Groups and Symmetries in Numerical Linear Algebra

Hans Z. Munthe-Kaas[1]

Department of Mathematics, University of Bergen,
Postbox 7803, N-5020 Bergen, Norway
hans.munthe-kaas@uib.no

Summary. Groups are fundamental objects of mathematics, describing symmetries of objects and also descrtibing sets of motions moving points in a domain, such as translations in the plane and rotations of a sphere. The topic of these lecture notes is applications of group theory in computational mathematics. We will first cover fundamental properties of groups and continue with an extensive discussion of commutative (abelian) groups and their relationship to computational Fourier analysis. Various numerical algorithms will be discussed in the setting of group theory. Finally we will, more briefly, discuss generalisation of Fourier analysis to non-commutative groups and discuss problems in linear algebra with non-commutative symmetries. The representation theory of non-commutative finite groups is used as a tool to efficiently solve linear algebra problems with symmetries, exemplified by the computation of matrix exponentials.

1 Introduction

'Symmetry' is a vaguely defined concept deeply rooted in nature, physics, biology, art, culture and mathematics. In everyday language it refers to a harmonious proportion and balance. In mathematics, the symmetries of an object are more precisely defined as a set of transformations leaving the object invariant. Examples in art are tessellations and mosaics invariant under translations and reflections. In mechanics, symmetry can refer to invariance of a Lagrangian function under transformations such as spatial rotations and translation in time, and the famous theorem of Emmy Noether relates such symmetries to conservation laws. Sophus Lie (1842-1899) revolutionised the theory of differential equations by considering the symmetries sending solution curves to other solutions. A huge part of signal processing and Fourier analysis is based on invariance of linear operators under time– or space translations. Classical Fourier analysis extends to non-commutative harmonic analysis and group representation theory when the symmetry transformations do not commute (when $ab \neq ba$).

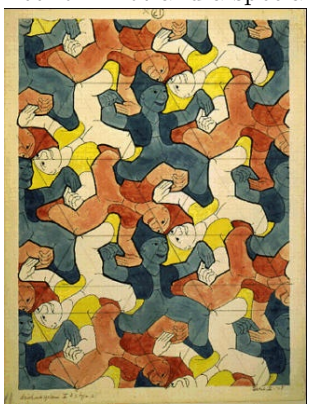
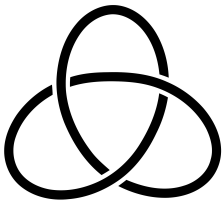
When designing an algorithm for solving some computational problem, it is usually a good idea to look for the symmetries of the problem. An algorithm which preserves symmetries often results in more accurate or stable numerical computations, and potentially also leads to huge savings in terms of time and space consumption. For these reasons, knowledge of the mathematics of symmetries (group theory) should be very important for students of computational science.

In these lectures we will focus our attention to applications of group theory in numerical linear algebra, Fourier analysis and signal processing. We will in particular focus on a unified treatment of classical Fourier analysis, based on translational invariance in space or time (commutative symmetries), and continue with a treatment of non-commutative groups of reflection symmetries such as the symmetries generated by the mirrors in a kaleidoscope. This is applied to fast solution of boundary value problems, computation of matrix exponentials and applications to sampling theory and numerical computations on regular lattices.

It is our goal that most of the material in these lectures should be accessible to advanced undergraduate students in applied and computational mathematics, requiring only basic knowledge of linear algebra and calculus, and not any prior knowledge of group theory nor abstract algebra.

1.1 Motivation for the main topics of the lectures

Consider the objects below, a 13'th century mosaic from Alhambra, a tesselation by Maurice Escher, a three-foil knot and a special matrix:

$$A = \begin{pmatrix} a_0 & a_2 & a_1 & a_3 & a_4 & a_5 \\ a_1 & a_0 & a_2 & a_5 & a_3 & a_4 \\ a_2 & a_1 & a_0 & a_4 & a_5 & a_3 \\ a_3 & a_5 & a_4 & a_0 & a_1 & a_2 \\ a_4 & a_3 & a_5 & a_2 & a_0 & a_1 \\ a_5 & a_4 & a_3 & a_1 & a_2 & a_0 \end{pmatrix}$$

Quiz:

1. Do any of these objects have the same symmetries?
2. How can we compute the eigenvalues and eigenvectors of A?

In order to try to answer 1) we must define what we mean by 'invariance under a set of transformations'. The two tessellations (Alhambra and Escher) can be seen to be invariant under certain Euclidean (rigid) motions of the plane. The exact group of symmetries depends on whether or not one considers the colouring, or just the shapes. Without regarding the colouring, they are both invariant under 120^o rotations in certain points and translations in two different directions. In a certain sense, which has not yet been made clear, it seems as the two tessellations have the same symmetries.

