutions of $\mathcal{L}(G)$ is $\mathcal{C}(G) = \{\mathsf{C}(f) : f \in G\}$.

The mapping $\mathsf{C} : \mathcal{L}(G) \to \mathcal{C}(G)$ satisfies

$$\mathsf{C}(f+g) = \mathsf{C}(f) + \mathsf{C}(g), \qquad f, g \in \mathcal{L}(G),$$

$$\mathsf{C}(\alpha f) = \alpha\mathsf{C}(f), \qquad \alpha \in \mathbb{C},\ f \in \mathcal{L}(G).$$

These properties imply that $\mathcal{C}(G)$ is a vector space under addition and scalar multiplication of operators and C is a linear isomorphism from $\mathcal{L}(G)$ onto $\mathcal{C}(G)$. Furthermore, since $\mathsf{T}(y) = \mathsf{C}(e_y),\ y \in G$,

$$\mathcal{T}(G) \subset \mathcal{C}(G)$$

For $f, g \in \mathcal{L}(G)$,

$$\mathsf{C}(f)\mathsf{C}(g) = \sum_{y\in G}\sum_{z\in G} f(y)g(z)\mathsf{T}(y)\mathsf{T}(z) = \sum_{y\in G}\sum_{z\in G} f(y)g(z)\mathsf{T}(yz).$$

By the change of variables $u = yz$,

$$\mathsf{C}(f)\mathsf{C}(g) = \sum_{u\in G}\left(\sum_{y\in G} f(y)g(y^{-1}u)\right)\mathsf{T}(y). \tag{2.1}$$

The inner summation is a function in $\mathcal{L}(G)$, so $\mathcal{C}(G)$ is closed under composition and is a subalgebra of the algebra of all linear operators on $\mathcal{L}(G)$.

For $f, g \in \mathcal{L}(G)$, define the *convolution product*, denoted $*$, as follows:

$$f * g = \mathsf{C}(f)g.$$

By (2.1),

$$\mathsf{C}(f)\mathsf{C}(g) = \mathsf{C}(f * g).$$

The vector space $\mathcal{L}(G)$ paired with the convolution product is an algebra, the *convolution algebra over* G.

Theorem 1. *The mapping* $\mathsf{C} : \mathcal{L}(G) \to \mathcal{C}(G)$ *is an algebra isomorphism of the convolution algebra* $\mathcal{L}(G)$ *onto the algebra of left convolutions of* $\mathcal{L}(G)$.

Note that for the general case in which G is nonabelian, the convolution product is noncommutative. For, by the change of variables $y = z^{-1}x$, it is clear that

$$(g * f)(x) = \sum_{z\in G} g(z)f(z^{-1}x) = \sum_{y\in G} f(y)g(xy^{-1}), \qquad f, g \in \mathcal{L}(G),\ x \in G.$$

Since G is nonabelian, $xy^{-1} \neq y^{-1}x$ for some $x, y \in G$. Thus convolution is not a commuting product.

2.1.2 The Group Algebra $\mathbb{C}G$

The *group algebra* $\mathbb{C}G$ is the space of all formal sums

$$f = \sum_{x\in G} f(x)x, \qquad f(x) \in \mathbb{C}$$

with the following operations:

$$f+g=\sum_{x\in G}(f(x)+g(x))x, \qquad f,g\in\mathbb{C}G,$$

$$\alpha f=\sum_{x\in G}(\alpha f(x))x, \qquad \alpha\in\mathbb{C}, f\in\mathbb{C}G,$$

and

$$fg=\sum_{x\in G}\left(\sum_{y\in G}f(y)g(y^{-1}x)\right)x, \qquad f,g\in\mathbb{C}G.$$

The adds and multiplies inside the parentheses are the familiar operations in $\mathbb{C}$.

$\mathbb{C}G$ is an algebra and the mapping $\Theta:\mathcal{L}(G)\to\mathbb{C}G$ defined by

$$\Theta(f)=\sum_{x\in G}f(x)x, \qquad f\in\mathcal{L}(G)$$

is an algebra isomorphism of the convolution algebra $\mathcal{L}(G)$ onto the group algebra $\mathbb{C}G$. Thus we are free to denote $\Theta(f)$ by f, using context to decide whether f symbolizes a function in $\mathcal{L}(G)$ or a formal sum in $\mathbb{C}G$.

Remark 1. For $y\in G$, the formal sum consisting of a single nonzero term y will be denoted by y. We can view G as a subset of $\mathbb{C}G$. For $x,y\in G$, the product xy in G is equal to the product $xy\in\mathbb{C}G$, and the identity element 1 of G is the identity element in the group algebra $\mathbb{C}G$. Thus G is a basis of the space $\mathbb{C}G$ corresponding to the canonical basis $\{e_y : y\in G\}$ of $\mathcal{L}(G)$.

Remark 2. Keep in mind that multiplication in $\mathbb{C}G$ corresponds to convolution in $\mathcal{L}(G)$ and, as such, does not behave like simple multiplication over $\mathbb{R}$ or $\mathbb{C}$. In particular, the group algebra can have *zero divisors*. This means that there exist $f,g\in\mathbb{C}G$ such that $fg=0$, but $f\neq 0$ and $g\neq 0$.

2.1.3 Left Ideal Decompositions

For $g\in\mathbb{C}G$, the mapping $\mathsf{L}(g)$ of $\mathbb{C}G$ defined by

$$\mathsf{L}(g)f=gf, \qquad f\in\mathbb{C}G$$

is a linear operator on the space $\mathbb{C}G$ called *left multiplication by* g.

Recall remark 1 which identifies $u\in G$ with $u\in\mathbb{C}G$. The mapping L defined above can be thought of as an "overloaded" operator in the sense that it represents translation or convolution depending on its argument. More precisely, left translation by $u\in G$ corresponds to the left multiplication $\mathsf{L}(u)$, whereas left convolution by $g\in\mathcal{L}(G)$ corresponds to the left multiplication $\mathsf{L}(g)$.

Denote by $L(G)$ the collection of all left multiplications $\mathsf{L}(y),\ y\in G$. Since

$$\mathsf{L}(yz)=\mathsf{L}(y)\mathsf{L}(z), \qquad y,z\in G,$$

$$\mathsf{L}(y^{-1})=\mathsf{L}(y)^{-1}, \qquad y\in G,$$

$L(G)$ is a group under operator composition and the mapping $\mathsf{L}:G\to L(G)$ is a group isomorphism from G onto $L(G)$. $L(G)$ is a noncommuting family of linear operators of $\mathbb{C}G$ and, consequently, there cannot exist an $L(G)$-eigenvector basis.

A subspace $\mathcal{V}$ of the space $\mathbb{C}G$ is called a *left ideal* if

$$u\mathcal{V}=\{uf : f\in\mathcal{V}\}\subset\mathcal{V}, \qquad u\in G.$$

A left ideal of $\mathbb{C}G$ corresponds to a subspace of $\mathcal{L}(G)$ invariant under all left translations. If $\mathcal{V}$ is a left ideal, then, by linearity,

$$g\mathcal{V}\subset\mathcal{V}, \qquad g\in\mathbb{C}G.$$

For $g \in \mathbb{C}G$, the set $\mathbb{C}Gg$ defined by

$$\mathbb{C}Gg = \{fg : f \in \mathbb{C}G\}$$

is a left ideal of $\mathbb{C}G$ called *the left ideal generated by* g in $\mathbb{C}G$. $\mathbb{C}Gg = \mathbb{C}G$ if and only if g is an invertible element in $\mathbb{C}G$.

A left ideal $\mathcal{V}$ of $\mathbb{C}G$ is called *irreducible* if the only left ideals of $\mathbb{C}G$ contained in $\mathcal{V}$ are $\{0\}$ and $\mathcal{V}$. The sum of two distinct, irreducible left ideals is always a direct sum ([1], p. 129).

For an abelian group, A, a direct sum decomposition of $\mathbb{C}A$ into ideals is constructed from the character basis A^*. In the abelian case, these irreducible ideals are one-dimensional. Direct sum decompositions of $\mathbb{C}G$ into left ideals play the same role with respect to left translations defined by G (i.e., $\mathsf{L}(u),\ u \in G$) as direct sum decompositions of $\mathbb{C}A$ into ideals play with respect to classical translations. However, *a potential advantage in using nonabelian groups as data indexing sets is the wide scope and complexity of nonabelian group translations.* Projections of data into these left ideals can usually reveal more complicated partitions and structures in the data as compared with the Fourier components in the abelian group case.

2.2 Fourier Analysis on Finite Nonabelian Groups

A *character* of G is a group homomorphism of G into $\mathbb{C}^\times$, where $\mathbb{C}^\times$ denotes the non-zero complex numbers. In other words, a mapping $\varrho : G \to \mathbb{C}^\times$ is a character of G if it satisfies

$$\varrho(xy) = \varrho(x)\varrho(y), \qquad x, y \in G$$

Denote the collection of all characters of G by G^*. Under the usual identification between $\mathcal{L}(G)$ and $\mathbb{C}G$, we can view a character $\varrho \in G^*$ as a formal sum in $\mathbb{C}G$,

$$\varrho = \sum_{x \in G} \varrho(x)x, \tag{2.2}$$

and G^* is a subset of $\mathbb{C}G$.

There always exists at least one character in G^* – namely, the *trivial character* of G, which takes on the value 1 for all $y \in G$.

Characters of nonabelian groups share many properties with abelian group characters.

Theorem 2. *If G has a nontrivial character ϱ, then*

$$\frac{1}{N}\sum_{x \in G} \varrho(x) = 0$$

Proof. For any $y \in G$, by a change of variables,

$$\varrho(y)\sum_{x \in G} \varrho(x) = \sum_{x \in G} \varrho(yx) = \sum_{x \in G} \varrho(x)$$

Therefore, either (a) $\varrho(y) = 1, \forall y \in G$ or (b) $\sum_{x \in G} \varrho(x) = 0$. Since (a) contradicts the assumption that G has a nontrivial character, (b) must be true. □

Theorem 3. *Suppose ϱ is a character of G. If $y \in G$,*

$$y\varrho = \varrho y = \varrho(y^{-1})\varrho$$

Proof. Exercise.
Hint: a straight-forward, two-line proof follows from the standard machine (i.e., write ϱ as the formal sum (2.2) and change variables; [1], p.131).

Corollary 1. Suppose ϱ is a character of G. If $f \in \mathbb{C}G$, then

$$f\varrho = \varrho f = \hat{f}(\varrho)\varrho$$

where $\hat{f}(\varrho) = \sum_{y\in G} f(y)\varrho(y^{-1})$.

Proof. By Theorem 3,

$$f\varrho = \sum_{y\in G} f(y)y\varrho = \left(\sum_{y\in G} f(y)\varrho(y^{-1})\right)\varrho$$

and

$$\varrho f = \sum_{y\in G} f(y)\varrho y = \left(\sum_{y\in G} f(y)\varrho(y^{-1})\right)\varrho.$$

□

2.2.1 Abelian by Abelian Semidirect Products

In order to determine whether a particular group is useful for a DSP application, we must specify exactly how this group represents the data and decide whether the resulting indexing scheme is helpful in some way. For example, the group representation may (or may not) reduce computational complexity, or it may simply allow us to more easily state, understand, or model the given problem.

In this section we describe procedures for specifying and studying a simple class of nonabelian groups that have proven useful in applications – the *abelian by abelian semidirect products*. These are groups of the form $G = A \leftthreetimes B$, where A and B are abelian groups, are they perhaps the simplest generalizations of abelian groups. Not surprisingly, DSP over such groups closely resembles that of abelian groups. However, the resulting processing tools can have vastly different characteristics.

2.2.2 Examples

Example 1 (Semidirect Product, $G_4 = A \leftthreetimes C_2(k_c)$).
(An and Tolimieri [1], p.198)

- $G_4 = A \leftthreetimes C_2(k_c)$, where $A = C_N(x) \times C_N(y)$,

$$c = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

- Group multiplication

$$x^N = y^N = k_c^2 = 1, \quad xy = yx, \quad k_c x^m y^n = x^m y^{m-n} k_c.$$

A function $f \in \mathbb{C}G_4$ is given by the formal sum

$$f = \sum_{l=0}^{1}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} f(x^m y^n k_c^l) x^m y^n k_c^l$$

Specifying an ordering of elements in G_4 by the index array,

$$\begin{pmatrix} 1 \\ k_c \end{pmatrix} \otimes \left(\{x^m y^n\}_{m,n}\right) \begin{pmatrix} 1 & y & \dots & y^{N-1} \\ x & xy & & \\ \vdots & & \ddots & \\ x^{N-1} & & & x^{N-1}y^{N-1} \\ \hline k_c & yk_c & \dots & y^{N-1}k_c \\ xk_c & xyk_c & & \\ \vdots & & \ddots & \\ x^{N-1}k_c & & & x^{N-1}y^{N-1}k_c \end{pmatrix}$$

uniquely specifies a $2N \times N$ matrix representation of f given by the coefficients,

$$\mathbf{f} = \begin{pmatrix} \{f(x^m y^n)\}_{m,n} \\ \hline \{f(x^m y^n k_c)\}_{m,n} \end{pmatrix}$$

Example 2 (Semidirect Product, $G_5 = A \leftthreetimes C_6(k_d)$).
(An and Tolimieri [1], p.204)

- $G_5 = A \leftthreetimes C_6(k_d)$, where $A = C_N(x) \times C_N(y)$, $N = 3 \times 2^K$ for an integer $K \geq 2$, and

$$d = \begin{pmatrix} -1 & 2^K + 1 \\ 2^K - 1 & 2^K \end{pmatrix}$$

- Group multiplication
  Let $M = 2^K$. Since $N = 3 \times 2^K$, this yields the mod N relations

$$M^2 \equiv \begin{cases} M, & \text{if } K \text{ is even,} \\ 2M, & \text{if } K \text{ is odd.} \end{cases}$$

  Consider K even, in which case,

$$M^2 \equiv M, \quad (M+1)^2 \equiv 1, \quad (2M-1)^2 \equiv 1$$

  Therefore,

$$d = \begin{pmatrix} -1 & M+1 \\ M-1 & M \end{pmatrix}, \quad d^2 = \begin{pmatrix} M & M-1 \\ 2M+1 & 2M-1 \end{pmatrix}, \quad d^3 = \begin{pmatrix} M+1 & 2M \\ 0 & 1 \end{pmatrix}$$

$$d^4 = \begin{pmatrix} 2M-1 & 2M+1 \\ M-1 & M \end{pmatrix}, \quad d^5 = \begin{pmatrix} 2M & -1 \\ 2M+1 & 2M-1 \end{pmatrix}, \quad d^6 = 1.$$

  and group multiplication satisfies the following relations:

$$x^N = y^N = k_d^6 = 1, \quad xy = yx, \quad k_d x^m y^n = x^{-m+(M+1)n} y^{(M-1)m+Mn} k_d.$$

A function $f \in \mathbb{C}G_5$ is given by the formal sum

$$f = \sum_{l=0}^{5} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f(x^m y^n k_d^l) x^m y^n k_d^l$$

Part II

Applications

3

Digital Audio Processing: group filters

William DeMeo

williamdemeo@yahoo.com

3.1 Introduction

The translation-invariance of most classical signal processing transforms and filtering operations is largely responsible for their widespread use, and is crucial for efficient algorithmic implementation and interpretation of results. Underlying most digital signal processing (DSP) algorithms is the group $\mathbb{Z}/N$ of integers modulo N, which serves as the data indexing set. Translations are defined using addition modulo N, and basic operations, including convolutions and Fourier expansions, are developed relative to these translations.

DSP on finite abelian groups such as $\mathbb{Z}/N$ is well-understood and has great practical utility. Recently, however, interest in the practical utility of nonabelian groups has grown significantly. Although the theoretical foundations of finite nonabelian groups is well established, application of the theory to DSP has yet to become common-place; cf. the NATO ASI "Computational Non-commutative Algebras," Italy, 2003. Another notable exception is [1], which develops theory and algorithms for indexing data with nonabelian groups, defining translations with a (non-commutative) group multiply operation, and performing typical DSP operations relative to these translations. The work of An and Tolimieri demonstrates that including nonabelian groups among the possible data indexing strategies significantly broadens the range of useful signal processing techniques.

The present work demonstrates the use of nonabelian groups for indexing audio signals, and discusses the insights gained from this approach computational advantages can be gained. For a standard noise reduction problem, we show how the new framework simplifies the algorithm, while simultaneously generalizing it to handle more complex noise reduction tasks. Numerical results are provided along with the Matlab source code for reproducing them.

3.2 Nonabelian group DSP

3.2.1 Two distinctions of consequence

Abelian group DSP can be completely described in terms of a special class of signals called the *characters* of the group. (For $\mathbb{Z}/N$, the characters are simply the exponentials.) Each character of an abelian group represents a one-dimensional translation-invariant subspace, and the set of all characters spans the space of signals indexed by the group; any such signal can be uniquely expanded as a linear combination over the characters.

In contrast, the characters of a nonabelian group G, do not determine a basis for expanding signals indexed by G. However, a basis can be constructed by extending the characters of an abelian subgroup A of G, and then taking certain translations of these extensions. Some of the characters of A cannot be extended to characters of G, but only to proper subgroups of G. This presents some difficulties involving the underlying translation-invariant subspaces, some of which are now multi-dimensional. However, it also

presents opportunities for alternative views of local signal domain information on these translation-invariant subspaces.

The other abelian/nonabelian distinction of primary importance concerns translations defined on the group. In the abelian group case, translations represent simple linear shifts in space or time. When nonabelian groups index the data, however, translations are no longer so narrowly defined.

3.2.2 Basic notations and definitions

This section quickly summarizes the notations, definitions and important facts needed below. (See [1] for elaboration.) Throughout, $\mathbb{C}$ denotes complex numbers, G an arbitrary (nonabelian) group, and $\mathcal{L}(G)$ the collection of complex valued functions on G.

Translations

For $y \in G$, the mapping $\mathsf{T}(y)$ of $\mathcal{L}(G)$ defined by

$$(\mathsf{T}(y)f)(x) = f(y^{-1}x), \quad x \in G$$

is a linear operator of $\mathcal{L}(G)$ called *left translation by* y. The mapping $\mathsf{C}(f)$ of $\mathcal{L}(G)$ defined by

$$\mathsf{C}(f) = \sum_{y \in G} f(y)\mathsf{T}(y), \quad f \in \mathcal{L}(G)$$

is a linear operator of $\mathcal{L}(G)$ called *left convolution by* f. By definition,

$$(\mathsf{C}(f)g)(x) = \sum_{y \in G} f(y)g(y^{-1}x), \quad g \in \mathcal{L}(G),\ x \in G$$

For $f, g \in \mathcal{L}(G)$, the composition

$$f * g = \mathsf{C}(f)g$$

is called the *convolution product*. The vector space $\mathcal{L}(G)$ paired with the convolution product is an algebra, the *convolution algebra over* G.