The tre-foil knot, understood as a curve in $\mathbb{R}^3$ is invariant under transformation α being a 120^o rotation around the centre as well as transformation β being a 180^o rotation around the vertical line through the plane of the knot. Any product of α, β and their inverses are also symmetries, so that the total group of symmetries becomes $\{1, \alpha, \alpha^2, \beta, \alpha\beta, \alpha^2\beta\}$, where 1 denotes the identity transformation (do nothing). We can verify that α and β satisfy the relations

$$\alpha^3 = \beta^2 = \alpha\beta\alpha\beta = 1. \tag{1}$$

The symmetries of A are less evident. We can verify that A *commutes* with some particular (permutation) matrices; we have that $P_i A = AP_i$, or equivalently $A = P_i A P_i^{-1}$ for both

$$P_1 = \begin{pmatrix} 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 1&0&0&0&0&0\\ 0&0&0&0&0&1\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0 \end{pmatrix} \quad \text{and} \quad P_2 = \begin{pmatrix} 0&0&0&0&0&1\\ 0&0&0&0&1&0\\ 0&0&0&1&0&0\\ 0&0&1&0&0&0\\ 0&1&0&0&0&0\\ 1&0&0&0&0&0 \end{pmatrix}$$

and hence also for any P_i which is given as a product of these and their inverses, $P_i \in \{I, P_1, P_1^2, P_2, P_1P_2, P_1^2P_2\}$. It can be shown that P_1 and P_2 satisfy exactly the same relations (1) as α and β. However, to claim that the three-foil and A have the same symmetries, we need to abstract the notion of a symmetry group and the action of a group on a set so that we can discuss the properties of the abstract group independently of the concrete transformations the group performs on a given object. We start Section 3 with the modern definition of a group and group actions.

Now, back to Question 2 in the Quiz. Once we understand that A commutes with a set of matrices, we are closer to finding the eigenvectors. From elementary linear algebra we know that matrices with a common complete set of eigenvectors do commute, and conversely, under quite mild conditions (e.g. distinct eigenvalues), commuting matrices share a complete set of common eigenvectors. However, P_1 and P_2 do not have distinct eigenvalues and furthermore they do not commute among themselves, since $P_1P_2 = P_2P_1^{-1}$,

so we cannot possibly find a complete set of common eigenvectors. However, *groups representation theory* provides something almost as good; a complete list of *irreducible representations*, which yields a particular basis change such that A becomes block diagonalised. Applications of this theory is the topic of Chapter 4. A very important and special case of structured matrices appears in classical Fourier analysis, where A commutes with a set of matrices P_i such that also $P_iP_j = P_jP_i$. If the set of such matrices is sufficiently large, we find a complete set of common eigenvectors for all these P_i and these also form a complete set of eigenvectors for A. In the case of finite dimensional A, the mathematical analysis becomes particularly simple, the common eigenvectors are exponential functions and the change of basis is given by the Discrete Fourier Transform. Also Fourier series on the continuous circle and the Fourier transform for functions on the real line can be described within a common group theoretical framework. A detailed study of these cases, with applications, is the topic of Chapter 3.

2 Prelude: Introduction to Group Theory

Before going into a detailed study of abelian (= commutative) groups, we will for future reference introduce some general concepts in group theory. A detailed understanding of this chapter is not needed for the applications in Fourier analysis and sampling theory, where these concepts become somewhat simpler. I suggest that this chapter is read lightly in first pass, and studied more carefully whenever needed later.

2.1 Groups and actions

Definition 1 (Group). *A group is a set G with a binary operation $\cdot\colon G\times G \to G$, called the group product, such that*

1. *The product is associative, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$ for all $x, y, z \in G$.*
2. *There exits an identity element $\mathbf{1} \in G$ such that $x \cdot \mathbf{1} = \mathbf{1} \cdot x = x$ for all $x \in G$.*
3. *Every element $x \in G$ has an inverse $x^{-1} \in G$ such that $x \cdot x^{-1} = \mathbf{1}$.*

Sometimes we write the group product without the dot, as xy instead of $x\cdot y$. The special groups where $x\cdot y = y\cdot x$ for all $x, y \in G$ are called *commutative* or *abelian groups*. In the case of abelian groups we will often (but not always) write $+$ instead of $\cdot$, $-x$ instead of x^{-1} and $\mathbf{0}$ instead of $\mathbf{1}$.

Example 1. We list some common groups that we will encounter later.

- **Zero group $\{\mathbf{0}\}$.** This is the trivial additive abelian group consisting of just the identity element. Sometimes we write $\mathbf{0}$ instead of $\{\mathbf{0}\}$.
- **Additive group of reals $(\mathbb{R}, +)$.**

- **Additive circle group** $(T, +)$**:** This is the real numbers $[0, 1)$ under addition modulo 1. It is also defined as $T = \mathbb{R}/\mathbb{Z}$, a quotient group (see below). The name T refers to this being the 1-dimensional torus.
- **Additive group of integers** $(\mathbb{Z}, +)$**.**
- **Additive cyclic group** $(\mathbb{Z}_n, + \bmod n)$ This consists of the integers $\{0, 1, \ldots, n-1\}$ with group operation being addition modulo n, and is also given as the quotient $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$.
- **Identity group** $\{\mathbf{1}\}$**:** This is the trivial multiplicative group consisting of just the identity element. The two trivial groups $\{\mathbf{0}\}$ and $\{\mathbf{1}\}$ are *isomorphic* (abstractly the same group). Isomorphisms are discussed below.
- **Multiplicative cyclic group** C_n**:** This consists of the complex n'th roots of 1 under multiplication, $\{e^{2\pi ij/n}\}_{j=0}^{n-1}$, and it is isomorphic to $\mathbb{Z}_n$.
- **Multiplicative circle group** $\mathbb{T}$**:** The multiplicative group of complex numbers with modulus 1, $\mathbb{T} = \{e^{2\pi i\theta}\}$ for $\theta \in [0, 1)$, isomorphic to T.
- **Dihedral group** D_n**:** The symmetries of an regular n-gon form a group called the *Dihedral group*, D_n. In particular D_3 are the 6 symmetries of an equilateral triangle (three rotations and three reflected rotations). Abstractly, D_n is generated by two elements α and β satisfying the relations

$$\alpha^n = \beta^2 = \alpha\beta\alpha\beta = \mathbf{1}. \tag{2}$$

  The last relation is equivalent to $\alpha\beta = \beta\alpha^{-1}$, thus this is our first example of a non-commutative group.
- **General linear group GL**(V)**:** For a vector space V (e.g. $V = \mathbb{R}^n$), this group consists of all invertible linear operators on V (e.g. all invertible real $n \times n$ matrices), and the group product is composition of linear operators (matrix product).
- **Orthogonal group** $O(V)$ The set of all orthogonal matrices ($A^T = A^{-1}$ in $\mathrm{GL}(V)$.
- **The symmetric group** S_n This is the group of all permutations of n objects, thus S_n has $n!$ elements.
- **The alternating group** A_n The subset of all *even* permutations of n objects, with $n!/2$ elements.
- **Lie groups** There are two main classes of groups, discrete groups and Lie groups. Lie groups are groups which also has a differentiable structure, so that one can define continuous and smooth families of transformations. Among the examples above, the Lie groups are $(\mathbb{R}, +)$, $(T, +)$, $\mathbb{T}$, $\mathrm{GL}(V)$ and $O(V)$. The remaining groups are discrete.

We want to connect an abstract group to a concrete group of transformations of some 'object'. This is done by the concept of a *group action*.

Definition 2 (Group action). *A group G acts*[1] *on a set X if there is a function ('group action') $\cdot : G \times X \to X$ satisfying*

[1] This definition is more precisely called a *left action*.

$$1 \cdot x = x \quad \textit{for all } x \in X$$
$$g \cdot (h \cdot x) = (g \cdot h) \cdot x \quad \textit{for all } g, h \in G,\ x \in X.$$

Example 2. We define an action $\cdot\colon D_n \times \mathbb{C} \to \mathbb{C}$ on the generators of D_n as $\alpha \cdot z = e^{2\pi i/n} z$ (rotation counterclockwise through the angle $2\pi/n$) and $\beta \cdot z = \overline{z}$ (complex conjugation). These two symmetry operations are compatible with (2), we have $\alpha \cdot (\alpha \cdots (\alpha \cdot z)) = z$ (n times application of α), $\beta \cdot (\beta \cdot z) = z$ and $\alpha \cdot (\beta \cdot (\alpha \cdot (\beta \cdot z))) = z$. Therefore, we can extend this to a group action of D_n on $\mathbb{C}$. Consider the regular n-gon in $\mathbb{C}$, with vertices in the n'th roots of unity $C_n \subset \mathbb{C}$. It is straightforward to check that the set of vertices of the n-gon is invariant under this action.

Example 3. The dihedral group D_3 acts on the set of all 6×6 matrices as $\alpha \cdot X = P_1 X P_1^{-1}$ and $\beta \cdot X = P_2 X P_2^{-1}$, where P_1 and P_2 are given above.

Example 4. Any group G can act on itself in various ways. We can let G act on itself by left multiplication $L_g g' := gg'$ or by right multiplication $R_g g' = g' g^{-1}$ or by conjugation $\mathrm{Conj}_g\, g' := gg'g^{-1}$. Check that all these are well defined actions.

Definition 3 (Types of actions). *An action $\cdot\colon G \times X \to X$ is:*

- transitive *if for any pair $x, y \in X$ there exists a $g \in G$ such that $g \cdot x = y$,*
- free *if the identity $\mathbf{1} \in G$ is the only group element which has a fixed point on X, i.e. for $g \in G$ there exists an an $x \in X$ such that $g \cdot x = x$ only if $g = \mathbf{1}$,*
- regular *if it is both free and transitive,*
- effective *if whenever $g, h \in G$ and $g \neq h$ there exist an $x \in X$ such that $g \cdot x \neq h \cdot x$.*

Exercise 1.

1. Show that *free* $\Rightarrow$ *effective*.
2. Is the action of D_n on C_n defined in Example 2 regular?
3. Show that if an action is regular, then there is a 1–1 correspondence between elements of G and X. Find a subset of $2n$ points in $\mathbb{C}$ on which the action of D_n defined in Example 2 is regular.