The group algebra

The *group algebra* $\mathbb{C}G$ is the space of all formal sums

$$f = \sum_{x \in G} f(x)x, \quad f(x) \in \mathbb{C}$$

with the following operations:

$$f + g = \sum_{x \in G} (f(x) + g(x))x, \quad f, g \in \mathbb{C}G,$$

$$\alpha f = \sum_{x \in G} (\alpha f(x))x, \quad \alpha \in \mathbb{C}, f \in \mathbb{C}G,$$

$$fg = \sum_{x \in G} \left(\sum_{y \in G} f(y)g(y^{-1}x) \right) x, \quad f, g \in \mathbb{C}G.$$

For $g \in \mathbb{C}G$, the mapping $\mathsf{L}(g)$ of $\mathbb{C}G$ defined by

$$\mathsf{L}(g)f = gf, \quad f \in \mathbb{C}G$$

is a linear operator on the space $\mathbb{C}G$ called *left multiplication by* g.

Since $y \in G$ can be identified with the formal sum $e_y \in \mathbb{C}G$ consisting of a single nonzero term,

$$yf = \mathsf{L}(e_y)f = \sum_{x \in G} f(y^{-1}x)x \tag{3.1}$$

In relation to translation of $\mathcal{L}(G)$, (3.1) is the $\mathbb{C}G$ analog.

The mapping $\Theta : \mathcal{L}(G) \to \mathbb{C}G$ defined by

$$\Theta(f) = \sum_{x \in G} f(x)x, \quad f \in \mathcal{L}(G)$$

is an algebra isomorphism of the convolution algebra $\mathcal{L}(G)$ onto the group algebra $\mathbb{C}G$. Thus we can identify $\Theta(f)$ with f, using context to decide whether f refers to the function in $\mathcal{L}(G)$ or the formal sum in $\mathbb{C}G$.

An important aspect of the foregoing isomorphism is the correspondence between the translations of the spaces. Translation of $\mathcal{L}(G)$ by $y \in G$ corresponds to left multiplication of $\mathbb{C}G$ by $y \in G$. Convolution of $\mathcal{L}(G)$ by $f \in \mathcal{L}(G)$ corresponds to left multiplication of $\mathbb{C}G$ by $f \in \mathbb{C}G$. To put it symbolically,

$$\begin{aligned} \mathcal{L}(G) &\leftrightarrow \mathbb{C}G \\ \mathsf{T}(y) &\leftrightarrow \mathsf{L}(y) \\ \mathsf{C}(f) &\leftrightarrow \mathsf{L}(f) \end{aligned}$$

Left ideals

A subspace $\mathcal{V}$ of the space $\mathbb{C}G$ is called a *left ideal* if

$$u\mathcal{V} = \{uf : f \in \mathcal{V}\} \subset \mathcal{V}, \quad u \in G.$$

A left ideal of $\mathbb{C}G$ corresponds to a subspace of $\mathcal{L}(G)$ invariant under all left translations. If $\mathcal{V}$ is a left ideal, then, by linearity, $g\mathcal{V} \subset \mathcal{V}$, $g \in \mathbb{C}G$.

For $g \in \mathbb{C}G$, the set $\mathbb{C}Gg$ defined by $\{fg : f \in \mathbb{C}G\}$ is called *the left ideal generated by* g in $\mathbb{C}G$.

A left ideal $\mathcal{V}$ of $\mathbb{C}G$ is called *irreducible* if the only left ideals of $\mathbb{C}G$ contained in $\mathcal{V}$ are $\{0\}$ and $\mathcal{V}$. The sum of two distinct, irreducible left ideals is always a direct sum.

For an abelian group A, the group algebra $\mathbb{C}A$ is decomposed into a direct sum of irreducible ideals. The ideals are one-dimensional translation-invariant subspaces.

Similarly, for a nonabelian group G, the group algebra $\mathbb{C}G$ is decomposed into a direct sum of left ideals. Again, the ideals are translation-invariant subspaces, but some of them must now be multi-dimensional, and herein lies the potential advantage of using nonabelian groups for indexing the data. The left translations are more general and represent a broader class of transformations. Therefore, projections of data into the resulting left ideals can reveal more complicated partitions and structures as compared with the Fourier components in the abelian group case.

3.2.3 Fourier analysis on finite nonabelian groups

A *character* of G is a group homomorphism of G into $\mathbb{C}^\times$, where $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$. In other words, the mapping $\varrho : G \to \mathbb{C}^\times$ is a character of G if it satisfies $\varrho(xy) = \varrho(x)\varrho(y)$, $x, y \in G$.

By the aforementioned identification between $\mathcal{L}(G)$ and $\mathbb{C}G$, a character $\varrho \in G^*$ can be viewed as a formal sum,

$$\varrho = \sum_{x \in G} \varrho(x)x, \tag{3.2}$$

and therefore $G^* \subset \mathbb{C}G$.

Theorem 1. *If G has a nontrivial character ϱ, then*

$$\frac{1}{N}\sum_{x\in G}\varrho(x)=0$$

Proof. For any $y\in G$, by a change of variables,

$$\varrho(y)\sum_{x\in G}\varrho(x)=\sum_{x\in G}\varrho(yx)=\sum_{x\in G}\varrho(x)$$

Therefore, either (a) $\varrho(y)=1,\forall y\in G$, or (b) $\sum\varrho(x)=0$. Since (a) contradicts the hypothesis, (b) must be true. □

Theorem 2. *Suppose ϱ is a character of G. If $y\in G$,*

$$y\varrho=\varrho y=\varrho(y^{-1})\varrho$$

Proof. As above, write ϱ as a formal sum and change variables.

Corollary 1. Suppose ϱ is a character of G. If $f\in\mathbb{C}G$, then

$$f\varrho=\varrho f=\hat{f}(\varrho)\varrho$$

where $\hat{f}(\varrho)=\sum_{y\in G}f(y)\varrho(y^{-1})$.

Proof. By Theorem 2,

$$f\varrho=\sum_{y\in G}f(y)y\varrho=\left(\sum_{y\in G}f(y)\varrho(y^{-1})\right)\varrho$$

Similarly for ϱf, mutatis mutandis. □

3.2.4 Abelian by abelian semidirect products

To determine whether a particular group is useful for a DSP application, we must specify exactly how this group represents the data. The group representation may reduce computational complexity, or it may simply make it easier to state, understand, or model a given problem.

In this section we describe procedures for specifying and studying a simple class of nonabelian groups that have proven useful in applications – the *abelian by abelian semidirect products*. These are perhaps the simplest extension of abelian groups and DSP over such groups closely resembles that over abelian groups. However, the resulting processing tools can have vastly different characteristics.

Semidirect product

Let G be a finite group of order N, K a subgroup of G, and H a normal subgroup of G. If $G=HK$ and $H\cap K=\{1\}$, then we say that G is the *internal semidirect product* $G=H\leftthreetimes K$. It can be shown that $G=H\leftthreetimes K$ if and only if every $x\in G$ has a *unique representation* of the form $x=yz,\ y\in H,z\in K$.

The mapping $\Psi:K\to Aut(H)$, defined by

$$\Psi_z(x)=zxz^{-1},\quad z\in K,x\in H$$

is a group homomorphism. Define the binary composition in G in terms of Ψ as follows:

$$x_1x_2=(y_1z_1)(y_2z_2)=y_1\Psi_{z_1}(y_2)z_1z_2,$$

$$y_1,y_2\in H,\ z_1,z_2\in K.$$

If $G=H\leftthreetimes K$ and K is a normal subgroup of G, then G is the "usual" direct product (i.e., the cartesian product $H\times K$ along with component-wise multiplication). What is new in the semidirect product is the possibility that K acts nontrivially on H.

Cyclic groups

Denote by $C_N(x)$ the cyclic group of order N having generator x, $\{x^n : 0 \leq n < N\}$, and define binary composition by $x^m x^n = x^{m+n}$, $0 \leq m, n < N$, where $m + n$ is addition modulo N.

For an integer L, $0 \leq L < N$, denote by $gp_N(x^L)$ the subgroup generated by x^L in $C_N(x)$. If L divides N, then

$$gp_N(x^L) = \{x^m L : 0 \leq m < M\}, \quad LM = N$$

and $gp_N(x^L)$ is a cyclic group of order M.

Group of units

Multiplication modulo N is a ring product on the group of integers $\mathbb{Z}/N$. An element $m \in \mathbb{Z}/N$ is called a unit if there exists an $n \in \mathbb{Z}/N$ such that $mn = 1$. The set $U(N)$ of all units in $\mathbb{Z}/N$ is a group with respect to multiplication modulo N, and is called the *unit group* of $\mathbb{Z}/N$.

The unit group $U(N)$ can be characterized as the set of all integers $0 < m < N$ such that m and N are relatively prime. For example, $U(8) = \{1, 3, 5, 7\}$.

Semidirect product example

The mapping $\Psi : U(N) \to Aut(C_N(x))$ is a group isomorphism. Under this identification, we can form $C_N(x) \leftthreetimes K$ for any subgroup K of $U(N)$. A typical point in $C_N(x) \leftthreetimes K$ is denoted (x^n, u), $0 \leq n < N$, $u \in K$ with multiplication given by

$$(x^m, u)(x^n, v) = (x^{m+un}, uv), \quad 0 \leq m, n < N,\ u, v \in K$$

where $m + un$ is taken modulo N.

4

Digital Image Processing: atmospheric turbulence, anisoplanatism

Paul Billings[1] and William DeMeo[2]

[1] Textron Systems Corporation, Hawaii Operations; pbilling@systems.textron.com
[2] Textron Systems Corporation, Hawaii Operations; williamdemeo@yahoo.com

4.1 Preliminaries

This chapter describes an approach to *multi-frame blind deconvolution* (MFBD), which is a method for reducing image degradation due to atmospheric turbulence. Section 4.2 is based on Paul Billing's notes on this topic [2], but the present exposition is set in the group theoretic framework.[3] This formalism not only aides general understanding of the MFBD algorithm, but also yields some simplifications that are crucial for developing fast algorithms, even as the imaging model grows in complexity. Sections 4.2.3 and 4.1.2 support these claims by setting up the space-variant blur problem[4] over finite groups, and then demonstrating how this set-up renders the model computationally tractable.

4.1.1 Abelian Groups

We first collect some required notations and definitions. Recall that, by isomorphism, we take $\mathbb{Z}/N$ to denote the group $\{0, 1, \ldots, N-1\}$, and

$$\mathbb{Z}/N_1 \times \mathbb{Z}/N_2 = \{(x_1, x_2) : 0 \leq x_1 < N_1,\, 0 \leq x_2 < N_2\}$$

is a direct product group, each element of which represents a 2-dimensional spatial coordinate.

For a finite set A of order N, let $\mathcal{L}(A)$ denote the vector space of all complex valued functions on A with addition and scalar multiplication defined by

$$(f+g)(a) = f(a) + g(a), \qquad f, g \in \mathcal{L}(A),\, a \in A,$$

$$(\alpha f)(a) = \alpha f(a), \qquad f, g \in \mathcal{L}(A),\, \alpha \in \mathbb{C},\, a \in A.$$

The space $\mathcal{L}(A)$ has dimension N.

Quotient Groups and Periodic Functions

Suppose A is an abelian group of order N and B is a subgroup of A. The vector space $\mathcal{L}(A/B)$ is the subspace of all functions in $\mathcal{L}(A)$ which are constant on the B-cosets in A,

$$f(a+B) = f(a), \qquad f \in \mathcal{L}(A),\, a \in A.$$

Such functions are called *B-periodic*. If B has order M, where $LM = N$, and $\{a_l : 0 \leq l < L\}$ is a complete system of B-coset representatives in A, then functions in $\mathcal{L}(A/B)$ are completely determined by their values on the set $\{a_l : 0 \leq l < L\}$.

[3] See also Tolimieri and An [7], for a lucid treatment of the mathematical prerequisites.
[4] "Space-variant blur" (or *spatially varying blur*, or *anisoplanatism*), is discussed in [5], among other places.

$$f(a_l + b) = f(a_l), \qquad 0 \le l < L,\, b \in B.$$

Since the collection of B-cosets, $A/B = \{a_l + B : 0 \le l < L\}$, is a partition of A, it determines a direct sum decomposition,

$$\mathcal{L}(A) = \bigoplus_{l=0}^{L} \mathcal{L}(a_l + B).$$

Such decompositions underlie many divide-and-conquer strategies; e.g., the fast Fourier transform (FFT).

The following is a generic example of a periodic function in two variables. Such functions are useful in image processing applications.

Example 1. Identify $\mathcal{L}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)$ with the space of all $N_1 \times N_2$ complex matrices $\mathcal{M}(N_1, N_2)$. In particular, represent $f \in \mathcal{L}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)$ by the $N_1 \times N_2$ matrix

$$Mat(f) = [f(n_1, n_2)]_{\substack{0 \le n_1 < N_1 \\ 0 \le n_2 < N_2}}$$

Suppose $A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ and $B = L_1\mathbb{Z}/N_1 \times L_2\mathbb{Z}/N_2$, where $N_1 = L_1M_1$ and $N_2 = L_2M_2$. The set

$$\{(l_1, l_2) : 0 \le l_1 < L_1,\, 0 \le l_2 < L_2\}$$

is a complete system of B-coset representatives in A. Therefore, the $(L_1\mathbb{Z}/N_1 \times L_2\mathbb{Z}/N_2)$-periodic functions in $\mathcal{L}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)$ can be identified with the collection of all complex $N_1 \times N_2$ matrices of the form

$$\begin{pmatrix} M & \cdots & M \\ \vdots & \ddots & \vdots \\ M & \cdots & M \end{pmatrix}, \quad M = [f(l_1, l_2)]_{\substack{0 \le l_1 < L_1 \\ 0 \le l_2 < L_2}}$$

having block matrix structure with M_1M_2 identical copies of the matrix $M \in \mathbb{C}^{L_1 \times L_2}$.

In image processing applications, we often work on the abelian group $A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$. The following definitions specialize translation and convolution operations for this group. Unless otherwise noted, we assume A denotes the direct product group $\mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ throughout this section.

Definition 1 (Translation by $y \in A$).
For $A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ and $y \in A$, the mapping $\mathsf{T}(y)$ of $\mathcal{L}(A)$ takes $f \in \mathcal{L}(A)$ to a function $\mathsf{T}(y)f \in \mathcal{L}(A)$ having the following values

$$(\mathsf{T}(y)f)(x) = f(x - y) = f(x_1 - y_1, x_2 - y_2), \qquad x \in A$$

$\mathsf{T}(y)$ is a linear operator on $\mathcal{L}(A)$ called *translation by y*.

Definition 2 (Convolution by $g \in \mathcal{L}(A)$).
Let $A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ and $g \in \mathcal{L}(A)$. The mapping $C(g)$ of $\mathcal{L}(A)$ defined by

$$C(g)f = \sum_{y \in A} g(y)\mathsf{T}(y)f \qquad f \in \mathcal{L}(A)$$

is a linear operator of $\mathcal{L}(A)$ called *convolution by g*. Evaluated at a point,

$$(C(g)f)(x) = \sum_{y_1 \in \mathbb{Z}/N_1} \sum_{y_2 \in \mathbb{Z}/N_2} g(y_1, y_2) f(x_1 - y_1, x_2 - y_2), \qquad x = (x_1, x_2) \in A$$

Finally, we define the characters for our special case. First recall the general case. The group of all N^{th} roots of unity is the set

$$U_N = \{\mathrm{e}^{\mathrm{i}2\pi\frac{n}{N}} : 0 \le n < N\}$$

If τ is a character of the group A, then τ is a group homomorphism from A into U_N, where N is the least common multiple of the orders of the elements in A.

Definition 3 (Characters of $\mathbb{Z}/N_1 \times \mathbb{Z}/N_2$).
Let $A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$. For each $a = (a_1, a_2) \in A$, the mapping $\tau_a : A \to U_{[N_1,N_2]}$ defined by

$$\tau_a(x) = \mathrm{e}^{\mathrm{i}2\pi \frac{x_1 a_1}{N_1}} \mathrm{e}^{\mathrm{i}2\pi \frac{x_2 a_2}{N_2}}, \qquad x = (x_1, x_2) \in A.$$

is a *character of* A.

Any isomorphism from a finite abelian group A onto its character group A^* is called a *presentation* of A. The presentation given by definition 3 is called the *standard presentation* of A. We use this presentation exclusively, and for the remainder assume the identifications

$$A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2 = (\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)^* = A^*.$$

Cyclic Groups

Consider the abelian group A with elements in the set

$$C_N(x) \times C_N(y) = \{x^m y^n : 0 \le m, n < N\},$$

where $C_N(x) = \{x^m : 0 \le m < N\}$ denotes a cyclic group of order N with generator x and multiplication satisfying $x^m x^n = x^{m+n \bmod N}$ and $x^N = 1$. A typical point $x^m y^n, 0 \le m, n < N$, is subject to the relations $x^N = 1 = y^N$ and $xy = yx$. Implicit are the standard identifications, such as $x^1 y^0 = (x, 1) = x$, $x^0 y^1 = (1, y) = y$, which cause no ambiguity.

Let B be the subgroup of A defined by

$$B = gp_N(x^L) \times gp_N(y^L) = \{x^{pL} y^{qL} : 0 \le p, q < M\}, \quad \text{where } LM = N,$$

Thus, B is a direct product of cyclic groups of order M. The *factor group* A/B is given by

$$A/B = \{x^j y^k B : 0 \le j, k < L\}$$

Each element $x^j y^k B \in A/B$ is a direct product of cyclic subgroups of A. The element $x^j y^k B$ is called the *B-coset of A with representative $x^j y^k$*. The elements of a particular coset are called *equivalent modulo B*. A complete set of *B-coset representatives in A* is

$$H = \{x^j y^k : 0 \le j, k < L\} = C_L(x) \times C_L(y)$$

Thus, H is a direct product of cyclic groups of order L. Furthermore, any element $a \in A$ can be uniquely written

$$a = hb, \qquad h \in H, b \in B$$

where h specifies that a belongs to the coset hB, and b identifies a within that coset. We give concrete examples of B-cosets for a few special cases.