2.2 Subgroups and quotients

It is important to understand some basic ways of obtaining groups from other groups, by decompositions (subgroups and quotients) and compositions (direct– and semidirect products).

Definition 4 (Subgroup). *A non-empty subset $H \subset G$ which is closed under the group product and inversion is called a subgroup, denoted $H < G$.*

A subgroup $H < G$ decomposes G into subsets called *cosets*, these can be defined from *left* or from *right*:

$$gH := \{gh \colon g \in G, h \in H\}$$
$$Hg := \{hg \colon g \in G, h \in H\}.$$

Note that for $g, g' \in G$ we have either $gH = g'H$ or $gH \cap g'H = \varnothing$, so the collection of all left (or all right) cosets form a disjoint partition of G.

Example 5. The dihedral group $D_3 = \{\mathbf{1}, \alpha, \alpha^2, \beta, \beta\alpha, \beta\alpha^2\}$ has four subgroups. The *trivial* subgroup consists of just the identity $\{\mathbf{1}\} < D_3$, and the *improper* subgroup is the whole group $D_3 < D_3$. The two proper and non-trivial subgroups are $H = \{\mathbf{1}, \alpha, \alpha^2\}$ and $\widetilde{H} = \{\mathbf{1}, \beta\}$. The left cosets of H are H and $\beta H = \{\beta, \beta\alpha, \beta\alpha^2\}$, and these form a disjoint partition $D_3 = H \cup \beta H$. The right cosets are H and $H\beta = \{\beta, \alpha\beta, \alpha^2\beta\} = \{\beta, \beta\alpha^2, \beta\alpha\} = \beta H$. The three left cosets of $\widetilde{H}$ are $\widetilde{H}$, $\alpha\widetilde{H} = \{\alpha, \alpha\beta\} = \{\alpha, \beta\alpha^2\}$ and $\alpha^2\widetilde{H} = \{\alpha^2, \beta\alpha\}$. The three right cosets are $\widetilde{H}$, $\widetilde{H}\alpha = \{\alpha, \beta\alpha\}$ and $\widetilde{H}\alpha^2 = \{\alpha^2, \beta\alpha^2\}$. Note that all left cosets of H are also right cosets, $gH = Hg$ for all $g \in G$. This is *not* the case for $\widetilde{H}$.

Definition 5 (Normal subgroup). *A subgroup $H < G$ is called normal if $gH = Hg$ for every $g \in G$. We write a normal subgroup as $H \lhd G$.*

The collection of cosets of a subgroup $H < G$ can be turned into a group if and only if H is normal.

Definition 6 (Quotient group). *For a normal subgroup $H \lhd G$ we define the quotient group G/H as a group where the elements of G/H are the cosets gH and the product of two cosets are defined as*

$$gH \cdot g'H = gg'H,$$

where gg' is the product of g and g' in G.

Example 6. Continuing Example 5, we obtain the quotient group D_3/H with two elements H and βH and the multiplication rule $H \cdot H = H$, $\beta H \cdot H = H \cdot \beta H = \beta H$ and $\beta H \cdot \beta H = H$. The group D_3/H can be identified with the group $C_2 = \{1, -1\} \subset \mathbb{R}$ with multiplication as group product, meaning that if we define the map $\varphi \colon D_3/H \to C_2$ as $\varphi(H) = 1$, $\varphi(\beta H) = -1$, we find that $\varphi(g_1 g_2) = \varphi(g_1)\varphi(g_2)$ for $g_1, g_2 \in \mathrm{d}_3/H$. This is an example of a *group isomorphism*, which identifies the two groups as being abstractly the same.

2.3 Homomorphisms and exact sequences

Definition 7 (Group homomorphism). *Let H and G be two groups. A homomorphism is a map $\varphi \colon H \to G$ such that$\varphi(h_1 \cdot h_2) = \varphi(h_1) \cdot \varphi(h_2)$ for all $h_1, h_2 \in H$. The set of all such homomorphisms is denoted* $\hom(H, G)$.

Definition 8 (Kernel and image). *The* kernel *and* image *of* $\varphi \in \hom(H, G)$ *are defined as*

$$\ker(\varphi) = \{h \in H \colon \varphi(h) = \mathbf{1}\}$$
$$\operatorname{im}(\varphi) = \{g \in G \colon g = \varphi(h) \text{ for some } h \in H.\}$$

Definition 9 (Epimorphism, monomorphism and isomorphism). *If* $\ker(\varphi) = \mathbf{1}$ *then* φ *is* injective*, meaning that* $\varphi(h) = \varphi(h') \Rightarrow h = h'$*. If* $\operatorname{im}(\varphi) = G$ *we say that* φ *is* surjective *(onto* G*). A surjective homomorphism is called an* epimorphism*, denoted* $\phi \in \operatorname{epi}(G_1, G_2)$ *and an injective homomorpism is called a* monomorphism*, denoted* $\phi \in \operatorname{mono}(G_1, G_2)$*. A homomorphism which is both injective and surjective is called an* isomorphism*, denoted* $\phi \in \operatorname{iso}(G_1, G_2)$*. If there exists an isomorphism between* G_1 *and* G_2 *we write* $G_1 \simeq G_2$ *and say that* H *and* G *are isomorphic groups, meaning that they are structurally identical.*

Exercise 2. Show that the additive group of real numbers $(\mathbb{R}, +)$ and the multiplicative group of positive reals $(\mathbb{R}^+, \cdot)$ are isomorphic. Hint: use the exponential map.