Example 2. First, let $N = 8$, $M = 2$ and $L = 4$. Then,

$$A = C_8(x) \times C_8(y) = \{x^m y^n : 0 \le m, n < 8\} \tag{4.1}$$

and

$$B = gp_8(x^4) \times gp_8(y^4) = \{x^{p4} y^{q4} : 0 \le p, q < 2\}$$

In the following figure, the numbers denote exponents mn on the elements $x^m y^n \in A$ in (4.1).

m \ n	0	1	2	3	4	5	6	7
0	**00**	01	02	03	**04**	05	06	07
1	10	11	12	13	14	15	16	17
2	20	21	22	23	24	25	26	27
3	30	31	32	33	34	35	36	37
4	**40**	41	42	43	**44**	45	46	47
5	50	51	52	53	54	55	56	57
6	60	61	62	63	64	65	66	67
7	70	71	72	73	74	75	76	77

The boldface exponents comprise the B-coset with representative x^0y^0; that is,

$$x^0y^0B = B = \begin{bmatrix} 00 & 04 \\ 40 & 44 \end{bmatrix}$$

The B-coset with representative x^1y^0 is

$$x^1y^0B = xB = \begin{bmatrix} 10 & 14 \\ 50 & 54 \end{bmatrix}$$

A few more examples are the sets

$$yB = \begin{bmatrix} 01 & 05 \\ 41 & 45 \end{bmatrix}, \quad xyB = \begin{bmatrix} 11 & 15 \\ 51 & 55 \end{bmatrix}$$

$$x^2B = \begin{bmatrix} 20 & 24 \\ 60 & 64 \end{bmatrix}, \quad x^2yB = \begin{bmatrix} 21 & 25 \\ 61 & 65 \end{bmatrix}$$

which are the B-cosets with representatives x^0y^1, x^1y^1, x^2y^0, and x^2y^1, respectively.

4.1.2 Nonabelian Groups

This section describes a few tools from nonabelian group theory that can be usefully applied to the space-varying blur problem. Some of the material merely re-iterates or emphasizes previously stated facts about groups and factor groups so that this section is less dependent on those preceding it.

Action Group

Denote by $\mathsf{GL}(2, \mathbb{Z}/N)$ the set of all 2×2 invertible matrices with coefficients in $\mathbb{Z}/N$. Suppose $c \in \mathsf{GL}(2, \mathbb{Z}/N)$ is such that $c^M = 1$ – the identity in $\mathsf{GL}(2, \mathbb{Z}/N)$ – and consider the *action group* K_c with elements

$$C_M(k_c) = \{k_c^m : 0 \le m < M\}, \quad c = \begin{pmatrix} c_0 & c_1 \\ c_2 & c_3 \end{pmatrix} \in \mathsf{GL}(2, \mathbb{Z}/N).$$

The semi-direct product of H and K_c has elements

$$H \leftthreetimes K_c = \{x^j y^k k_c^m : 0 \le j, k < L, 0 \le m < M\}$$

and binary composition satisfying the following relations:

$$x^L = y^L = k_c^M = 1,$$

$$x^{-1} = x^{L-1}, \quad y^{-1} = y^{L-1}, \quad k_c^{-1} = k_c^{M-1},$$

$$k_c x^j y^k = x^{c_0 j + c_1 k} y^{c_2 j + c_3 k} k_c.$$

where the summands in the exponents are modulo $|H| = L$.

Rotation

Let $A = C_N(x) \times C_N(y)$ with binary composition satisfying

$$(x^m y^j)(x^n y^k) = x^{m+n \bmod N} y^{j+k \bmod N},$$

$$x^N = y^N = 1, \quad x^{-1} = x^{N-1}, \quad y^{-1} = y^{N-1}.$$

Consider the action group K_c, $c \in \mathsf{GL}(2, \mathbb{Z}/N)$, with $c^M = 1$. The group generated by k_c is the cyclic group of order M with elements $C_M(k_c) = \{k_c^m : 0 \le m < M\}$. Now suppose

$$c(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Example 3 (Rotation by $\pi/2$). The action group $K_{c(\pi/2)}$ has

$$c(\pi/2) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Since $c^4(\pi/2)$ is the identity, the group has order $M = 4$.

The semi-direct product $A \leftthreetimes K_{c(\theta)}$ has elements $\{x^j y^k k_{c(\theta)}^m : 0 \le j, k < N, 0 \le m < M\}$, and binary composition satisfying

$$x^N = y^N = k_{c(\theta)}^M = 1,$$

$$x^{-1} = x^{N-1}, \quad y^{-1} = y^{N-1}, \quad k_{c(\theta)}^{-1} = k_{c(\theta)}^{M-1},$$

and

$$k_{c(\theta)} x^j y^k = x^{j\cos\theta - k\sin\theta} y^{j\sin\theta + k\cos\theta} k_{c(\theta)},$$

Additive operations in the exponents are modulo $|A| = N$.

4.1.3 Ideal Image Model

Assume the object of interest is $f \in \mathbb{C}G$, where $G = H \leftthreetimes K_c$ is as defined above. Then,

$$f = \sum_{a \in G} f(a) a = \sum_{j,k} \sum_{m} f(x^j y^k k_c^m) x^j y^k k_c^m$$

Take the function $g \in \mathbb{C}H$ to be some imperfect representation of f. For example, a blurry image is formed by mixing the object with a point spread function $s \in \mathbb{C}H$.

$$g = fs = \left(\sum_{j,k} \sum_{m} f(x^j y^k k_c^m) x^j y^k k_c^m \right) s \tag{4.2}$$

This is a superposition in which the function s is left-multiplied by elements of G. Recall that left-multiplication of $s \in \mathbb{C}H$ generalizes translation of $s \in \mathcal{L}(H)$. Thus, fs in $\mathbb{C}G$ is a generalized convolution of f and s in $\mathcal{L}(G)$. The following further elucidates.

Consider a single left-multiply of $s \in \mathbb{C}H$ by $hk \in H \leftthreetimes K_c$.

$$\begin{aligned} \mathsf{L}(hk)s &= x^j y^k k_c^m s \\ &= \sum_{p,q} s(x^p y^q)(x^j y^k k_c^m) x^p y^q \\ &= \sum_{p,q} s((x^j y^k k_c^m)^{-1} x^p y^q) x^p y^q \end{aligned}$$

So, by (4.2), the function $g \in \mathbb{C}H$ is defined by the values

$$g(x^p y^q) = \sum_{j,k} \sum_m f(x^j y^k k_c^m) s((x^j y^k k_c^m)^{-1} x^p y^q) \tag{4.3}$$

The important thing to take from equation (4.3) is that left-multiplying by an element of a nonabelian group is a linear transformation that is quite general and includes classical translation as a special case. When the operand is an element of the abelian group H, left-multiply is a simple shift operator; e.g.,

$$\mathsf{L}(x^j y^k) s(x^p y^q) = s((x^j y^k)^{-1} x^p y^q) = s(x^{p-j} y^{q-k}) \tag{4.4}$$

where, as usual, arithmetic in exponents is performed modulo $|H|$. A superposition of such left-multiplies of $\mathbb{C}H$ is equivalent to classical convolution of $\mathcal{L}(H)$,

$$(C(f)s)(x) = \sum_{y \in H} f(y) y s(x) = \sum_{y \in H} f(y) s(y^{-1} x)$$

However, when the operand is an element of the action group K_c, a much richer class of transformations is available, and this class can be constructed to include rotations, scale changes, and other revealing transformations. This is an important and powerful consideration since it allows us to apply our existing fast algorithms for the standard convolution to operations that are more general than simple spatial or temporal shifts.

4.2 Noisy Image Model

This section is based on Paul Billing's notes [2], but the present exposition is set in the algebraic framework described above.

Let f denote the object of interest and define

$$\begin{aligned} g &= \text{ ideal (noiseless) image}, & d &= \text{ detected image}, \\ \hat{h} &= \text{ coherent OTF}, & \hat{s} &= \text{ incoherent OTF}. \end{aligned}$$

and assume $f, g, d \in \mathbb{C}G$, and $\hat{h} \in \mathbb{C}G^*$; that is the domain of $\hat{h}$ is G^*, the character group of G, which can be interpreted as frequency space. The incoherent OTF, $\hat{s} \in \mathbb{C}G^*$, is defined as the autocorrelation of the coherent OTF. We denote this autocorrelation by $\hat{h} \star \hat{h}$ and define it as follows:

$$\hat{s}(\tau) = (\hat{h} \star \hat{h})(\tau) = \sum_{\lambda \in G^*} \hat{h}(\lambda) \overline{\hat{h}(\tau^{-1}\lambda)}, \quad \tau \in G^*.$$

As an element of the group algebra $\mathbb{C}G^*$, $\hat{s}$ is represented as the formal sum

$$\hat{s} = \sum_{\tau \in G^*} \hat{s}(\tau) \tau.$$

The coefficients $\hat{s}(\tau)$ in this basis can be derived using the following identities:

$$\hat{h}(\lambda) = \sum_{x \in G} h(x) \lambda(x^{-1}), \quad \overline{\tau(x)} = \tau^{-1}(x) = \tau(x^{-1}), \quad (\tau^{-1}\lambda)(x) = \tau^{-1}(x)\lambda(x)$$

The coefficient of $\hat{s}$ at $\tau \in G^*$ is now readily derived as follows:

$$
\begin{aligned}
\hat{s}(\tau) = (\hat{h} \star \hat{h})(\tau) &= \sum_{\lambda \in G^*} \hat{h}(\lambda)\overline{\hat{h}(\tau^{-1}\lambda)} \\
&= \sum_{\lambda \in G^*} \left(\sum_{y \in G} h(y)\lambda(y^{-1}) \right) \overline{\left(\sum_{x \in G} h(x)(\tau^{-1}\lambda)(x^{-1}) \right)} \\
&= \sum_{x \in G} \sum_{y \in G} h(y)\overline{h(x)}\,\overline{\tau^{-1}(x^{-1})} \sum_{\lambda \in G^*} \lambda(y^{-1})\overline{\lambda(x^{-1})} \\
&= \sum_{x \in G} \sum_{y \in G} h(y)\overline{h(x)}\tau(x^{-1}) \sum_{\lambda \in G^*} \lambda(y^{-1}x) \\
&= |G| \sum_{x \in G} h(x)\overline{h(x)}\tau(x^{-1})
\end{aligned}
$$

The last equality holds by the following character formula:

$$
\sum_{\lambda \in G^*} \lambda(y^{-1}x) = \begin{cases} |G|, & x = y, \\ 0, & x \neq y. \end{cases}
$$

Poisson/Extended Object Model (old) This section is based on Paul Billing's notes [2], but the present exposition is set in the algebraic framework described above.

Define

$$\begin{aligned} f &= \text{ object} & s &= \text{ incoherent PSF} \\ g &= \text{ noiseless image} & d &= \text{ detected image} \end{aligned}$$

and assume $f, s, g, d \in \mathcal{L}(A)$. By an image "ensemble" we mean a set of K images, or frames, indexed by $k \in K = \{0, 1, 2, \ldots, K-1\}$. For the k^{th} frame in the ensemble, suppose[5]

$$g_k = C(f)s_k \tag{4.5}$$

In (4.5) $C(f)s_k$ denotes convolution by f of $s_k \in \mathcal{L}(A)$. More explicitly, for any point $x = (x_1, x_2)$ in the image plane, we have

$$g_k(x) = \sum_{y_1 \in \mathbb{Z}/N_1} \sum_{y_2 \in \mathbb{Z}/N_2} f(y_1, y_2) s_k(x_1 - y_1, x_2 - y_2), \qquad k \in K$$

Recall the Poisson distribution with mean λ has a probability mass function (pmf) given by:

$$p(x|\lambda) = \frac{\lambda^x \mathrm{e}^{-\lambda}}{x!}$$

Suppose that, at each point $x \in A$, the value $d_k(x)$ is a realization of a Poisson process with mean $g_k(x)$. Then

$$p\left(d_k(x)|g_k(x)\right) = \frac{g_k(x)^{d_k(x)} \mathrm{e}^{-g_k(x)}}{d_k(x)!}$$

Denote the set of all pixels in images belonging to the detected ensemble by

$$\mathbf{d} = \{d_k(x)\}_{\substack{k \in K \\ x \in A}}$$

Similarly, denote the set of all pixels in images from the noiseless ensemble by

$$\mathbf{g} = \{g_k(x)\}_{\substack{k \in K \\ x \in A}}$$

Assuming independence, the joint pmf for pixels in the detected ensemble is

$$p\left(\mathbf{d}|\mathbf{g}\right) = \prod_k \prod_x p\left(d_k(x)|g_k(x)\right) = \prod_k \prod_x \frac{g_k(x)^{d_k(x)}\ \mathrm{e}^{-g_k(x)}}{d_k(x)!} \tag{4.6}$$

Here – and throughout unless otherwise noted – k ranges over K and x ranges over A.

Given that we observed $\mathbf{d}$, we want to find the values $\mathbf{g}$ which yield a joint density, $p\left(\cdot|\mathbf{g}\right)$, from which the observed data set, $\mathbf{d}$, is the most likely outcome. In other words, we seek a maximum likelihood estimate (MLE) of $\mathbf{g}$.

Given the data $\mathbf{d}$, we maximize the right hand side of (4.6) over all values of $\mathbf{g}$. Therefore, it makes sense to write (4.6) as a function of $\mathbf{g}$ given $\mathbf{d}$:

$$L\left(\mathbf{g}|\mathbf{d}\right) = \prod_k \prod_x \frac{g_k(x)^{d_k(x)}\ \mathrm{e}^{-g_k(x)}}{d_k(x)!}$$

[5] We previously used $g_k = t \cdot C(f)s_k$, where $t \in \mathcal{L}(A)$ is a truncation window, and $\cdot$ denotes point-wise multiplication. However, since the function t does not influence any of the derivations, it is notationally cleaner to postpone inclusion of the truncation factor.

This is typically called the likelihood function. It's easier, and equivalent, to maximize the natural log of the likelihood function, which is:

$$\ell\left(\mathbf{g}|\mathbf{d}\right) = \sum_{k=0}^{K-1} \sum_{x} d_k(x) \log g_k(x) - g_k(x) - \log d_k(x)! \tag{4.7}$$

The last term on the right hand side is constant, so it can be ignored for the purposes of maximizing the so called "log likelihood function," ℓ.

Recall from (4.5), g_k is defined as $C(f)s_k$. Here, t is known and we are trying to estimate f and s_k. As a consequence, we arrive at an estimate of g_k. Thus, instead of maximizing the likelihood over values of g_k, we maximize over values of f and s_k simultaneously.

4.2.1 Gradient with respect to f

Maximizing the log likelihood function over possible values of the object leads to the MLE of the object, which we denote by $\tilde{f}$. To derive this estimate we find that value of f at which the derivative, with respect to $f(z)$, equals 0 for each $z \in A$.

$$\begin{aligned}
\frac{\partial \ell\left(\mathbf{g}|\mathbf{d}\right)}{\partial f(z)} &= \sum_{k=0}^{K-1} \sum_{x \in A} \frac{\partial}{\partial f(z)} \left[d_k(x) \log g_k(x) - g_k(x) \right] \\
&= \sum_{k=0}^{K-1} \sum_{x \in A} \frac{\partial}{\partial g_k(x)} \left[d_k(x) \log g_k(x) - g_k(x) \right] \frac{\partial g_k(x)}{\partial f(z)} \\
&= \sum_{k=0}^{K-1} \sum_{x \in A} \left[\frac{d_k(x)}{g_k(x)} - 1 \right] \frac{\partial g_k(x)}{\partial f(z)}
\end{aligned}$$

and

$$\frac{\partial g_k}{\partial f(z)} = \frac{\partial}{\partial f(z)} \sum_{y \in A} f(y) \mathsf{T}(y) s_k = \mathsf{T}(z) s_k$$

Therefore,

$$\frac{\partial \ell\left(\mathbf{g}|\mathbf{d}\right)}{\partial f(z)} = \sum_{k=0}^{K-1} \sum_{x \in A} \left[\frac{d_k(x)}{g_k(x)} - 1 \right] \mathsf{T}(z) s_k(x) \tag{4.8}$$

In simple optimization problems, the gradient is a function of the variable over which we optimize. In the present case, the gradient is a function of g_k, which in turn is a function of f via equation (4.5). Therefore, given a set of K psf estimates $\{s_k\}_{0 \le k < K}$, we seek a vector $\tilde{f}$, which, through (4.5), minimizes the magnitude of (4.8) for each $z \in A$. That is, we must find the vector $\tilde{f}$ which minimizes

$$\left| \sum_{k=0}^{K-1} \sum_{x \in A} \left[\frac{d_k(x)}{g_k(x)} - 1 \right] \mathsf{T}(z) s_k(x) \right|, \qquad z \in A$$

4.2.2 Gradient with respect to s_k

In order to maximize the likelihood function with respect to the PSF, s_k, we could proceed by estimating s_k directly, but it is often useful to write s_k as a function of phase-aberration, and then estimate the form of this function.

PSF Phase Parameterization

There are unknown phase errors across the aperture that distort the image. We represent this distortion with a *phase-aberration function*, Φ_k, which may vary with k; in other words, the phase-aberration may be different for each frame.

In deriving the MLE of s_k using the second approach mentioned above, we first expand the phase-aberration function in a suitable basis. (Often a *Zernike basis* is used.) If we denote the set of basis functions by $\{\varphi_j\}_{0\le j<J}$, then the phase-aberration for the k^{th} frame is expanded as follows:

$$\Phi_k = \sum_{j=0}^{J-1} \alpha_{kj}\varphi_j$$

We usually take the basis functions φ_j as given and not depending on k. On the other hand, the coefficients α_{kj} must be estimated for each $k \in K$.

Remark 1. If the choice of basis was physically motivated, it can help our intuition to think of α_{kj} as the projection of Φ_k onto the basis function φ_j. In fact, when $\{\varphi_j\}_{0\le j<J}$ is an orthonormal basis, the coefficients really are projections. In that case, we often write the expansion as follows:

$$\Phi_k = \sum_{j=0}^{J-1} \langle \Phi_k, \varphi_j\rangle \varphi_j$$

The incoherent PSF for the k^{th} frame is

$$s_k = |h_k|^2 = h_k h_k^* \tag{4.9}$$

where h_k is the coherent PSF, which is the inverse Fourier transform of the coherent transfer function, or *pupil function*.

Above we defined a character $\tau_a : A \to U_N$ by the mapping

$$\tau_a(x) = \mathrm{e}^{\mathrm{i}2\pi \frac{x_1 a_1}{N_1}} \mathrm{e}^{\mathrm{i}2\pi \frac{x_2 a_2}{N_2}}, \qquad x \in \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$$

Using this notation, we state some standard Fourier transform relations on which our model depends.

$$h_k = \frac{1}{N_1 N_2}\mathcal{F}^{-1} H_k = \frac{1}{N_1 N_2} \sum_{u\in A^*} H_k(u)\tau_u \tag{4.10}$$

The index set of the summation in (4.10) is A^*, the group of characters of A. However, as noted above, we identify the elements of A^* with those of A by the standard presentation, and simply write the character group as $A^* = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$.