Exercise 3. Let $\varphi \in \hom(H, G)$. Show that $\ker(\varphi) \lhd H$ and that $\operatorname{im}(\varphi) < G$.

Definition 10 (Coimage). *Let* $\varphi \in \hom(H, G)$*. Since* $\ker(\varphi) \lhd G$ *(always normal subgroup), we can form the quotient. This is called the* coimage

$$\operatorname{coim}(\varphi) := H/\ker(\varphi).$$

Definition 11 (Cokernel). *Let* $\varphi \in \hom(H, G)$*. If* $\operatorname{im}(\varphi) \lhd G$ *we can form the quotient* $C = G/\operatorname{im}(\varphi)$*. This is called the* cokernel *of* φ*.*

It is very useful to present homomorphisms in terms of *exact sequences*.

Definition 12 (Exact sequence). *A sequence*

$$G_0 \xrightarrow{\varphi_1} G_1 \xrightarrow{\varphi_2} G_2 \xrightarrow{\varphi_3} \cdots \xrightarrow{\varphi_n} G_n$$

of groups and group homomorphisms is called an exact sequence *if* $\operatorname{im}(\varphi_i) = \ker(\varphi_{i+1})$ *for every* i*.*

Let $\mathbf{1}$ denote the trivial group containing just the identity element. An exact sequence

$$\mathbf{1} \longrightarrow H \xrightarrow{\varphi} G$$

indicates that $\varphi \in \hom(H, G)$ is a monomorphism. To see this we note that the only homomorphism in $\hom(\mathbf{1}, H)$ is the trivial map $\mathbf{1} \mapsto \mathbf{1}$, thus $\ker(\varphi) = \mathbf{1}$. We will frequently also use a hooked attow $H \xhookrightarrow{\phi} G$ to indicate that ϕ is a monomorphism.

Exactness of the sequence

$$H \xrightarrow{\varphi} G \longrightarrow \mathbf{1}$$

means that $\varphi \in \hom(H, G)$ is an epimorphism, since the only homomorphism in $\hom(G, \mathbf{1})$ is the map sending G to $\mathbf{1}$ and hence $\operatorname{im}(\varphi) = G$. We will also use a double arrow $H \xrightarrow{\phi} \!\!\!\!\rightarrow G$ to visualise $\phi \in \operatorname{epi}(H, G)$. The exact sequence

$$\mathbf{1} \longrightarrow H \xrightarrow{\varphi} G \longrightarrow \mathbf{1}$$

means that φ is both epi- and mono- and hence it is an isomorphism and $H \simeq G$.

Definition 13 (Short exact sequence). *A* short exact sequence *is an exact sequence of the form*

$$\mathbf{1} \longrightarrow H \xrightarrow{\varphi_G} G \xrightarrow{\varphi_K} K \longrightarrow \mathbf{1} \ . \tag{3}$$

This indicates that $H \simeq \operatorname{im}(\varphi_G) \lhd G$ and that $K \simeq G/\operatorname{im}(\varphi_G) = \operatorname{coker}(\varphi_G)$, or by a slight abuse of notation (identification by isomorphisms) we write this as $H \lhd G$ and $K = G/H$.

We ask the reader to think through the meaning of the short exact sequence carefully! Since φ_G is injective, it must define an isomorphism between H and its image in G. To see that $H \lhd G$ is a normal subgroup and that φ_K is a projection of G onto G/H, we compute for $g \in G$ and $h \in \operatorname{im}(\varphi_G) = \ker(\varphi_K)$:

$$\varphi_K(gh) = \varphi_K(g)\varphi_K(h) = \varphi_K(g)\mathbf{1} = \mathbf{1}\varphi_K(g) = \varphi_K(h)\varphi_K(g) = \varphi_K(hg),$$

so all elements of gH and Hg are sent to the same element in C. Furthermore, if $\varphi_K(g) = \varphi_K(g')$, we must have $\mathbf{1} = \varphi_K(g')^{-1}\varphi_K(g) = \varphi_K(g'^{-1}g)$ thus $g'^{-1}g = h \in \ker(\varphi_K)$ and $g = g'h$. We conclude that $\varphi_K(g) = \varphi_K(g')$ if and only if g and g' belong to the same left and right coset. Finally we check that $\varphi_K(gH \cdot g'H) = \varphi_K(gg'H)$ and hence $K \simeq G/H$.