At any $x \in A$, (4.10) evaluates to

$$h_k(x) = \frac{1}{N_1 N_2} \sum_{u\in A^*} H_k(u)\, \mathrm{e}^{\mathrm{i}2\pi \frac{x_1 u_1}{N_1}} \mathrm{e}^{\mathrm{i}2\pi \frac{x_2 u_2}{N_2}}$$

The transfer function is the Fourier transform of the point spread function.

$$H_k = \mathcal{F} h_k = \sum_{x\in A} h_k(x)\tau_x^* \tag{4.11}$$

At any $u \in A^*$, (4.11) evaluates to

$$H_k(u) = \sum_{x\in A} h_k(x)\, \mathrm{e}^{-\mathrm{i}2\pi \frac{x_1 u_1}{N_1}} \mathrm{e}^{-\mathrm{i}2\pi \frac{x_2 u_2}{N_2}}$$

Again, the summations are over the group $A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ and its character group, A^*. By isomorphism, the character group is the same index set, $A^* = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$.

The pupil function is complex-valued, and we take as its complex argument the phase-aberration function. That is,

$$H_k(u) = |H_k(u)|\, \mathrm{e}^{\mathrm{i}\Phi_k(u)} = |H_k(u)|\, \mathrm{e}^{\mathrm{i}\sum_j \alpha_{kj}\varphi_j(u)}, \qquad u \in A \tag{4.12}$$

From equations (4.9)–(4.12) we can write s_k as a function of α_{kj}.

$$\begin{aligned} s_k &= h_k h_k^* \\ &= \left(\frac{1}{N_1 N_2}\sum_{u\in A^*} H_k(u)\tau_u\right)\left(\frac{1}{N_1 N_2}\sum_{v\in A^*} H_k^*(v)\tau_v^*\right) \\ &= \frac{1}{(N_1 N_2)^2}\sum_{u\in A^*}\sum_{v\in A^*} H_k(u)H_k^*(v)\tau_u\tau_{-v} \end{aligned} \tag{4.13}$$

Remark 2. At this point, we diverge slightly to make an observation which simplifies equation (4.13). Such a simplification may only hold when working in the *group algebra* $\mathbb{C}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)$ (see [1]). Therefore, we make this point to indicate the potential benefits of switching to the $\mathbb{C}(\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)$ framework for future analysis. Thereafter, we resume without benefit of this simplifying assumption.
Assume the following character formula:

$$\tau_u\tau_v = \begin{cases} o(A)\tau_u, & \tau_v = \tau_u \\ 0, & \text{otherwise} \end{cases} \tag{4.14}$$

Expression (4.14) holds for $\tau_u, \tau_v \in \mathbb{C}A$. Under this assumption, (4.13) simplifies to

$$s_k = \frac{1}{N_1 N_2}\sum_{u\in A^*} H_k(u)H_k^*(-u)\tau_u$$

Returning to expression (4.13), since $H_k^*(u) = |H_k(u)|\, \mathrm{e}^{-\mathrm{i}\varphi_k(u)}$, we have

$$s_k = \frac{1}{(N_1 N_2)^2}\sum_{u\in A^*}\sum_{v\in A^*} |H_k(u)||H_k(v)|\, \mathrm{e}^{i\sum_j \alpha_{kj}[\varphi_j(u)-\varphi_j(v)]}\tau_{u-v} \tag{4.15}$$

Now that we have written s_k as a function of $\{\alpha_{kj}\}_{0\le j<J}$, we can maximize the likelihood function (4.7) with respect to α_{kj}. For each index pair (a, b), we have

$$\frac{\partial \ell\,(\mathbf{g}|\mathbf{d})}{\partial \alpha_{ab}} = \sum_{k=0}^{K-1}\sum_{x\in A}\left[\frac{d_k(x)}{g_k(x)} - 1\right]\frac{\partial g_k(x)}{\partial \alpha_{ab}}$$

and

$$\frac{\partial g_k}{\partial \alpha_{ab}} = \sum_{y\in A}\frac{\partial g_k}{\partial s_k(y)}\frac{\partial s_k(y)}{\partial \alpha_{ab}} \tag{4.16}$$

By the definition of g_k and commutativity of convolution,

$$g_k = C(f)s_k = C(s_k)f = \sum_{x\in A} s_k(x)\mathsf{T}(x)f$$

Therefore,

$$\frac{\partial g_k}{\partial s_k(y)} = \mathsf{T}(y)f \tag{4.17}$$

It remains only to compute the partial of s_k with respect α_{ab}. From (4.15),

$$\frac{\partial s_k}{\partial \alpha_{ab}} = \frac{1}{(N_1 N_2)^2} \sum_{u \in A^*} \sum_{v \in A^*} |H_k(u)||H_k(v)|\, i[\varphi_b(u) - \varphi_b(v)]\delta(k-a)\, \mathrm{e}^{i \sum_j \alpha_{kj}[\varphi_j(u) - \varphi_j(v)]} \tau_{u-v}$$

Over the domain $(u, v) \in A^* \times A^*$, define

$$\mathcal{H}_{ab}(u,v) = \frac{1}{(N_1 N_2)^2} |H_a(u)||H_a(v)|\, [\varphi_b(u) - \varphi_b(v)]\, \mathrm{e}^{i \sum_j \alpha_{aj}[\varphi_j(u) - \varphi_j(v)] + i\frac{\pi}{2}}$$

Then,

$$\frac{\partial s_a}{\partial \alpha_{ab}} = \sum_{u \in A^*} \sum_{v \in A^*} \mathcal{H}_{ab}(u,v) \tau_{u-v}, \qquad \text{and} \qquad \frac{\partial s_k}{\partial \alpha_{ab}} = 0, \quad k \neq a \tag{4.18}$$

Inserting equations (4.17) and (4.18) into equation (4.16) yields

$$\begin{aligned} \frac{\partial g_a}{\partial \alpha_{ab}} &= \sum_{y \in A} \sum_{u \in A^*} \sum_{v \in A^*} \mathcal{H}_{ab}(u,v) \tau_{u-v}(y) \mathsf{T}(y) f \\ &= \sum_{u \in A^*} \sum_{v \in A^*} \mathcal{H}_{ab}(u,v) C(\tau_{u-v}) f \\ &= \sum_{u \in A^*} \sum_{v \in A^*} \mathcal{H}_{ab}(u,v) \hat{f}(u-v) \tau_{u-v} \end{aligned}$$

where

$$\hat{f}(u) = \sum_{y \in A} f(y) \tau_u(y)$$

denotes the Fourier coefficient of f at u.

We can now formally express the gradient of ℓ with respect to α_{ab} as follows:

$$\frac{\partial \ell\,(\mathbf{g}|\mathbf{d})}{\partial \alpha_{ab}} = \sum_{x \in A} \sum_{u \in A^*} \sum_{v \in A^*} \left[\frac{d_a(x)}{g_a(x)} - 1 \right] \mathcal{H}_{ab}(u,v) \hat{f}(u-v) \tau_{u-v}(x) \tag{4.19}$$

If we let $r_a = \frac{d_a}{g_a} - 1$, another Fourier coefficient in expression (4.19) becomes apparent.

$$\hat{r}_a(u-v) = \sum_{x \in A} r_a(x) \tau_{u-v}(x)$$

from which

$$\frac{\partial \ell\,(\mathbf{g}|\mathbf{d})}{\partial \alpha_{ab}} = \sum_{u \in A^*} \sum_{v \in A^*} \mathcal{H}_{ab}(u,v) \hat{f}(u-v) \hat{r}_a(u-v) \tag{4.20}$$

Equation (4.20) provides a formal expression of the gradient. However, we require an expression which better facilitates algorithmic implementation. Note that

$$s_a = h_a h_a^* = \mathrm{Re}[h_a]^2 + \mathrm{Im}[h_a]^2$$

is real valued. Therefore, we can write equation (4.15) as

$$s_k(x) = \mathrm{Re}[s_k(x)] = \frac{1}{(N_1 N_2)^2} \sum_{u,v} |H_k(u)||H_k(v)| \cos\left(\sum_j \alpha_{kj}[\varphi_j(u) - \varphi_j(v)] + 2\pi \frac{(u_1 - v_1)x_1}{N_1} + 2\pi \frac{(u_2 - v_2)x_2}{N_2} \right)$$

from which, we have

$$\frac{\partial s_a(y)}{\partial \alpha_{ab}} = \frac{1}{(N_1 N_2)^2} \sum_{u,v} |H_a(u)||H_a(v)|\, [\varphi_b(u) - \varphi_b(v)] \sin\left(\sum_j \alpha_{aj}[\varphi_j(u) - \varphi_j(v)] + 2\pi \frac{(u_1 - v_1)y_1}{N_1} + 2\pi \frac{(u_2 - v_2)y_2}{N_2} \right) \tag{4.21}$$

Substituting (4.21) for $\sum_{u,v} \mathcal{H}_{ab}(u,v)\tau_{u-v}$ in equation (4.19) yields

$$\frac{\partial \ell\,(\mathbf{g}|\mathbf{d})}{\partial \alpha_{ab}} = \frac{1}{(N_1N_2)^2} \sum_{x,y} \sum_{u,v} \left[\frac{d_a(x)}{g_a(x)} - 1\right] t(x) f(x-y) |H_a(u)||H_a(v)| \left[\varphi_b(u) - \varphi_b(v)\right] \tag{4.22}$$

$$\times \sin\left(\sum_j \alpha_{aj}[\varphi_j(u) - \varphi_j(v)] + 2\pi \frac{(u_1 - v_1)y_1}{N_1} + 2\pi \frac{(u_2 - v_2)y_2}{N_2}\right)$$

By (4.22) we see that one way to find a parameter set $\{\alpha_{aj}\}_j$ yielding a zero gradient is to find one which satisfies

$$\sum_{j=0}^{J-1} \alpha_{aj}[\varphi_j(u) - \varphi_j(v)] + 2\pi \frac{(u_1 - v_1)y_1}{N_1} + 2\pi \frac{(u_2 - v_2)y_2}{N_2} = n\pi, \qquad n \in \mathbb{Z} \tag{4.23}$$

4.2.3 The Spatially Varying Case

The foregoing assumes that the point spread function is spatially invariant. We now state the problem for the more general spatially varying case. Focusing at first on a single image frame, we need not involve the frame index k until later.

As before, identify $A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2$ with its character group by the standard presentation.

$$A = \mathbb{Z}/N_1 \times \mathbb{Z}/N_2 = (\mathbb{Z}/N_1 \times \mathbb{Z}/N_2)^* = A^*$$

Let $y \in A$ be a spatial coordinate in the object plane. Then we model the noiseless image as the superposition,

$$g = \sum_{y \in A} f(y) s_y \tag{4.24}$$

where g, f, $s_y \in \mathcal{L}(A)$. Note that in this model our point spread function varies with y.

Assume $s_y = h_y h_y^*$, where

$$h_y = \frac{1}{N_1N_2} \mathcal{F}^{-1} H_y = \frac{1}{N_1N_2} \sum_{u \in A^*} H_y(u) \tau_u \tag{4.25}$$

At any point x in the focal plane, (4.25) evaluates to

$$h_y(x) = \frac{1}{N_1N_2} \sum_{u \in A^*} H_y(u)\, \mathrm{e}^{\mathrm{i}2\pi \frac{x_1u_1}{N_1}} \mathrm{e}^{\mathrm{i}2\pi \frac{x_2u_2}{N_2}}, \qquad x \in A$$

Thus far, the equations defining the psf's and transfer function are identical to the invariant case, with the frame index k replaced by a spatial coordinate in the object frame. However, a basic difference arises in the functional form of the coherent transfer function.

$$H_y(u) = |H_y(u)|\, \mathrm{e}^{\mathrm{i}\Psi_y(u)} \tag{4.26}$$

where

$$\Psi_y(u) = \sum_{\ell=0}^{L-1} \Phi_\ell(\beta_\ell y + (1 - \beta_\ell)u)$$

The index ℓ represents locations along the elevation axis. Thus the total phase-aberration for the pair u, y takes contributions from phase-aberration functions Φ_ℓ at various altitudes, evaluated at a point along the line connecting u and y.

If we denote the set of basis functions by $\{\varphi_j\}_{0 \le j < J}$, then the phase-aberration for a point ℓ along the elevation axis is represented by the following expansion:

$$\Phi_\ell = \sum_{j=0}^{J-1} \alpha_{\ell j} \varphi_j$$

The basis functions φ_j usually do not depend on ℓ, whereas the coefficients $\alpha_{\ell j}$ must be estimated for each $\ell \in \{0, 1, \dots, L-1\}$. Finally, we have

$$\Psi_y(u) = \sum_{\ell=0}^{L-1} \sum_{j=0}^{J-1} \alpha_{\ell j} \varphi_j(\beta_\ell y + (1-\beta_\ell)u) \tag{4.27}$$

Remark 3. During implementation, we should be able to exploit periodicities of the basis functions. This would facilitate, among other things, estimation of the coefficients $\alpha_{\ell j}$ in equation (4.27). In particular, suppose φ_j is B-periodic; that is,

$$\varphi_j(x) = \varphi_j(x+y), \qquad y \in B \tag{4.28}$$

If the order of B divides L, say $M = L/o(B)$, then

$$\sum_{\ell=0}^{L-1} \alpha_{\ell j} \varphi_j(\beta_\ell y + (1-\beta_\ell)u) = \sum_{m=0}^{M-1} \left(\sum_{y \in B} \alpha_{(m+y)j} \right) \varphi_j(\beta_m y + (1-\beta_m)u) \tag{4.29}$$

We see that, without the benefit of periodicity, there are L coefficients to compute. Exploiting B-periodicity, we compute only $M = L/o(B)$ coefficients.

N.B. the form of (4.29) is probably wrong due to the form of the operand of φ_j. To correct for this, we need to represent the periodicities of φ_j along the line connecting y and u, rather than in the simple form given by (4.28).

4.2.4 Periodic PSF

Since s is ultimately a function of the basis $\{\varphi_j\}$, it should be possible to exploit periodicities in φ_j when working with s. This would greatly facilitate algorithmic implementation of the space-variant model.

Suppose the group A has order N. Let B be a subgroup of A with order $L = o(B)$, where $L = N/M$. Later we address the problem of writing s_y as a periodic function. For now, assume s_y is B-periodic in y; that is, for each $x \in A$,

$$s_a(x) = s_{a+b}(x), \qquad a \in A,\, b \in B$$

Then, by (4.24),

$$\begin{aligned} g &= \sum_{y \in A} f(y) s_y \\ &= \sum_{a \in A/B} \sum_{b \in B} f(a+b) s_{a+b} \\ &= \sum_{a \in A/B} \left(\sum_{b \in B} f(a+b) \right) s_a \end{aligned}$$

This reduces the order of spatial variance to $o(A/B) = M$.

In order to find periodicities in s_y, consider its functional form. By 4.25

$$\begin{aligned} s_y &= h_y h_y^* \\ &= \frac{1}{(N_1 N_2)^2} \left(\sum_{u \in A^*} H_y(u) \tau_u \right) \left(\sum_{v \in A^*} H_y^*(v) \tau_v \right) \\ &= \frac{1}{(N_1 N_2)^2} \sum_{u \in A^*} \sum_{v \in A^*} H_y(u) H_y^*(v) \tau_{u-v} \end{aligned}$$

Thus, we must locate periodicities in

$$H_y(u) = |H_y(u)|\, \mathrm{e}^{\mathrm{i}\Psi_y(u)}$$

where

$$\Psi_y(u) = \sum_{\ell=0}^{L-1} \Phi_\ell(\beta_\ell y + (1-\beta_\ell)u)$$

4.2.5 Status Reports

[2004.02.01]

Recap of previous status:
I had derived the necessary formalism for incorporating non-abelian group methods into our MFBD framework. Such methods enable the current framework to incorporate some basic models of anisoplanatism in a computationally tractable way.

Recall, one of our goals:

> – Better synthesis of old and new theory –
> The new theory is very general and, as yet, little has been done to incorporate the finer details and special properties of our application. The old theory is very specialized and highly adapted to our application. We need to bridge this gap.

I attained this goal by deriving concrete expressions for some basic transformations (translation, rotation, scaling) of a point spread function (PSF). The significance of deriving these in the nonabelian group context is that all operations can be expressed as left-multiplications, which greatly reduces computational complexity. In fact, this reduction is essential to the model's implementation; without them, the anisoplanatic component renders the model computationally infeasible.

Currently, I am implementing the aforementioned PSF trans- formations in the Octave language, and experimenting on simulated data. Kyle Cooper has helped me access some sample FM6 data, on which I will test the new methods.

[2003.12.01]

Last Month

1. Organized the collection of 200 Matlab routines acquired at the NATO meeting, adding basic documentation to aid in sorting through and understanding it all.
2. Studied these subroutines to better understand the correspondence between the underlying theory and its Matlab implementation.
3. Took steps to identify a subset of the NATO code which is most relevant and useful for the anisoplanatism problem.

Current/Future Goals

1. Achieve better understanding of the software, and learn how best to apply it under the imaging conditions of greatest interest to us. This can be accomplished through experimentation, simulations, and comparisons of results. By relating these experiences to the underlying theory, we can provide a coherent explanation of the results.
2. Identify and implement the necessary extensions, modifications, or specializations of the code in light of our immediate concerns, which include the following: new methods and algorithms, if they are to be included among existing tools, must either demonstrate clear performance advantages, or provide complementary information.

Over the next two weeks, I will perform experiments with the new methods on simulated and real data. I hope to demonstrate that performance gains are attainable under anisoplanatic conditions, and that complementary information is provided irrespective of the degree of spatial variability.

4.2.6 Future Work for Anisoplanatism R & D

Overview.

What follows is a brief, general outline of our immediate concerns and objectives, after which appears more specific details describing how we plan to address our concerns and carry out our objectives.

1. **Better synthesis of old and new theory**
   The new theory is very general and, as yet, little has been done to incorporate the finer details and special properties of our application. The old theory is very specialized and highly adapted to our application. We need to bridge this gap.
2. **Experimentation, Simulation, Real Data**
   Demonstrate new capabilities and characterize performance using both simulated and real data.

Immediate Objectives and Action Items.

By breaking down the items in the foregoing section, we construct a list of more concrete goals, stated in terms of observable actions with measurable results.

1. **Better synthesis of old and new theory**
   a) Review the following references:
      - Paul Billing's handwritten notes
      - section 4.2 above
      - section 4.2.3 above
   b) Consider special conditions (e.g., assumptions made, constraints imposed) of the existing model as stated in Paul's notes and section 4.2.
   c) Determine which of the special conditions are not reflected in the new paradigm and its corresponding model as described in section 4.2.3.
   d) Consider how the special conditions could be incorporated into the new model and, if incorporated, what resulting improvements or advantages could we expect.
2. **Experimentation, Simulation, Real Data**
   a) Develop Matlab prototypes for experimenting with and applying the new theory and methods.
      - Work with and learn from existing MFBD matlab code in CVS repository.
      - Work with and learn from matlab code acquired at NATO meeting.
      - Build upon and extend these assets by developing prototype code for implementing the model of section 4.2.3 above.
   b) Experiment with various models of psf variation and consider what transformations provide the best model for anisoplanatism.

Part III

Back Matter

A

Symbols and Acronyms

A.1 List of Symbols

$\mathbb{C}$	the complex numbers
$\mathbb{R}$	the real numbers
$\mathbb{Q}$	the rational numbers
$\mathbb{Z}$	the integers
$\mathbb{N}$	the natural numbers
$\mathbb{F}$	an arbitrary field (e.g. $\mathbb{R}$)
$\mathbb{F}^n$	the set of n-tuples with elements in $\mathbb{F}$ (e.g., $\mathbb{R}^n$)
Re	real part of a complex number
Im	imaginary part of a complex number
$\langle G, + \rangle$	a group with elements in G and binary addition operator $+$
$\langle G, \cdot \rangle$	a group with elements in G and binary multiplication operator $\cdot$
$\langle G, * \rangle$	a group with elements in G and binary operator $*$
G_*	the dual of the group G
$\langle R, +, \cdot \rangle$	a ring with elements in R and binary add/multiply operators $+/\cdot$
$\mathbb{Z}/N$	the finite additive group of order N (equivalent to $\mathbb{Z}/N\mathbb{Z} \simeq \mathbb{Z}_N \simeq \{0, 1, \ldots, N-1\}$)
$\mathbb{Z}_2$	the finite additive group of order 2
$\mathcal{H}$	a Hilbert space
$\mathcal{B}$	a Banach space
$\mathcal{L}$(A)	the space of complex valued functions on A
$\mathcal{L}$(G)	the space of complex valued functions on G
$\mathbb{C}A$	the group algebra of A
$\mathbb{C}G$	the group algebra of G
$\mathcal{B}$	a basis
T	an operator (e.g. translation)
I	the identity operator
W	the Wigner-Ville transform
Per	a periodizing operator
$\mathcal{N}$	the null space of an operator, or the normal distribution, depending on context
$\mathcal{R}$	the range of an operator, or a binary relation, depending on context
$\mathcal{F}$	the Fourier transform
$\mathcal{F}^{-1}$	the inverse Fourier transform
Poi	the Poisson distribution
$\overline{x}$	the sample mean of x
e.g.	for example (latin: *exempli gratia*)
i.e.	that is (latin: *id est*)
etc.	etcetera (latin: *et cetera*)

A.2 List of Acronyms

DSP digital signal processing
FFT fast Fourier transform
FOV field of view
OTF optical transfer function
PSF point spread function
SNR signal to noise ratio

B

Mathematical Tools

This chapter sets down many of the notations, tools, and modeling conventions that we find useful for developing and applying mathematical and statistical theory.

B.1 Vector Spaces

The main source of most material in this section is Horn and Johnson, *Matrix Analysis*, [4].

B.1.1 Subspace, Span, Basis

Definition 1 (Scalar Field). Underlying a vector space is the *field*, or set of scalars, on which addition and multiplication occurs. For our purposes, that underlying field will usually be the real numbers $\mathbb{R}$ or the complex numbers $\mathbb{C}$ under the usual addition and multiplication operations, but it could be the rational numbers $\mathbb{Q}$, the integers modulo a specified number $\mathbb{Z}/N\mathbb{Z}$, or some other field. When the field is unspecified, we use the symbol $\mathbb{F}$. To qualify as a field, a set of scalars must be closed under two specified binary operations ("addition" and "multiplication"); both operations must be associative and commutative and have an identity element in the set; inverses must exist in the set for all elements under the addition operation and for all elements except the additive identity (0) under the multiplication operation; the multiplication operation must also be distributive over the addition operation. (See section B.2.5 for a more formal definition of a field.)

Definition 2 (Vector Space). A *vector space* $\mathcal{V}$ over a field $\mathbb{F}$ is a set $\mathcal{V}$ of objects ("vectors") which is closed under a binary operation ("addition") and such that an identity (0) and additive inverses exist in the set $\mathcal{V}$. The set is also closed under an associative and commutative operation of left multiplication of vectors by elements of the scalar field, with the following properties: for all $\alpha, \beta \in \mathbb{F}$ and all $\mathbf{x}, \mathbf{y} \in \mathcal{V}$:

1. $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$
2. $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$
3. $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$
4. $e\mathbf{x} = \mathbf{x}$, for the multiplicative identity $e \in \mathbb{F}$

For a given field $\mathbb{F}$, the set $\mathbb{F}^n$ of n-tuples (n a positive integer) with components from $\mathbb{F}$ forms a vector space over $\mathbb{F}$ under the obvious operations (component-wise addition in $\mathbb{F}^n$). The special cases with which we are usually concerned are the n-dimensional vector spaces over the real and complex fields, $\mathbb{R}^n$ and $\mathbb{C}^n$, respectively.

Definition 3 (Subspace). A *subspace* $\mathcal{U}$ of a vector space $\mathcal{V}$ is a subset of $\mathcal{V}$ that is, by itself, a vector space over the same scalar field.

Usually a subspace of a vector space $\mathcal{V}$ is defined by some relation that identifies particular elements of $\mathcal{V}$ in such a way that the resulting set is closed under the addition operation – for example, the elements of $\mathbb{R}^3$ with the last component 0, denoted $\{(\alpha, \beta, 0) : \alpha, \beta \in \mathbb{R}\}$, is a subspace of $\mathbb{R}^3$. It is in this regard that we find it useful to think of the resulting set as a subspace rather than as a vector space in its own right. In any event, the intersection of two subspaces is again a subspace.

Definition 4 (Span, n.). If S is a subset of a vector space $\mathcal{V}$, denoted $S \subset \mathcal{V}$, then the *span* of S is the set

$$\text{span}(S) = \{\alpha_0\mathbf{v}_0 + \cdots + \alpha_{k-1}\mathbf{v}_{k-1} \mid \alpha_0, \ldots, \alpha_{k-1} \in \mathbb{F}, \mathbf{v}_0, \ldots, \mathbf{v}_{k-1} \in S, k = 1, 2, \ldots\}$$

Notice that $\text{span}(S)$ is always a subspace even if S is not a subspace.

The subset $S \subset \mathcal{V}$ is said to "span" the vector space $\mathcal{V}$ if $\text{span}(S) = \mathcal{V}$. In other words, span is also a verb, defined as follows:

Definition 5 (Span, v.). The subset S *spans* $\mathcal{V}$ if every element of $\mathcal{V}$ may be written as a linear combination of elements of S.

For example, the set

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \tag{B.1}$$

spans $\mathbb{R}^3$.

Definition 6 (Basis). A a linearly independent set which spans a vector space $\mathcal{V}$ is called a *basis* for $\mathcal{V}$.

The set (B.1) does not comprise a basis since the vectors are linearly dependent. However, three vectors from that set do form a basis.

The basis for $\mathbb{R}^n$ given by

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\} \tag{B.2}$$

is called the "elementary basis," often denoted $\{\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_{n-1}\}$.

Bases are highly non-unique, but are very efficient in that each element of $\mathcal{V}$ can be represented uniquely in terms of the basis. In this sense a basis is a minimal spanning set. For example, if $\mathcal{B}_0 = \{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ is a basis for $\mathcal{V}$, and $\mathbf{x} \in \mathcal{V}$, then $\mathbf{x}$ can be written uniquely as the linear combination

$$\mathbf{x} = \alpha_0 \mathbf{v}_0 + \alpha_1 \mathbf{v}_1 + \cdots + \alpha_{n-1} \mathbf{v}_{n-1} \tag{B.3}$$

Exercise 1. Show that the set of coefficients describing $\mathbf{x}$ in terms of $\mathcal{B}_0$ is unique.

Bases are important because we can represent all the objects of our vector space in terms of the elements of a given basis for the space. Thus, the choice of basis will determine how the objects under investigation appear, and an analysis that appears complex or intractable in one basis might appear simple when represented in another basis.

B.1.2 Linear Transformation, Similar Matrices, Change of Basis

Throughout, $\mathcal{M}(m, n, \mathbb{F})$ denotes the set of all $m \times n$ matrices with elements in $\mathbb{F}$, and $\mathcal{M}(n, \mathbb{F})$ the set of $n \times n$ matrices over $\mathbb{F}$. When the field is of no consequence, or is clear from the context, these are abreviated to $\mathcal{M}(n, m)$ and $\mathcal{M}(n)$. The subset of all invertible elements, or *units*, in $\mathcal{M}(n, \mathbb{F})$ is denoted $GL(n, \mathbb{F})$.

Let $\mathcal{U}$ be an n-dimensional vector space, $\mathcal{V}$ an m-dimensional vector space over the same scalar field $\mathbb{F}$, and let $\mathcal{B}_0$ and $\mathcal{B}_1$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively. Recall that, as a linear combination of the basis elements, $\mathbf{u}_i \in \mathcal{B}_0$,

$$\mathbf{x} = \alpha_0 \mathbf{u}_0 + \alpha_1 \mathbf{u}_1 + \cdots + \alpha_{n-1} \mathbf{u}_{n-1} = \sum_{k=0}^{n-1} \alpha_k \mathbf{u}_k \tag{B.4}$$

We denote the vector of coefficients by $[\mathbf{x}]_{\mathcal{B}_0} = (\alpha_0, \alpha_1, \ldots, \alpha_{n-1}) \in \mathbb{F}^n$. Similarly, any vector $\mathbf{y} \in \mathcal{V}$ represented in terms of $\mathcal{B}_1$ is denoted $[\mathbf{y}]_{\mathcal{B}_1}$.

Definition 7. The vector $[\mathbf{x}]_{\mathcal{B}_0}$ is called the $\mathcal{B}_0$-*basis representation of* $\mathbf{x}$.

It is sometimes helpful to let $\mathbf{x}(\mathbf{u}_k)$ denote the coefficient α_k, by which equation (B.4) becomes

$$\mathbf{x} = \sum_{k=0}^{n-1} \mathbf{x}(\mathbf{u}_k) \mathbf{u}_k$$

The scalar quantities $\mathbf{x}(\mathbf{u}_k) \in \mathbb{F}$ are called the coordinates of $\mathbf{x}$ with respect to $\mathcal{B}_0$.

Linear Transformation

Definition 8 (Linear Transformation). A mapping T from a vector space $\mathcal{U}$ to a vector space $\mathcal{V}$ is a *linear transformation* if

$$\mathsf{T}(\alpha \mathbf{x}_0 + \beta \mathbf{x}_1) = \alpha \mathsf{T}\mathbf{x}_0 + \beta \mathsf{T}\mathbf{x}_1 \tag{B.5}$$

for all vectors $\mathbf{x}_0, \mathbf{x}_1 \in \mathcal{U}$ and all scalars $\alpha, \beta \in \mathbb{F}$.

A linear transformation $\mathsf{T} : \mathcal{U} \to \mathcal{V}$ always has a matrix representation, A, which satisfies the following correspondence:

$$\mathsf{T}\mathbf{x} = \mathbf{y} \quad \Leftrightarrow \quad A[\mathbf{x}]_{\mathcal{B}_0} = [\mathbf{y}]_{\mathcal{B}_1}$$

The matrix A represents the linear transformation T relative to the bases $\mathcal{B}_0$ and $\mathcal{B}_1$. When we study a particular matrix, we are studying a linear transformation relative to a particular choice of bases. Thus we see that, when changing the basis used to represent objects, though the appearance of the objects may change, some fundamental structure is preserved.

Similar Matrices

The notion that two matrices have a common "fundamental structure" is an important idea in linear algebra, and it has a precise definition.

Definition 9 (Similar Matrices). A matrix $B \in M(n)$ is *similar* to a matrix $A \in M(n)$ if there exists a nonsingular matrix $S \in M(n)$ such that

$$B = S^{-1}AS$$

The transformation $A \to S^{-1}AS$ is called a *similarity transformation* and S is called the *similarity matrix*. The relation "A and B are similar" is sometimes abbreviated $A \sim B$. Similarity is an *equivalence relation* on $M(n)$. Equivalence relations are discussed further in section B.2.1. The important point here is that similarity, like any equivalence relation, partitions the set $M(n)$ into disjoint equivalence classes. Each equivalence class is the set of all matrices in $M(n)$ that are similar to a given matrix, a representative of the class.

To summarize the foregoing, the statement that two distinct matrices are similar means that they are merely different basis representations of the same linear transformation – they have a common "fundamental structure." Such structure partitions the set of all matrices into classes, thus reducing it to a subset consisting of one "canonical" matrix per class.

Change of Basis

Let $\mathcal{U}$ be an n-dimensional vector space over the field $\mathbb{F}$. Let $\mathcal{B}_0 = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}\}$ be a basis for $\mathcal{U}$. As remarked above,

$$\mathbf{x} = \sum_{k=0}^{n-1} \mathbf{x}(\mathbf{u}_k)\mathbf{u}_k, \qquad \mathbf{x} \in \mathcal{U}$$

The linear mapping

$$\mathbf{x} \to [\mathbf{x}]_{\mathcal{B}_0} = \begin{pmatrix} \mathbf{x}(\mathbf{u}_0) \\ \mathbf{x}(\mathbf{u}_1) \\ \vdots \\ \mathbf{x}(\mathbf{u}_{n-1}) \end{pmatrix}$$

from $\mathcal{U}$ to $\mathbb{F}^n$, is well defined, one-to-one, and onto.

Given a linear transformation, $\mathsf{T} : \mathcal{U} \to \mathcal{U}$, the action of T on $\mathbf{x} \in \mathcal{U}$ is determined once we know $[\mathbf{x}]_{\mathcal{B}_0}$ and the n vectors $\mathsf{T}\mathbf{u}_0, \mathsf{T}\mathbf{u}_1, \ldots, \mathsf{T}\mathbf{u}_{n-1}$. This is clear by the linearity – equation (B.5) – according to which,

$$\mathsf{T}\mathbf{x} = \mathsf{T}\left(\sum_{k=0}^{n-1} \mathbf{x}(\mathbf{u}_k)\mathbf{u}_k\right) = \sum_{k=0}^{n-1} \mathbf{x}(\mathbf{u}_k)\mathsf{T}\mathbf{u}_k$$

Let $\mathcal{B}_1 = \{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ be another basis for $\mathcal{U}$. For $j = 0, 1, \ldots, n-1$, let

$$[\mathsf{T}\mathbf{u}_j]_{\mathcal{B}_1} = \begin{pmatrix} \mathsf{T}\mathbf{u}_j(\mathbf{v}_0) \\ \mathsf{T}\mathbf{u}_j(\mathbf{v}_1) \\ \vdots \\ \mathsf{T}\mathbf{u}_j(\mathbf{v}_{n-1}) \end{pmatrix} = \begin{pmatrix} t_{0j} \\ t_{1j} \\ \vdots \\ t_{n-1j} \end{pmatrix}$$

denote the $\mathcal{B}_1$-basis representation of $\mathsf{T}\mathbf{u}_j$. Then, for $\mathbf{x} \in \mathcal{U}$,

$$
\begin{aligned}
[\mathsf{T}\mathbf{x}]_{\mathcal{B}_1} &= \left[\sum_{j=0}^{n-1} \mathbf{x}(\mathbf{u}_j)\mathsf{T}\mathbf{u}_j\right]_{\mathcal{B}_1} = \sum_{j=0}^{n-1} \mathbf{x}(\mathbf{u}_j)\,[\mathsf{T}\mathbf{u}_j]_{\mathcal{B}_1} \\
&= \sum_{j=0}^{n-1} \mathbf{x}(\mathbf{u}_j) \begin{pmatrix} t_{0j} \\ t_{1j} \\ \vdots \\ t_{n-1j} \end{pmatrix} = \begin{pmatrix} t_{00} & \cdots & t_{0n-1} \\ \vdots & \ddots & \vdots \\ t_{n-10} & \cdots & t_{n-1n-1} \end{pmatrix} \begin{pmatrix} \mathbf{x}(\mathbf{u}_0) \\ \vdots \\ \mathbf{x}(\mathbf{u}_{n-1}) \end{pmatrix}
\end{aligned}
\tag{B.6}
$$

The n-by-n array $\{t_{ij}\}_{0 \le i,j < n}$ in equation (B.6) depends on T and on the choice of bases $\mathcal{B}_0$ and $\mathcal{B}_1$, but it does not depend on $\mathbf{x}$. The following definition generalizes this discussion slightly by letting T take $\mathbf{x} \in \mathcal{U}$ into another vector space $\mathcal{V}$.

Definition 10. Given a basis $\mathcal{B}_0 = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}\}$ for $\mathcal{U}$, a basis $\mathcal{B}_1 = \{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{m-1}\}$ for $\mathcal{V}$, and a linear transformation $\mathsf{T} : \mathcal{U} \to \mathcal{V}$, the $\mathcal{B}_0$-$\mathcal{B}_1$*-basis representation of* T is given by the $m \times n$ array

$$
{}_{\mathcal{B}_1}[\mathsf{T}]_{\mathcal{B}_0} = \big([\mathsf{T}\mathbf{u}_0]_{\mathcal{B}_1} \cdots [\mathsf{T}\mathbf{u}_{m-1}]_{\mathcal{B}_1}\big) = \begin{pmatrix} \mathsf{T}\mathbf{u}_0(\mathbf{v}_0) & \cdots & \mathsf{T}\mathbf{u}_{n-1}(\mathbf{v}_0) \\ \vdots & \ddots & \vdots \\ \mathsf{T}\mathbf{u}_0(\mathbf{v}_{m-1}) & \cdots & \mathsf{T}\mathbf{u}_{n-1}(\mathbf{v}_{m-1}) \end{pmatrix}
$$

By the foregoing definition and equation (B.6),

$$
[\mathsf{T}\mathbf{x}]_{\mathcal{B}_1} = {}_{\mathcal{B}_1}[\mathsf{T}]_{\mathcal{B}_0}\,[\mathbf{x}]_{\mathcal{B}_0}, \qquad \mathbf{x} \in \mathcal{U} \tag{B.7}
$$

In practice, the case $\mathcal{B}_1 = \mathcal{B}_0$ is the most common one for presenting a basis representation of T, and the array ${}_{\mathcal{B}_0}[\mathsf{T}]_{\mathcal{B}_0}$ is called the $\mathcal{B}_0$*-basis representation of* T.