Example 7. Let D_n be the dihedral group and $C_n = \{e^{2\pi ij/n}\}_{j=0}^{n-1} \subset \mathbb{C}$ the cyclic group of n elements, identified with the multiplicative group of complex n'th roots of unity. There is a short exact sequence

$$\mathbf{1} \longrightarrow C_n \xrightarrow{\varphi_1} D_n \xrightarrow{\varphi_2} C_2 \longrightarrow \mathbf{1} \ , \tag{4}$$

where $\varphi_1(e^{2\pi ij/n}) = \alpha^j$ and $\varphi_2(\alpha^j) = 1$, $\varphi_2(\beta\alpha^j) = -1$.

An exact sequence of the form

$$\mathbf{1} \longrightarrow K \xrightarrow{\ker(\varphi)} H \xrightarrow{\varphi} G \xrightarrow{\operatorname{coker}(\varphi)} C \longrightarrow \mathbf{1} \tag{5}$$

indicates that $K \simeq \ker(\varphi)$ and $C \simeq \operatorname{coker}(\varphi)$. Note that we have now called the injection arrow of K into H for the kernel of φ and the projection arrow from G onto C the cokernel of φ. This definition of kernels and cokernels as arrows rather than objects (groups) is standard in category theory language, where any mathematical property of an object is defined in terms of arrows into and out of the object. We will call both the arrows and their images for kernels and cokernels. If we really need to distinguish, we call the arrow *'(co)kernel homomorphism'* and the group *'(co)kernel group'*.

Definition 14 (Kernel and cokernel homomorphisms[2]). *The kernel homomorphism of* $\varphi \in \hom(H,G)$ *is defined as a monomorphism, denoted* $\ker(\varphi) \in \operatorname{mono}(K,H)$*, such that the image of* $\ker(\varphi)$ *is the kernel group of* φ*. The cokernel homomorphism of* $\varphi \in \hom(H,G)$ *is defined as an epimorphism, denoted* $\operatorname{coker}(\varphi) \in \operatorname{epi}(G,C)$*, such that the image of* $\operatorname{coker}(\varphi)$ *is the cokernel group of* φ*.*

Definition 15 (Image and coimage homomorphisms). *Let* $\varphi \in \hom(G_1,G_2)$ *and let* $K = G_1/\ker(\varphi)$ *be the coimage group. The coimage homomorphism is defined as an epimorphism* $\operatorname{coim}(\varphi) \in \operatorname{epi}(G_1,K)$ *and the image homomorphism is a monomorphism* $\operatorname{im}(\phi) \in \operatorname{mono}(K,G_2)$ *such that*

$$\varphi = \operatorname{im}(\varphi) \circ \operatorname{coim}(\varphi).$$

Note that these homomorphisms are defined up to an isomorphism of K, so there is a freedom in how to represent $K = G_1/\ker(\varphi)$. However, the image and coimage homomorphisms must be consistent with this choice. The following example is illustrating this point.

Example 8. This example should make the above discussion more familiar to computational scientists. Consider the set of all abelian groups $(\mathbb{R}^n,+)$ for all $n \in \mathbb{N}$ and the continuous homomorphisms between these. This is an example of a category[3], where $\mathbb{R}^n$ are 'objects' and homomorphisms are the 'arrows'. The set $\hom(\mathbb{R}^n,\mathbb{R}^m)$ can be identified with the set of $m \times n$ matrices

$$\hom(\mathbb{R}^n,\mathbb{R}^m) \approx \mathbb{R}^{m\times n}$$

and composition of homomorphisms is given as matrix products. A monomorphism is a matrix with full column-rank, and an epimorphism a matrix with full row-rank. The isomorphisms are the invertible matrices.

For $A \in \mathbb{R}^{m\times n}$ we want to compute the homomorphisms (matrices) $\operatorname{im}(A)$, $\operatorname{coim}(A)$, $\ker(A)$ and $\operatorname{coker}(A)$. Recall the singular value decomposition

[2] Check Wikipedia for a proper categorical definition of kernel and cokernel which only refers to properties of arrows.

[3] A category is a collection of objects and arrows between the objects such that the composition of an arrow from A to B and an arrow from B to C yields an arrow from A to C.

$$A = U\Sigma V,$$

where $\Sigma \in \mathbb{R}^{m\times k}$ is a diagonal matrix with non-negative diagonal elements $\sigma_i = \Sigma_{i,i}$ called singular values. We assume that $\sigma_i \geqslant \sigma_{i+1}$ and $\sigma_{k+1} = 0$, so there are k positive singular values. The two matrices $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal. We block up the three matrices as

$$U = \begin{pmatrix} U_1 \; U_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} \; \Sigma_{12} \\ \Sigma_{21} \; \Sigma_{22} \end{pmatrix}, \quad V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix},$$

where $U_1 \in \mathbb{R}^{m\times k}$, $U_2 \in \mathbb{R}^{m\times(m-k)}$, $\Sigma_{11} \in \mathbb{R}^{k\times k}$, $\Sigma_{12} \in \mathbb{R}^{k\times(n-k)}$, $\Sigma_{21} \in \mathbb{R}^{(n-k)\times k}$, $\Sigma_{22} \in \mathbb{R}^{(n-k)\times(n-k)}$, $V_1 \in \mathbb{R}^{k\times n}$ and $V_2 \in \mathbb{R}^{(n-k)\times n}$. The matrix Σ_{11} is diagonal with positive diagonal and Σ_{12}, Σ_{21} and Σ_{22} are all zero. Since U and V are orthogonal, their inverses are $U^{-1} = U^T$ and $V^{-1} = V^T$. From $A \in \hom(\mathbb{R}^n, \mathbb{R}^m)$ we get the four homomorphisms