Consider the identity transformation $\mathsf{I} : \mathcal{U} \to \mathcal{U}$, defined by $\mathsf{I}\mathbf{x} = \mathbf{x}$, for all $\mathbf{x} \in \mathcal{U}$. By (B.7),

$$
[\mathbf{x}]_{\mathcal{B}_1} = [\mathsf{I}\,\mathbf{x}]_{\mathcal{B}_1} = {}_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0}\,[\mathbf{x}]_{\mathcal{B}_0}
$$

Thus, for $\mathbf{x} \in \mathcal{U}$, the $\mathcal{B}_0$-$\mathcal{B}_1$-basis representation of I maps $[\mathbf{x}]_{\mathcal{B}_0}$ to $[\mathbf{x}]_{\mathcal{B}_1}$, and the matrix ${}_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0}$ is called the $\mathcal{B}_0$-$\mathcal{B}_1$ *change of basis matrix*. Furthermore,

$$
\begin{aligned}
{}_{\mathcal{B}_1}[\mathsf{T}]_{\mathcal{B}_1}\,[\mathbf{x}]_{\mathcal{B}_1} &= [\mathsf{T}\mathbf{x}]_{\mathcal{B}_1} = [\mathsf{I}\,\mathsf{T}\,\mathsf{I}\,\mathbf{x}]_{\mathcal{B}_1} \\
&= {}_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0}\,{}_{\mathcal{B}_0}[\mathsf{T}]_{\mathcal{B}_0}\,[\mathsf{I}\,\mathbf{x}]_{\mathcal{B}_0} \\
&= {}_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0}\,{}_{\mathcal{B}_0}[\mathsf{T}]_{\mathcal{B}_0}\,{}_{\mathcal{B}_0}[\mathsf{I}]_{\mathcal{B}_1}\,[\mathbf{x}]_{\mathcal{B}_1}
\end{aligned}
$$

Hence,

$$
{}_{\mathcal{B}_1}[\mathsf{T}]_{\mathcal{B}_1} = {}_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0}\,{}_{\mathcal{B}_0}[\mathsf{T}]_{\mathcal{B}_0}\,{}_{\mathcal{B}_0}[\mathsf{I}]_{\mathcal{B}_1}
$$

and

$$
{}_{\mathcal{B}_0}[\mathsf{I}]_{\mathcal{B}_1}\,{}_{\mathcal{B}_1}[\mathsf{T}]_{\mathcal{B}_1} = {}_{\mathcal{B}_0}[\mathsf{T}]_{\mathcal{B}_0}\,{}_{\mathcal{B}_0}[\mathsf{I}]_{\mathcal{B}_1}
$$

Exercise 2. Given two bases, $\mathcal{B}_0 = \{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1}\}$ and $\mathcal{B}_1 = \{\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_{m-1}\}$, let U and V be the matrices in $GL(n, \mathbb{F})$ having the basis vectors as their columns; that is,

$$
U = (\,\mathbf{u}_0\ \mathbf{u}_1\ \cdots\ \mathbf{u}_{n-1}\,), \qquad V = (\,\mathbf{v}_0\ \mathbf{v}_1\ \cdots\ \mathbf{v}_{m-1}\,)
$$

Show that the $\mathcal{B}_0$-$\mathcal{B}_1$ change of basis matrix satisfies

$$
V\,{}_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0} = U \tag{B.8}
$$

As Excercise 2 shows, given two matrices, U and V, whose column vectors define bases $\mathcal{B}_0$ and $\mathcal{B}_1$, respectively, it is trivial to write down the $\mathcal{B}_0$-$\mathcal{B}_1$ change of basis matrix in terms of these two matrices:

$$_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0} = V^{-1}U$$

Every invertible matrix is a change-of-basis matrix, and every change-of-basis matrix is invertible. Thus, if $\mathcal{B}_0$ is a basis for vector space $\mathcal{U}$, if T is a linear transformation on $\mathcal{U}$, and if $A =_{\mathcal{B}_0} [\mathsf{T}]_{\mathcal{B}_0}$ denotes the $\mathcal{B}_0$-basis representation of T, then the set of all possible basis representations of T is

$$\{_{\mathcal{B}_1}[\mathsf{I}]_{\mathcal{B}_0}\ _{\mathcal{B}_0}[\mathsf{T}]_{\mathcal{B}_0}\ _{\mathcal{B}_0}[\mathsf{I}]_{\mathcal{B}_1} : \mathcal{B}_1 \text{ is a basis of } \mathcal{U}\} = \left\{S^{-1}AS : S \text{ is an invertible matrix } \right\} \tag{B.9}$$

This is just the set of all matrices that are *similar* to the given matrix A. Therefore, two distinct matrices that are similar are different basis representations of a single linear transformation. We use the term *similarity transformation* when referring to the operation $A \to S^{-1}AS$.

Exercise 3. [1] Generalize the set in (B.9) to include the case: $A =_{\mathcal{B}_1} [\mathsf{T}]_{\mathcal{B}_0}$ is an $m \times n$ matrix representation of the transformation $\mathsf{T} : \mathcal{U} \to \mathcal{V}$.

One would expect similar matrices to share some significant properties – at least, those properties that are intrinsic to the underlying linear transformation. In fact, this is an important theme in linear algebra. *It is often useful to step back from a question about a matrix to a question about some intrinsic property of the unerlying linear transformation, of which this matrix is only one particular manifestation.*

B.1.3 The Four Fundamental Subspaces

Let $\mathcal{U}$ be an n dimensional vector space and let $\mathcal{V}$ be an m dimensional vector space. Suppose we represent an operator $\mathsf{T} : \mathcal{U} \to \mathcal{V}$ by the $m \times n$ matrix $A \in M(m, n)$, using the elementary basis (B.2). That is

$$\mathsf{T}(\mathbf{x}) = \mathbf{y} \ \Leftrightarrow\ A\mathbf{x} = \mathbf{y}$$

Then the matrix A is a linear transformation with *domain* $\mathcal{U}$. The domain is one of four fundamental subspaces associated with A. Another one is the *range*. The range of A is the subspace of $\mathcal{V}$ given by

$$\mathcal{R}(A) = \{A\mathbf{x} : \mathbf{x} \in \mathcal{U}\} \subset \mathcal{V}$$

A third fundamental subspace is the *nullspace* of A,

$$\mathcal{N}(A) = \{\mathbf{x} : A\mathbf{x} = 0\} \subset \mathcal{U}$$

The nullspace of A is sometimes called the *kernel* of T. The range of A is sometimes called the *column space* of A because is represents the space of all linear combinations of the columns of A. Similarly, the space of all linear combinations of the rows of A is called the *row space* of A and is essentially the same space as $\mathcal{R}(A^t)$ – the column space of the transpose of A.

[1] *Hint:* see Definition 10, and the comment preceeding it.

B.2 Abstract Algebra

The main source of most material in this section is Fraleigh, *A First Course in Abstract Algebra*, [3].

B.2.1 Partitions and Equivalence Relations

Definition 11 (Partition). A *partition of a set* S is a decomposition into nonempty subsets such that every element of the set is in *one and only one* of the subsets. We call these subsets the *cells* of the partition.

Theorem 1. *Let S be a nonempty set and let $\sim$ be a relation between elements of S that satisfy the following properties for all $a, b, c \in S$:*

1. *(reflexive) $a \sim a$.*
2. *(symmetric) If $a \sim b$, then $b \sim a$.*
3. *(transitive) If $a \sim b$ and $b \sim c$, then $a \sim c$.*

Then $\sim$ yields a natural partition of S, where

$$\bar{a} = \{x \in S : x \sim a\}$$

is the cell containing a for all $a \in S$. Conversely, each partition of S gives rise to a natural relation $\sim$ satisfying the reflexive, symmetric, and transitive properties if $a \sim b$ is defined to mean $a \in \bar{b}$.

Definition 12 (Equivalence Relation). A relation $\sim$ on a set S satisfying the reflexive, symmetric, and transitive properties described in Theorem 1 is an *equivalence relation on S*. Each cell $\bar{a}$ in the natural partition given by an equivalence relation is an *equivalence class*.

(Notation: The symbol $\sim$ is usually reserved for an equivalence relation. We will use $\mathcal{R}$ for a relation between elements of a set S which is not necessarily an equivalence relation on S.)

Definition 13 (Congruence Modulo n). Let h and k be two integers in $\mathbb{Z}$ and let n be any positive integer. We define h *congruent to k modulo n*, written $h \equiv k \pmod{n}$, if $h - k$ is evenly divisible by n, so that $h - k = ns$ for some $s \in \mathbb{Z}$. Equivalence classes for congruence modulo n are *residue classes modulo n*.

Definition 14 (Addition modulo n). Let n be a fixed positive integer and let h and k be any integers. The remainder r when $h + k$ is divided by n is called the *sum of h and k modulo n*.

B.2.2 Groups and Subgroups

Binary Operations

Definition 15 (Binary Operation). A *binary operation $*$ on a set S* is a rule that assigns to each ordered pair (a, b) of elements of S some element of S.

Note that a binary operation on S must assign to each ordered pair (a, b) an element *that is again in S*. This requirement that the element be again in S is known as the *closure condition*; we require that S be *closed* under a binary operation on S.

Example 1. Our usual addition operator, $+$, is *not* a binary operation on the set $\mathbb{R}^+ = (0, \infty)$ of positive real numbers because $2 + (-2)$ is not in the set $\mathbb{R}^+$; that is, $\mathbb{R}^+$ is not closed under $+$.

There are two important points to remember when defining a binary operation $*$ on a set S. They are:

1. exactly one element is assigned to each possible ordered pair of elements of S
2. for each ordered pair of elements of S, the element assigned to it is again in S.

Definition 16 (Commutative Operation). A binary operation on a set S is *commutative* if $a * b = b * a$ for all $a, b \in S$.

Definition 17 (Associative Operation). A binary operation on a set S is *associative* if $(a*b)*c = a*(b*c)$ for all $a, b, c \in S$.

Theorem 2 (Associativity of Function Composition). *$f \circ (g \circ h) = (f \circ g) \circ h$ whenever this composition is defined.*

Proof. If this composition is defined, then for each x in the domain of h, we have

$$\begin{aligned}(f \circ (g \circ h))(x) &= f((g \circ h)(x)) = f(g(h(x))) \\ &= (f \circ g)(h(x)) = ((f \circ g) \circ h)(x)\end{aligned}$$

□

Remark 1. This is important because, for some operations (e.g., tranlation and scaling) that we use in Fourier analysis, order may matter. That is, the operations may not be commutative. However, the foregoing shows that composition of functions is always associative.

Groups

One motive for defining a binary operation on a set is the desire to solve simple linear equations involving that binary operation. For example, we might expect that the binary operation permits a solution x to the equation $2 * x = 3$. A little experimentation will show that such equations are solvable when there are special conditions on the set S and the operator $*$. A most basic set of conditions is given by the requirement that $\langle S, * \rangle$ define a *group*.

Definition 18 (Group). A *group* $\langle G, * \rangle$ is a set G, closed under a binary operation $*$, such that the following axioms are satisfied:

1. (associativity) The binary operation $*$ is associative.
2. (identity) There is an element e in G such that $e * x = x * e = x$ for all $x \in G$.
3. (inverse) For each a in G, there is an element a' in G with the property that $a * a' = a' * a = e$.

Theorem 3 (Unicity of identity and inverse). *The identity and inverses are unique in a group. To be precise, in a group $\langle G, * \rangle$, there is only one identity e such that*

$$e * x = x * e = x$$

for all $x \in G$. Likewise, for each $a \in G$, there is only one element a' such that

$$a * a' = a' * a = e$$

Definition 19 (Abelian Group). A group G is *Abelian* if its binary operation $*$ is commutative.

Here are a few examples from linear algebra:

Example 2. The axioms for a vector space $\mathcal{V}$ pertaining just to vector addition can be summarized by asserting that $\mathcal{V}$ with the operation of vector addition is an Abelian group.

Example 3. The set $M_{m \times n}(\mathbb{R})$ of all real-valued $m \times n$ matrices with binary operation matrix addition is a group. The $m \times n$ matrix with all entries 0 is the identity matrix. This group is Abelian.

Example 4. The set $M_n(\mathbb{R})$ of all real-valued $n \times n$ matrices with operation matrix multiplication is *not* a group. The $n \times n$ matrix with all entries 0 has no inverse.

Thus far our work deals primarily with integers, so the following example has special relevance:

Example 5. The set of non-negative integers, $\{0, 1, 2, \ldots\}$, with the usual addition operation $+$ is *not* a group. There is an identity element 0, but no inverse for, e.g., 2.

The last example is not intended to disparage the integers. Algebraic structures consisting of sets with binary operations for which not all of the group axioms hold can also be useful. Of these weaker structures, the *semigroup*, a set with an associative binary operation, has perhaps had the most attention in the math literature. Another example is the *monoid* – a semigroup that has an identity element for the binary operation.

Subgroups

We define a few more concepts from abstract algebra that are useful in our applied research.

Definition 20 (Subgroup). If a subset H of a group G is closed under the binary operation of G and if H itself is a group, then H is a *subgroup of* G. We shall let $H \leq G$ or $G \geq H$ denote that H is a subgroup of G, and $H < G$ or $G > H$ shall mean $H \leq G$ but $H \neq G$.

The elements of $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ are the possibilities for the remainder when an integer is divided by 4. With the operator $+$ denoting addition modulo 4, $\langle \mathbb{Z}_4, + \rangle$ satisfies the three group properties given in definition (18).[2] The only nontrivial proper subgroup of $\mathbb{Z}_4$ is $\{0, 2\}$ Note that $\{0, 3\}$ is *not* a subgroup of $\mathbb{Z}_4$, since $\{0, 3\}$ is not closed under $+$ (addition modulo 4). For example, $3 + 3 = 2$ and $2 \notin \{0, 3\}$. Similarly, the set $H = \{0, 1, 2\}$ does not satisfy the definition of a subgroup. For $1 + 2 = 3 \notin H$.

If a is a member of the group $\langle G, * \rangle$ then it is not hard to show that all elements of the set $\{a^n \,|\, n \in \mathbb{Z}\}$ are also members of $\langle G, * \rangle$. The box below contains the (trivial) demonstration of this fact. A few other propositions about the set $H \equiv \{a^n \,|\, n \in \mathbb{Z}\}$ are

1. H is a subgroup of G.
2. H is the smallest subgroup containing a.
3. H is called the *cyclic subgroup of G generated by a*, denoted $\langle a \rangle$.

Also, suppose we are given a group G and an element $a \in G$, such that

$$G \equiv \{a^n \,|\, n \in \mathbb{Z}\}$$

Then a is the *generator* of G and the group $G = \langle a \rangle$ is *cyclic.*

Notice that, if a is a member of the group $\langle G, * \rangle$ then the closure property guarantees that $a * a$ is also a member of G. Denote this element by $a^2 \equiv a * a$. Similarly, it must be the case that $a^2 * a \equiv a^3 \in G$. Proceeding by induction, it is clear that

$$a \in G \Rightarrow a^n \in G, \text{ for } n = 1, 2, \ldots \tag{B.10}$$

For any member a of the group $\langle G, * \rangle$ there exists an inverse a^{-1}, also a member of G. By an argument analogous to that of the preceding paragraph, it must be the case that $a^{-1} * a^{-1} \in G$, and, in general,

$$a \in G \Rightarrow a^{-n} \in G, \text{ for } n = 1, 2, \ldots \tag{B.11}$$

By convention, let $a^0 = 1$. Then equations (B.10) and (B.11) prove that, for any member a of the group $\langle G, * \rangle$ it must be the case that all elements of the set $\{a^n \,|\, n \in \mathbb{Z}\}$ are also members of $\langle G, * \rangle$

Below we will use the *direct product of groups*. The definition is made evident in the following theorem.

[2] To simplify notation, when the definition of the operator $+$ is clear from the context, let the group $\langle \mathcal{G}, + \rangle$ be denoted simply by $\mathcal{G}$, e.g., we will speak of the group $\mathbb{Z}_4$, when technically we mean $\langle \mathbb{Z}_4, + \rangle$.

Theorem 4. *Let $G_1, G_2, \ldots, G_n$ be groups and let $\cdot$ denote their respective binary operators. For $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ in*

$$\prod_i G_i \equiv G_1 \times G_2 \times \cdots \times G_n$$

define $(a_1, a_2, \ldots, a_n) \cdot (b_1, b_2, \ldots, b_n)$ to be $(a_1 \cdot b_1, a_2 \cdot b_2, \ldots, a_n \cdot b_n)$. Then $\prod_i G_i$ is a group, the direct product of the groups G_i*, under this binary operation.*

The following theorem is a fundamental result about direct products of integer groups.

Theorem 5. *The group $\mathbb{Z}_m \times \mathbb{Z}_n$ is isomorphic to $\mathbb{Z}_{mn}$ if and only if m and n are relatively prime, that is, if and only if the gcd of m and n is 1.*

As a result, when m and n are relatively prime, $\mathbb{Z}_m \times \mathbb{Z}_n$ is cyclic of order mn.

B.2.3 More Groups and Cosets

Groups of Permutations

Definition 21 (Permutation). A *permutation of a set A* is a function $\phi : A \to A$ that is both one to one and onto. In other words, a permutation of A is a one-to-one function from A onto A.

Function composition is a binary operation on the collection of all permutations of a set A. We call this operation *permutation multiplication*. Thus, if σ and τ are both permutations of the set A, then the composite function $\tau\sigma$ is also a permutation of A.

Theorem 6. *Let A be a nonempty set, and let S_A be the collection of all permutations of A. Then S_A is a group under permutation multiplication.*

Definition 22 (Symmetric Group). Let A be the finite set $\{1, 2, \ldots, n\}$. The group of all permutations of A is the *symmetric group on n letters*, and is denoted by S_n.

Note that S_n has $n!$ elements.

Orbits, Cycles, and the Alternating Groups

Orbits

Each permutation σ of a set A determines a natural partition of A into cells with the property that $a, b \in A$ are in the same cell if and only if $b = \sigma^n(a)$ for some $n \in \mathbb{Z}$. We establish this partition using an appropriate equivalence relation:

$$\text{For } a, b \in A, \text{ let } a \sim b \text{ if and only if } b = \sigma^n(a) \text{ for some } n \in \mathbb{Z} \tag{B.12}$$

Definition 23 (Orbits of σ). Let σ be a permutation of a set A. The equivalence classes in A determined by the equivalence relation (B.12) are the *orbits of σ*.

Cycles

Definition 24 (Cycle). A permutation $\sigma \in S_n$ is a *cycle* if it has at most one orbit containing more than one element. The *length* of a cycle is the number of elements in its largest orbit.

Theorem 7. *Every permutation of a finite set is a product of disjoint cycles.*

Even and Odd Permutations

Definition 25 (Transposition). A cycle of length 2 is a *transposition.*

Thus, a transposition leaves all but two elements fixed. A computation shows that

$$(a_1, a_2, \ldots, a_n) = (a_1, a_n)(a_1, a_{n-1}) \cdots (a_1, a_3)(a_1, a_2)$$

Therefore, any cycle is a product of transpositions. We then have the following as a corollary to Theorem 7.

Corollary 1. Any permutation of a finite set of at least two elements is a product of transpositions.

Corollary 1 simply states that any rearrangement of n objects can be achieved by successively interchanging pairs of them.

Lemma 1. Let $\sigma \in S_n$ and let τ be a transposition in S_n. The number of orbits of σ and the number of orbits of $\tau\sigma$ differ by 1.

Theorem 8. *No permutation can be expressed both as a product of an even number of transpositions and as a product of an odd number of transpositions.*

Definition 26 (Even or Odd Permutation). A permutation of a finite set is *even* (resp. *odd*) if it can be expressed as a product of an even (resp. odd) number of transpositions.

Proposition 1. If $n \geq 2$, then the collection of all even permutations of $\{1, 2, \ldots, n\}$ forms a subgroup of order $n!/2$ of the symmetric group S_n. The same is true for the subgroup of all odd permutations.

Definition 27 (Alternating Group). The subgroup of S_n consisting of the even permutations of n letters is the *alternating group A_n on n letters.*

Cosets and the Theorem of Lagrange

Cosets

Let H be a subgroup of a group G, which may be of finite or infinite order. We exhibit two partitions of G by defining two equivalence relations, $\sim_L$ and $\sim_R$ on G.

Theorem 9. *Let H be a subgroup of G. Let the relation $\sim_L$ be defined on G by*

$$a \sim_L b \;\Leftrightarrow\; a^{-1}b \in H$$

Let $\sim_R$ be defined by

$$a \sim_R b \;\Leftrightarrow\; ab^{-1} \in H$$

Then $\sim_L$ and $\sim_R$ are both equivalence relations on G.

Definition 28 (Cosets). Let H be a subgroup of a group G. The subset $aH = \{ah : h \in H\}$ of G is the *left coset* of H containing a, while $Ha = \{ha : h \in H\}$ is the *right coset* of H containing a.

Proposition 2. For an abelian subgroup H of G, the partition of G into left cosets of H and the partition into right cosets are the same.

The Theorem of Lagrange

Proposition 3. Every coset (left or right) of a subgroup H has the same number of elements as H.

Theorem 10 (Theorem of Lagrange). *Let H be a subgroup of a finite group G. Then the order of H is a divisor of the order of G.*

Proof. Let n be the order of G and let m be the order of H. By the preceding proposition, every coset of H has m elements. Let r be the number of cells in the partition of G into cosets of H. Then $n = rm$, so m is indeed a divisor of n. □

Corollary 2. Every group of prime order is cyclic

Proof. Let G be of prime order p and let a be an element of G different from the identity. Then the cyclic subgroup $\langle a \rangle$ of G generated by a has at least two elements, a and e. But by Theorem 10, the order $m \geq 2$ of $\langle a \rangle$ must divide the prime number p. Thus, we must have $m = p$ and $\langle a \rangle = G$, so G is cyclic. □

Theorem 11. *The order of an element of a finite group divides the order of the group.*

Proof. Recalling that the order of an element is the same as the order of the cyclic group generated by that element, we see that the theorem follows directly from Theorem 10. □

Definition 29 (Index of H in G). Let H be a subgroup of a group G. The number of left cosets of H in G is the *index* $(G : H)$ *of H in G.*

The index may be finite or infinite. If G is finite, then $(G : H)$ is finite and

$$(G : H) = |G|/|H|$$

since every coset of H contains $|H|$ elements.

Theorem 12. *Suppose H and K are subgroups of a group G such that $K \leq H \leq G$, and suppose $(H : K)$ and $(G : H)$ are both finite. Then $(G : K)$ is finite and $(G : K) = (G : H)(H : K)$.*

B.2.4 Homomorphisms and Factor Groups

Homomorphisms

Structure-Relating Maps

Let G and G' be groups. We are interested in a map $\phi : G \to G'$ that relates the group structure of G to the group structure of G'. Such a map often gives us information about the structure of G' from known structural properties of G, or information about the structure of G from known structural properties of G'. Now *group structure is completely determined by the binary operation on the group.* We define such a structure-relating map for groups, and then point out how the binary operations of G and G' are related by such a map.

Definition 30 (Homomorphism). A map ϕ of a group G into a group G' is a *homomorphism* if

$$\phi(ab) = \phi(a)\phi(b) \tag{B.13}$$

for all $a, b \in G$.

In equation (B.13), the product ab on the left-hand side takes place in G while the product $\phi(a)\phi(b)$ on the right takes place in G'. Thus ϕ gives a relation between the two binary operations, and hence between the two group structures.

Example 6 (Reduction Modulo n). Let γ be the natural map of $\mathbb{Z}$ into $\mathbb{Z}_n$ given by $\gamma(m) = r$ where r is the remainder when m is divided by n. Then γ is a homomorphism.

In more familiar notation, this example implies that, for all $s, t \in \mathbb{Z}$,

$$(s+t)(\mathrm{mod} n) = s(\mathrm{mod} n) + t(\mathrm{mod} n)$$

Remember: the addition on the right hand side takes place in $\mathbb{Z}_n$, thus it is addition modulo n. For instance, if we were to implement this example in Matlab, where the syntax `mod(s,n)` yields $s(\mathrm{mod} n)$, we would observe the following:

```
>> mod((s+t),n) == mod(s,n) + mod(t,n)

ans:   0

>> mod((s+t),n) == mod(mod(s,n)+mod(t,n),n)

ans:   1
```

In the first line of code, we used the standard $\mathbb{Z}$ addition operator on the right hand side, which does not yield the desired equality. The second line uses the appropriate addition modulo n operator.

Properties of Homomorphisms

We turn to structural features of G and G' that are preserved by a homomorphism $\phi : G \to G'$. We first give a set-theoretic definition. Note the use of square brackets when we apply a function to a *subset* of its domain.

Definition 31 (Image and Inverse Image). Let ϕ be a mapping of a set X into a set Y, and let $A \subseteq X$ and $B \subseteq Y$. The *image* $\phi[A]$ *of* A *in* Y *under* ϕ is $\{\phi(a) : a \in A\}$. The set $\phi[X]$ is sometimes called the *range of* ϕ. The *inverse image* $\phi^{-1}[B]$ *of* B *in* X is $\{x \in X : \phi(x) \in B\}$.

Theorem 13. *Let ϕ be a homeomorphism of a group G into a group G'.*

1. *If e is the identity in G, then $\phi(e)$ is the identity in G'.*
2. *If $a \in G$, then $\phi(a^{-1}) = \phi(a)^{-1}$.*
3. *If H is a subgroup of G, then $\phi[H]$ is a subgroup of G'.*
4. *If K' is a subgroup of G', then $\phi^{-1}[K']$ is a subgroup of G.*

Loosely speaking, the theorem states that a homomorphism preserves the identity, inverses, and subgroups.

Definition 32 (Fibre). Let $\phi : G \to G'$ be a group homomorphism. For each $a' \in G'$ the inverse image $\phi^{-1}[\{a'\}]$ is the *fibre over* a' *under* ϕ.

When the argument of a set function is a singleton, e.g., $\{a'\}$, we will omit the curly braces in order to simplify notation. Thus, $\phi^{-1}[\{a'\}]$ will appear as $\phi^{-1}[a']$.

Definition 33 (Kernel). Let $\phi : G \to G'$ be a group homomorphism. The subgroup $\phi^{-1}[e'] = \{x \in G : \phi(x) = e'\}$ is the *kernel of* ϕ, denoted by $\ker(\phi)$.

Theorem 14. *Let $\phi : G \to G'$ be a group homomorphism, and let $H = \ker(\phi)$. For $a \in G$, the set*

$$\phi^{-1}[\phi(a)] = \{x \in G : \phi(x) = \phi(a)\}$$

is the left coset aH of H, and is also the right coset Ha of H. Consequently, the two partitions of G into left cosets and into right cosets are the same.

Corollary 3. A group homomorphism $\phi : G \to G'$ is a one-to-one map if and only if $\ker(\phi) = \{e\}$.

Definition 34 (Normal Subgroup). A subgroup H of a group G is *normal* if its left and right cosets coincide, that is, if $gH = Hg$ for all $g \in G$.

Note that all subgroups of abelian groups are normal. Also, Theorem 14 shows that the kernel of a homomorphism $\phi : G \to G'$ is a normal subgroup of G.

Isomorphism and Cayley's Theorem

Definition and Elementary Properties

Definition 35 (Isomorphism). An *isomorphisms* $\phi : G \to G'$ is a homomorphism that is one-to-one and onto G'. The relation $G \simeq G'$ denotes the existence of an isomorphism from G onto G', in which case we call G and G' *isomorphic.*

Theorem 15. *Let $\mathcal{G}$ be any collection of groups, and define $G \simeq G'$ for G and G' in $\mathcal{G}$ if there exists an isomorphism $\phi : G \to G'$. Then $\simeq$ is an equivalence relation.*

Theorem 16 (Cayley's Theorem). *Every group is isomorphic to a group of permutations.*

Factor Groups

Factor Groups from Homomorphisms

Theorem 14 shows that for each $a \in G$, the fibre $\phi^{-1}[\phi(a)] = \{x \in G : \phi(x) = \phi(a)\}$ is the left coset aH of H and is also the right coset Ha of H. Since these left and right cosets coincide, we will simply refer to them as the cosets of H.

Now $\phi[G]$ is a group by Theorem 13. We associate with each $\phi(a) = a' \in \phi[G]$ the coset $\phi^{-1}[a']$ (the fibre over a' under ϕ). By renaming $a' \in \phi[G]$ by the name of the associated coset, $\phi^{-1}[a']$, we can consider the cosets to form a group. This group of cosets will be isomorphic to the group $\phi[G]$ since its elements are just the elements of $\phi[G]$ renamed.

In summary, the cosets of the kernel of a group homomorphism $\phi : G \to G'$ form a group isomorphic to the subgroup $\phi[G]$ of G. The binary operation on the cosets can be computed in terms of the group operation of G'. This group of cosets is the *factor group of G modulo H*, and is denoted by G/H.

Theorem 17. *Let $\phi : G \to G'$ be a group homomorphism with kernel H. Then the cosets of H form a group, G/H, whose binary operation defines the product $(aH)(bH)$ of two cosets by choosing elements a and b from the cosets, and letting $(aH)(bH) = (ab)H$. Also, the map $\mu : G/H \to \phi[G]$ defined by $\mu(aH) = \phi(a)$ is an isomorphism.*

Example 7. Example 6 considered the map $\gamma : \mathbb{Z} \to \mathbb{Z}_n$ where $\gamma(m)$ is the remainder when m is divided by n. We know that γ is a homomorphism and, of course, $\ker(\gamma) = n\mathbb{Z}$. By Theorem 17, we see that the factor group $\mathbb{Z}/n\mathbb{Z}$ is isomorphic to $\mathbb{Z}_n$. The cosets of $n\mathbb{Z}$ are the *residue classes modulo n* described in Definition 13. For example, taking $n = 12$, we see the cosets of $12\mathbb{Z}$ are

$$\begin{aligned} 12\mathbb{Z} &= \{\ldots, -24, -12, 0, 12, 24, \ldots\} \\ 1 + 12\mathbb{Z} &= \{\ldots, -23, -11, 1, 13, 25, \ldots\} \\ 2 + 12\mathbb{Z} &= \{\ldots, -22, -10, 2, 14, 26, \ldots\} \\ &\vdots \\ 11 + 12\mathbb{Z} &= \{\ldots, -13, -1, 11, 23, 35, \ldots\} \end{aligned}$$

Note that the isomorphism $\mu : \mathbb{Z}/12\mathbb{Z} \to \mathbb{Z}_{12}$ of Theorem 17 assigns to each coset of $12\mathbb{Z}$ its smallest nonnegative element. That is, $\mu(12\mathbb{Z}) = 0,\ \mu(1 + 12\mathbb{Z}) = 1$, etc.

The factor group $\mathbb{Z}/n\mathbb{Z}$ in the preceding example is classic. Recall that we refer to the cosets of $n\mathbb{Z}$ as *residue classes modulo n.* Two integers in the same coset are *congruent modulo n.* This terminology is carried over to other factor groups. A factor group G/H is called the *factor group of G modulo H.* Elements in the same coset of H are called *congruent modulo H.* By abuse of notation, we may sometimes write $\mathbb{Z}/n\mathbb{Z} = \mathbb{Z}_n$ and think of $\mathbb{Z}_n$ as the additive group of residue classes of $\mathbb{Z}$ modulo $\langle n \rangle$, or abusing notation further, modulo n.

Factor Groups from Normal Subgroups

So far, we have obtained factor groups only from homomorphisms. The following theorem again takes H to be a subgroup of G, but it does not assume that H is the kernel of a homomorphism.

Theorem 18. *Let H be a subgroup of a group G. Then left coset multiplication is well defined by the equation*

$$(aH)(bH) = (ab)H$$

if and only if left and right cosets coincide, so that $aH = Ha$ for all $a \in G$.

Corollary 4. Let H be a subgroup of a group G whose left and right cosets coincide. Then the cosets of H form a group G/H under the binary operation $(aH)(bH) = (ab)H$.

Definition 36 (Factor Group). The group G/H in the preceding corollary is the *factor group* (or *quotient group*) of G modulo H.

In summary, if H is a normal subgroup of G, then the cosets of H form the factor group, G/H. Recalling that the kernel of a homomorphism $\phi : G \to G'$ is a normal subgroup of G, we see again that $G/\ker(\phi)$ is a group.

The following three conditions are equivalent characterizations for a normal subgroup H of a group G.

1. $ghg^{-1} \in H$ for all $g \in G$ and $h \in H$.
2. $gHg^{-1} = H$ for all $g \in G$.
3. $gH = Hg$ for all $g \in G$.

Condition 2 is often taken as the definition of a normal subgroup of H of a group G.

The map $i_g : G \to G$ defined by $i_g(x) = gxg^{-1}$ is a homomorphism of G into itself. Clearly

$$i_g(a) = gag^{-1} = gbg^{-1} = i_g(b)$$

if and only if $a = b$, so i_g is one-to-one. Since $g(g^{-1}yg)g^{-1} = y$, we see that i_g is onto G, so it is an isomorphism of G with itself.

Definition 37 (Automorphism). An isomorphism $\phi : G \to G$ is an *automorphism* of G. The automorphism $i_g : G \to G$ where $i_g(x) = gxg^{-1}$ is the *inner automorphism of G by g.*

The equivalence of conditions 2. and 3. states that $gH = Hg$ for all $g \in G$ if and only if $i_g[H] = H$ for all $g \in G$; that is, if and only if H is *invariant* under all inner automorphisms of G. It is important to realize that $i_g[H] = H$ is an equation in sets; we need not have $i_g(h) = h$ for all $h \in H$. That is, i_g may perform a nontrivial permutation of the set H. Thus, we see that the normal subgroups of a group G are precisely those that are invariant under all inner automorphisms.

Fundamental Homomorphism Theorem

We have seen that every homomorphism $\phi : G \to G'$ gives rise to a natural factor group, namely $G/Ker(\phi)$. We now show that each factor group G/H gives rise to a natural homomorphism having H as kernel.

Theorem 19. *Let H be a normal subgroup of G. Then $\gamma : G \to G/H$ given by $\gamma(x) = xH$ is a homomorphism with kernel H.*

We have seen in Theorem 17 that if $\phi : G \to G'$ is a homomorphism with kernel H, then $\mu : G/H \to \phi[G]$ where $\mu(gH) = \phi(g)$ is an isomorphism. Theorem 19 shows that $\gamma : G \to G/H$ defined by $\gamma(g) = gH$ is a homomorphism. Thus, we see that the homomorphism ϕ can be *factored*, $\phi = \mu\gamma$, where γ is a homomorphism and μ is a one-to-one homomorphism with image $\phi[G]$. We state this as a theorem.

Theorem 20 (Fundamental Homomorphism Thoerem). *Let $\phi : G \to G'$ be a group homomorphism with kernel H. Then:*

1. $\phi[G]$ is a group,
2. the map $\mu : G/H \to \phi[G]$ given by $\mu(gH) = \phi(g)$ is an isomorphism,
3. the map $\gamma : G \to G/H$ given by $\gamma(g) = gH$ is a homomorphism,
4. for each $g \in G$ we have $\phi(g) = \mu\gamma(g)$.

The isomorphism μ is sometimes called the *natural* or *canonical* isomorphism, and the same adjectives are used to describe the homomorphism γ.

To summarize the foregoing, every homomorphism with domain G and kernel H gives rise to a factor group G/H, and every factor group gives rise to a homomorphism $\gamma : G \to G/H$.

Factor-Group Computations and Simple Groups

Let N be a normal subgroup of G. In the factor group G/N, the subgroup N acts as the identity element. We may regard N as being collapsed to a single element, either to 0 in additive notation or to e in multiplicative notation. This collapsing of N together with the algebraic structure of G require that other subsets of G, namely the cosets of N, also collapse into a single element in the factor group.

Theorem 21. *Let $G = H \times K$ be the direct product of groups H and K. Then $\bar{H} = \{(h, e) : h \in H\}$ is a normal subgroup of G, and $G/\bar{H} \simeq K$. Also, $\bar{K} = \{(e, k) : k \in K\}$ is a normal subgroup of G, and $G/\bar{K} \simeq H$.*

Example 8. Let $G = \mathbb{Z}_3 \times \mathbb{Z}_4$ and let $\bar{H} = \langle(1,0)\rangle = \{(0,0), (1,0), (2,0)\} \simeq \mathbb{Z}_3$. Then the factor group $G/\bar{H} = (\mathbb{Z}_3 \times \mathbb{Z}_4)/\langle(1,0)\rangle$ is the collection of cosets $\{(0,0) + \bar{H},\ (0,1) + \bar{H},\ (0,2) + \bar{H},\ (0,3) + \bar{H}\}$. Since we can compute in the factor group by choosing representative elements $\{(0,0),\ (0,1),\ (0,2),\ (0,3)\}$ from the cosets, it is clear that the factor group is isomorphic to $\mathbb{Z}_4$. Thus,

$$(\mathbb{Z}_3 \times \mathbb{Z}_4)/\langle(1,0)\rangle \simeq \mathbb{Z}_4$$

as predicted by the preceding theorem. Similarly, letting $\bar{K} = \langle(0,1)\rangle = \{(0,0), (0,1), (0,2), (0,3)\} \simeq \mathbb{Z}_4$, we see that

$$G/\bar{K} = (\mathbb{Z}_3 \times \mathbb{Z}_4)/\langle(0,1)\rangle \simeq \mathbb{Z}_3$$

Series of Groups

Subnormal and Normal Series

Definition 38 (Subnormal Series). A *subnormal* (or *subinvariant*) *series of a group G* is a finite sequence $\{H_k\}_{k=0}^n$ of subgroups of G such that $H_k < H_{k+1}$ and H_k is a normal subgroup of H_{k+1} with $H_0 = \{e\}$ and $H_n = G$.