$$\begin{aligned} \ker(A) &= V_2^T \in \mathrm{mono}(\mathbb{R}^{n-k}, \mathbb{R}^n) \\ \mathrm{coker}(A) &= U_2^T \in \mathrm{epi}(\mathbb{R}^m, \mathbb{R}^{m-k}) \\ \mathrm{coim}(A) &= V_1 \in \mathrm{epi}(\mathbb{R}^n, \mathbb{R}^k) \\ \mathrm{im}(A) &= U_1\Sigma_{11} \in \mathrm{mono}(\mathbb{R}^k, \mathbb{R}^m). \end{aligned}$$

We leave the verification of this to the reader. To check that the kernel and cokernel homomorphisms are correctly defined, you must verify that

$$\mathbf{0} \longrightarrow \mathbb{R}^{n-k} \xrightarrow{V_2^T} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m \xrightarrow{U_2^T} \mathbb{R}^{m-k} \longrightarrow \mathbf{0}$$

is an exact sequence.

To check the image and coimage, you must verify that the diagram

$$\begin{array}{ccccc} & & \mathbf{0} & & \\ & & \downarrow & & \\ \mathbb{R}^n & \xrightarrow{V_1} & \mathbb{R}^k & \longrightarrow & \mathbf{0} \\ & \searrow^{A} & \downarrow{\scriptstyle U_1\Sigma_{11}} & & \\ & & \mathbb{R}^m & & \end{array}$$

commutes (meaning that you get the same result if you follow different paths between two objects) and that the row and the column are exact.

The image- coimage factorisation is $A = (U_1\Sigma_{11})V_1$, where the left term has full column rank and the right has full row-rank. Such a factorisation is not unique, for any invertible $k \times k$ matrix X, we could instead do the factorisation as $A = (U_1\Sigma_{11}X)(X^{-1}V_1)$, which is another factorisation of A in a product of a matrix with full column rank and a matrix with full row-rank. The possibility of choosing X is expressed as *'defined up to isomorphisms'*.

Exercise 4. Repeat the example using the QR-factorisation instead of SVD.

2.4 Products of groups and split exact sequences

How can we construct more complicated groups from simpler ones? The two most important operations are called *direct product* and *semidirect product*.

Definition 16 (Direct product). *For two groups G and H we define their direct product $G \times H$ as a group defined on the set of pairs*

$$G \times H = \{(g, h) \colon g \in G, h \in H\}$$

with product defined componentwise

$$(g, h) \cdot (g', h') = (gg', hh').$$

Example 9. For additive abelian groups we write the direct product as $\oplus$ instead of $\times$. The abelian group $\mathbb{R}^2 = \mathbb{R} \oplus \mathbb{R}$ is defined on pairs of reals with the sum $(x, y) + (x', y') = (x + x', y + y')$ and $\mathbf{0} = (0, 0)$.

The semidirect product is a bit more involved. To motivate the definition, let us look at a particular group of all affine linear mappings on a vector space.

Example 10 (Affine group). Let $V = \mathbb{R}^n$ be a vector space. Any vector space is also an abelian group (by forgetting scalar multiplication), so we can let V act on itself by translation $v, w \mapsto v + w$. An other action is the linear action of $\mathrm{GL}(V)$ on V by matrix-vector product $A \cdot v = Av$. The affine action for $(A, b) \in \mathrm{GL}(V) \times V$ on V is given as

$$(A, b) \cdot v := Av + b.$$

What is the group structure on $\mathrm{GL}(V) \times V$ compatible with this action? We compute:

$$(A', b') \cdot ((A, b) \cdot v) = (A', b') \cdot (Av + b) = A'Av + A'b + b' = (AA', b' + A'b) \cdot v,$$

thus we obtain the group product

$$(A', b') \cdot (A, b) = (AA', b' + A'b).$$

The identity element is $(I, 0)$, where I is the identity matrix. This is an important example of a semidirect product. We write the affine group as $\mathrm{Aff}(V) := \mathrm{GL}(V) \rtimes V$.

Definition 17 (Semidirect product). *A semidirect product is defined from two groups G and H and an action $\cdot \colon G \times H \to H$. We write the products in G and H as gg' and hh', and the action as $g \cdot h$. The semidirect product of G and H, written $G \rtimes H$ is the set of pairs (g, h) with the product*

$$(g, h) \cdot (g', h') := (gg', h(g \cdot h').$$

The direct product is the special case of semidirect product where $g \cdot h = h$ for all g and h.