Definition 39 (Normal Series). A *normal* (or *invariant*) *series of a group G* is a finite sequence $\{H_k\}_{k=0}^n$ of normal subgroups of G such that $H_k < H_{k+1}$ with $H_0 = \{e\}$ and $H_n = G$.

B.2.5 Rings and Fields

Definition 40 (Ring). A ring $\langle R, +, \cdot\rangle$ is a set R together with two binary operations $+$ and $\cdot$ that satisfy the following three conditions:

1. $\langle R, +\rangle$ is an Abelian group.
2. The operator $\cdot$ is associative.
3. For all $a, b, c \in R$ the following left and right distributive laws hold:

$$a \cdot (b + c) = a \cdot b + a \cdot c$$

$$(a + b) \cdot c = a \cdot c + b \cdot c$$

We refer to $\langle R, + \rangle$ as the *additive group of the ring*.

Example 9. Consider the cyclic group $\langle \mathbb{Z}_n, + \rangle$. For any $a, b \in \mathbb{Z}_n$, define multiplication, $a \cdot b$, to be the remainder when the usual product, ab, is divided by n. Then $\langle \mathbb{Z}_n, +, \cdot \rangle$ is a ring. For instance, on $\mathbb{Z}_{10}$, we have $3 \cdot 7 = 1$. Thus defined, the operator $\cdot$ is called "multiplication modulo n."

Definition 41 (Ring Homomorphism). Let R and R' be rings. A map $\phi : R \to R'$ is a *homomorphism* if the following hold for all $a, b \in R$:

1. $\phi(a + b) = \phi(a) + \phi(b)$
2. $\phi(a \cdot b) = \phi(a) \cdot \phi(b)$

Definition 42 (Ring Isomorphism). An *isomorphism* $\phi : R \to R'$ from ring R to ring R' is a homomorphism that is one-to-one and onto. The rings R and R' are then *isomorphic*.

Definition 43 (Commutative Ring). A ring in which multiplication is commutative is a *commutative ring*.

Definition 44 (Ring with Unity). A ring with a multiplicative identity is a *ring with unity*. A multiplicative identity is called *unity*.

(In a ring with unity it is standard to let the symbol 1 denote that element representing unity.)

Example 10. Denote the greatest common divisor of two integers $r, s \in \mathbb{Z}$, by $\gcd(r, s)$. For integers $r, s \in \mathbb{Z}$, where $\gcd(r, s) = 1$, the cyclic rings $\mathbb{Z}_{rs}$ and $\mathbb{Z}_r \times \mathbb{Z}_s$ are isomorphic.

Definition 45 (Multiplicative Inverse). A *multiplicative inverse* of an element a in a ring R is an element $a^{-1} \in R$ such that $a^{-1} \cdot a = a \cdot a^{-1} = 1$

Finally, we arrive at a precise definition of a *field*.

Definition 46 (Unit, Division Ring, Field, Skew Field). Let R be a ring with unity. An element $u \in R$ is a *unit* if it has a multiplicative inverse in R. If every non-zero element of R is a unit, then R is a *division ring*. A *field* is a commutative division ring. A non-commutative division ring is a *skew field*.

Example 11. The ring $\mathbb{Z}_{12}$ is isomorphic to $\mathbb{Z}_3 \times \mathbb{Z}_4$ (see Example 10). The elements of $\mathbb{Z}_{12}$ are $\{0, 1, 2, \ldots, 11\}$. Let us find the units. Of course 1 is a units. Since 7^2 is 49, in the modulo 12 arithmetic of the ring $\mathbb{Z}_{12}$, $7 \cdot 7 = 1$. Therefore, 7 is its own inverse and is a unit. The other units are 5 and 11. It is clear, however, that the remaining members of $\mathbb{Z}_{12}$ are not units. We conclude that $\mathbb{Z}_{12}$ is not a division ring (hence, not a field).

In general, the units of $\mathbb{Z}_n$ are those $m \in \mathbb{Z}_n$ satisfying $\gcd(m, n) = 1$. An equivalent statement is that m is a unit of $\mathbb{Z}_n$ if and only if m and n are *relatively prime*. This fact makes it easy to check whether $\mathbb{Z}_n$ is a division ring. Since multiplication on this ring is always commutative, $\mathbb{Z}_n$ is a field if and only if it is a division ring.

Example 12. The previous example showed that not all non-zero elements of $\mathbb{Z}_{12}$ are units and, thus, $\mathbb{Z}_{12}$ is not a field. The ring $\mathbb{Z}_{12}$ is isomorphic to the ring $\mathbb{Z}_3 \times \mathbb{Z}_4$. The members of $\mathbb{Z}_3$ are $\{0, 1, 2\}$. Clearly 1 is a unit. Since $2 \cdot 2 = 1$, the element 2 is also a unit. Therefore, all non-zero members of $\mathbb{Z}_3$ are units, and it follows that $\mathbb{Z}_3$ is a field. On the other hand, $\mathbb{Z}_4$ is *not* a field since, for example, 2 does not have a multiplicative inverse in this ring.

To conclude the general discussion in the paragraph above example 12, let $+$ and $\cdot$ represent addition modulo n and multiplication modulo n, respectively. That is, $a \cdot b$ is equal to the remainder when the usual product, ab, is divided by n. Let 0 be the additive identity, and $-a$ the additive inverse of a, and 1 the multiplicative identity. To determine that $\mathbb{Z}_n$ is *not* a field, it is sufficient to find one non-zero member that does not have a multiplicative inverse (i.e., a member that is not a unit). This is equivalent to finding a member a of $\mathbb{Z}_n$ such that a and n are not relatively prime.

Example 13. Suppose $n = 32$ so that the elements of $\mathbb{Z}_n$ are $\{0, 1, 2, \ldots, 31\}$. In this case it is easy to find members of this ring – e.g., $2, 4, 6, \ldots$ – that are not relative primes of 32. That there exists a member with no multiplicative inverse proves $\mathbb{Z}_{32}$ is not a field.

B.3 Probability and Statistics

This section is a sparse collection of results from the theory of probability and mathematical statistics. The presentation style may seem terse and incoherent, especially to those without the necessary background in probability and statistics.

B.3.1 States of Nature, Events, and Random Variables

A concept which is often useful for describing natural phenomena is the "state of nature." It is common to let ω denote a particular state of nature. For example, ω may be the outcome of an experiment, the position of the planets, etc. The space of all possible states of nature is denoted Ω, and is sometimes called the "outcome space."

Another useful concept is that of a measurable subset $A \subset \Omega$, which we call an *event*. Denote by $\mathcal{S}$ the collection of all measurable subsets of Ω. Thus, events are elements of $\mathcal{S}$. An event may contain multiple states of nature or experimental outcomes; for instance,

$$A = \{\omega_{\alpha(0)}, \omega_{\alpha(1)}, \omega_{\alpha(2)}, \ldots, \omega_{\alpha(N-1)}\} \in \mathcal{S}$$

Exercise 4. For a six-sided die, the possible outcomes of a single roll is $\{1, 2, \ldots, 6\}$. For the experiment that consists of rolling such a die N times, identify the following:

1. the outcome space, Ω.
2. the cardinality, $|\Omega|$.

A *random variable* is a function which maps Ω into some other space. We usually denote such functions by capital letters. Some common examples are

$$X : \Omega \to \mathbb{R}, \qquad \mathbf{X} : \Omega \to \mathbb{R}^n$$

$$Y : \Omega \to \mathbb{C}, \quad \text{and} \quad \mathbf{Z} : \Omega \to \mathbb{C}^n.$$

We use bold font to indicate that the random variable is vector-valued.

Exercise 5. As in exercise 4, a six-sided die is rolled N times. Think of some random variables for this experiment. Write them down as functions Ω. Into what spaces do they map?

Recap

As described thus far, the probability modeling process reduces to only a few main ideas.

1. Conduct experiments and/or make observations; the observed state of nature is $\omega \in \Omega$.
2. Evaluate the function X at the point ω; the result is a particular realization of the random variable X, namely the *random sample* $x = X(\omega)$.

B.3.2 Probability Measures, Distributions, and Parameters

Denote by Θ the space of all parameters that we find interesting. It can be useful – especially from a Bayesian point of view – to adopt the modern convention of writing parameters in upper case; The reason becomes apparent when we view parameters as random variables. As such, they are functions on Ω.

To each state of nature $\omega_k \in \Omega$ is associated a particular *realization* of the random variable, which we denote by the corresponding lower case symbol. For example, a realization of the random variable $\mathbf{X} : \Omega \to \mathbb{R}^N$ is

$$\mathbf{x} = (x_0, x_1, \ldots, x_{N-1}) = (X_0(\omega_k), X_1(\omega_k), \ldots, X_{N-1}(\omega_k)) = \mathbf{X}(\omega_k)$$

Another example is the parameter-space analog of the foregoing, which we denote as follows:

$$\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_{N-1}) = (\Theta_0(\omega_k), \Theta_1(\omega_k), \ldots, \Theta_{N-1}(\omega_k)) = \boldsymbol{\Theta}(\omega_k)$$

A measure μ on the space Ω is called a *probability measure* if it satisfies the following:

1. $\mu : \Omega \to [0, 1]$
2. $A \subset B \Rightarrow \mu(A) \leq \mu(B)$
3. $\mu(\Omega) = 1$

A probability measure is sometimes called a probability distribution.

The first condition above can be stated equivalently as $0 \leq \mu(A) \leq 1,\ A \subset \Omega$. The second condition says that a probability distribution is a cumulative, or monotonically increasing set function. The final condition essentially says the probability that something (or nothing)[3] occurs is 1.

The triple $(\Omega, \mathcal{S}, \mu)$ – an outcome space, the collection of all measurable subsets, and a probability measure, respectively – is called a *probability space*. A probability measure μ often depends on some parameters of interest, say Θ, and it may be useful to make this explicit. In such cases we use μ_Θ to denote the probability measure.

We often wish to compute the probability that a random variable has a particular realization; for example, we seek the probability that X takes on a value in the interval $[x_0, x_1)$. This probability is *not* found by measuring the interval $[x_0, x_1)$. Instead, we measure the set of all ω for which $X(\omega) \in [x_0, x_1)$. That is,

$$\text{Probability}\{x_0 \leq X(\omega) < x_1\} = \mu_\Theta\{\omega : x_0 \leq X(\omega) < x_1\}$$

B.3.3 Exponential Families

Definition 47 (Exponential Family).
(Schervish [6], p. 102) A parametric family with parameter space Θ and density $f_{X|\Theta}(x|\theta)$ with respect to a measure ν on $(\mathcal{X}, \mathcal{B})$ is called an *exponential family* if

$$f_{X|\Theta}(x|\theta) = c(\theta)h(x)\exp\left[\sum_{k=0}^{K-1} \pi_k(\theta)t_k(x)\right] \tag{B.14}$$

for some measurable functions $\{\pi_k\}_{0\leq k<K}$, $\{t_k\}_{0\leq k<K}$, and some integer K.

Since $f_{X|\Theta}$ in Definition 47 is a probability density, the function $c(\theta)$ can be written as

$$c(\theta) = \left\{\int_{\mathcal{X}} h(x)\exp\left[\sum_{k=0}^{K-1} \pi_k(\theta)t_k(x)\right]\,\mathrm{d}\nu(x)\right\}^{-1}$$

so that dependence on θ is through the vector $\boldsymbol{\pi}(\theta) = (\pi_0(\theta), \ldots, \pi_{K-1}(\theta)) \in \mathbb{R}^K$.

Example 14 (Gaussian).
Suppose $\{X_n\}_{n=0}^{\infty}$ are i.i.d. normal μ, σ^2 random variables; let $\boldsymbol{\theta} = (\mu, \sigma^2)$ and $\mathbf{x} = (x_0, x_1, \ldots, x_{N-1})$. Then,

$$\begin{aligned} f_{X|\Theta}(\mathbf{x}|\boldsymbol{\theta}) &= \left(\sigma\sqrt{2\pi}\right)^{-N}\exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x_n-\mu)^2\right] \\ &= (2\pi)^{-N/2}\sigma^{-N}\exp\left(-\frac{N\mu^2}{2\sigma^2}\right)\exp\left(\frac{\mu}{\sigma^2}N\bar{\mathbf{x}} - \frac{1}{2\sigma^2}\sum_{n=0}^{N-1}x_n^2\right) \end{aligned}$$

In this form it is clear that $f_{X|\Theta}$ is expressed in terms of equation (B.14) with the following definitions:

$$K = 2, \quad c(\boldsymbol{\theta}) = \sigma^{-N}\exp\left(-\frac{N\mu^2}{2\sigma^2}\right), \quad h(\mathbf{x}) = (2\pi)^{-N/2},$$

$$\pi_0(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2}, \quad \pi_1(\boldsymbol{\theta}) = -\frac{1}{2\sigma^2}, \quad t_0(\mathbf{x}) = N\bar{\mathbf{x}}, \quad t_1(\mathbf{x}) = \sum_{n=0}^{N-1}x_n^2$$

[3] Of course, the empty set $\{0\}$ is a subset of Ω.

Example 15 (Poisson).
Suppose $\{X_n\}_{n=0}^{\infty}$ are i.i.d. Poisson λ random variables; let $\theta = \lambda$ and $\mathbf{x} = (x_0, x_1, \ldots, x_{N-1})$. Then,

$$f_{X|\Theta}(\mathbf{x}|\theta) = \prod_{n=0}^{N-1} \frac{\mathrm{e}^{-\theta}\theta^{x_n}}{x_n!} = \frac{\exp\left(-N\theta + \log\theta \cdot \sum_{n=0}^{N-1} x_n\right)}{\prod_{n=0}^{N-1} x_n!}$$
$$= \mathrm{e}^{-N\theta} \left(\prod_{n=0}^{N-1} x_n!\right)^{-1} \exp\left(\log\theta \cdot N\bar{\mathbf{x}}\right)$$

In this form it is clear that $K = 1$,

$$c(\theta) = \mathrm{e}^{-N\theta}, \quad h(\mathbf{x}) = \left(\prod_{n=0}^{N-1} x_n!\right)^{-1}, \quad \pi_0(\theta) = \log\theta, \quad t_0(\mathbf{x}) = N\bar{\mathbf{x}}.$$

A particular exponential family of distributions has a fixed K, a defines the (vector-valued) functionsions $\boldsymbol{\pi} : \Theta \to \mathbb{R}^K$ and $\mathbf{t} : \mathcal{X} \to \mathbb{R}^K$. By composition with the functions $\Theta : \Omega \to \Theta$ and $X : \Omega \to \mathcal{X}$, we arrive at the following functions of Ω:

$$\boldsymbol{\pi}(\Theta) : \Omega \to \mathbb{R}^K, \quad \mathbf{t}(X) : \Omega \to \mathbb{R}^K$$

The inner product of the resulting vectors is a map from Ω to $\mathbb{R}$, and the result is the value in the exponent of (B.14). In the statistics literature the function $\boldsymbol{\pi}(\Theta) = (\pi_0(\Theta), \ldots, \pi_{K-1}(\Theta))$ is called the *natural parameter* and

$$\Pi = \left\{\pi \in \mathbb{R}^K : \int_{\mathcal{X}} h(x) \exp\left[\boldsymbol{\pi}(\theta)^{\mathsf{t}}\mathbf{t}(x)\right] \mathrm{d}\nu(x) < \infty\right\}$$

the *natural parameter space*. The function $\mathbf{t}(X)$ is a *sufficient statistic* – the *natural sufficient statistic* – and is usually denoted $T(X)$. We adopt this notation in the sequel. We make two further abuses of notation letting t denote a particular realization of $\mathbf{t}(X)$ and θ a particular realization of $\boldsymbol{\pi}(\Theta)$.

Lemma 2. (Schervish [6], p. 103) If X has an exponential family distribution, then so does the natural sufficient statistic $T(X)$ and the natural parameter space for T is the same as that for X. In particular, there exists a dominating measure $\nu_{\mathcal{T}}$ such that

$$\frac{\mathrm{d}P_{\Theta,T}}{\mathrm{d}\nu_{\mathcal{T}}}(t) = c(\theta)\exp(\theta^{\mathsf{t}}t)$$

Theorem 22. (Schervish [6], p. 105)
Let the density of $T(X)$ with respect to a measure $\nu_{\mathcal{T}}$ be $c(\theta)\exp(\theta^{\mathsf{t}}t)$. If $\phi : \mathcal{T} \to \mathbb{R}$ is measurable and

$$\int |\phi(t)| \exp(\theta^{\mathsf{t}}t)\, \mathrm{d}\nu_{\mathcal{T}}(t) < \infty$$

then

$$f(z) = \int \phi(t)\exp(z^{\mathsf{t}}t)\, \mathrm{d}\nu_{\mathcal{T}}(t)$$

is an analytic function of z in the region where the real part of z is interior to the natural parameter space, and

$$\frac{\partial}{\partial z_k} f(z) = \int t_k \phi(t)\exp(z^{\mathsf{t}}t)\, \mathrm{d}\nu_{\mathcal{T}}(t)$$

Theorem 22 allows us to calculate moments of sufficient statistics in exponential families by taking derivatives of the function $\log c(\theta)$.

Example 16. Let $\phi(t) = 1$. Then,

$$\begin{aligned}\mathsf{E}_\Theta(T_k) &= \int c(\theta) t_k \exp(\theta^\mathsf{t} t)\, \mathrm{d}\nu_\mathcal{T}(t) = c(\theta)\frac{\partial}{\partial\theta_k}\int \exp(\theta^\mathsf{t} t)\, \mathrm{d}\nu_\mathcal{T}(t) \\ &= c(\theta)\frac{\partial}{\partial\theta_k}\frac{1}{c(\theta)} \qquad \text{(by Lemma 2)} \\ &= -\frac{1}{c(\theta)}\frac{\partial c(\theta)}{\partial\theta_k} = -\frac{\partial}{\partial\theta_k}\log c(\theta).\end{aligned}$$

B.3.4 Maximum Likelihood Estimation

In exponential families, there is a simple method for finding MLEs in most cases. The logarithm of the likelihood function will be $\log L(\theta)$

References

1. Myoung An and Richard Tolimieri. *Group Filters and Image Processing.* Psypher Press, Boston, 2003.
2. Paul Billings. Poisson noise model. personal notes, 2001.
3. John B. Fraleigh. *A First Course in Abstract Algebra.* Addison-Wesley Publishing Company, fifth edition, 1994.
4. R. Horn and C. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.
5. Richard Paxman, Brian Thelen, and John Seldin. Phase-diversity correction of turbulence-induced space-variant blur. *Optics Letters*, 1994.
6. Mark J. Schervish. *Theory of Statistics.* S-V, 1995.
7. Richard Tolimieri and Myoung An. *Time-Frequency Representations.* Birkhäuser, Boston, 1998.