Example 11. Let C_2 act on C_n by complex conjugation, $(-1) \cdot z = \overline{z}$ for all $z \in C_n$. We claim that $D_n \simeq C_2 \rtimes C_n$, with respect to this action. We have that $C_2 \times C_n = \{(\pm 1, \omega^j)\}_{j=0}^{n-1}$ where $\omega = e^{2\pi i/n}$. Let $\alpha = (1, \omega)$ and $\beta = (-1, 1)$. We ask the reader to verify that these two elements generate $C_2 \rtimes C_n$ and satisfy the relations (2).

From this example and (4) we might be tempted to believe that (3) implies $G \simeq G/H \rtimes H$. This is, however, NOT true in general.

Example 12. We have a short exact sequence

$$\mathbf{1} \longrightarrow \mathbb{Z}_2 \xrightarrow{\cdot 4} \mathbb{Z}_8 \xrightarrow{\text{mod } 4} \mathbb{Z}_4 \longrightarrow \mathbf{1} \ ,$$

however, there is no way $\mathbb{Z}_8$ can be written as a direct or semidirect product of $\mathbb{Z}_2$ and $\mathbb{Z}_4$. On the other hand, we have

$$\mathbf{1} \longrightarrow \mathbb{Z}_3 \xrightarrow{\cdot 4} \mathbb{Z}_{12} \xrightarrow{\text{mod } 4} \mathbb{Z}_4 \longrightarrow \mathbf{1} \ ,$$

corresponding to a decomposition $\mathbb{Z}_{12} \simeq \mathbb{Z}_3 \times \mathbb{Z}_4$. The difference between these two cases is that the latter *splits* in the sense defined below. We return to this example in Section 3.1.

Definition 18 (Split exact sequence). *The short exact sequence*

$$\mathbf{1} \longrightarrow H \xrightarrow{\varphi_G} G \underset{\varphi_s}{\overset{\varphi_K}{\rightleftarrows}} K \longrightarrow \mathbf{1} \tag{6}$$

is called right split *if there exists a homomorphism $\varphi_s \colon K \to G$ such that the composition $\varphi_K \circ \varphi_s = \mathrm{Id}_K$ (the identity map). The exact sequence*

$$\mathbf{1} \longrightarrow H \underset{\varphi_s}{\overset{\varphi_G}{\rightleftarrows}} G \xrightarrow{\varphi_K} K \longrightarrow \mathbf{1} \tag{7}$$

is called left split *if there exists a homomorphism $\varphi_s \colon G \to H$ such that $\varphi_s \circ \varphi_G = \mathrm{Id}_H$.*

Theorem 1.

- *A semidirect product decomposition $G = K \rtimes H$ is equivalent to a right split short exact sequence.*
- *A direct product decomposition $G = H \times K$ is equivalent to a left split short exact sequence.*
- *Any left split short exact sequence is also right split (but not vice versa).*

Before we prove this theorem, let us discuss decomposition of G with respect to *any* subgroup $H < G$. As a set, G decomposes disjoint union in right cosets $G = \cup_i H k_i$, where the subset $\{k_i\} \subset G$ consists of exactly *one* element k_i from each coset Hg. Such k_i are called *coset representatives*. Hence, we have a unique factorisation $g = hk$, $h \in H$, $k \in \{k_i\}$, identifying G and $\{k_i\} \times H$ as *sets*. An important question is whether or not the coset representatives can be chosen in a canonical (natural) way, so that this identification also carries along a (semidirect product) group structure.

Proof (Theorem 1). Given the right split short exact sequence (6). For any function $\varphi_s \colon K \to G$ such that $\varphi_K \circ \varphi_s = \mathrm{Id}_K$ we have that $\mathrm{im}(\varphi_s) \subset G$ is a set of coset representatives. This defines a set-mapping

$$\xi \colon K \times H \to G, \quad (k, h) \mapsto g = \varphi_G(h)\varphi_s(k),$$

with inverse given as $k = \varphi_K(g)$, and we find h from $\varphi_G(h) = g\varphi_s(k)^{-1}$. Now let $g = \xi(k, h)$ and $g' = \xi(k', h')$ be arbitrary. If φ_s is a homomorphism

$$gg' = \varphi_G(h)\varphi_s(k)\varphi_G(h')\varphi_s(k') = \varphi_G(h)\varphi_s(k)\varphi_G(h')\varphi_s(k^{-1})\varphi_s(kk').$$

The K-part of gg' is $\varphi_K(gg') = kk'$, hence $\varphi_G(h)\varphi_s(k)\varphi_G(h')\varphi_s(k^{-1}) \in \mathrm{im}(H)$, and we conclude that $k \cdot h \colon K \times H \to H$ defined such that

$$\varphi_G(k \cdot h) = \varphi_s(k)\varphi_G(h')\varphi_s(k^{-1})$$

is a well-defined action of K on H. We see that $K \rtimes H$ with the semidirect product $(k, h)(k', h') = (kk', h(k \cdot h'))$ is isomorphic to G.

Conversely, it is straightforward to check that if $G = K \rtimes H$ we have that $\varphi_G(h) = (1, h)$, $\varphi_K(k, h) = k$ and $\varphi_s(k) = (k, 1)$ defines a right split short exact sequence, and we have established the first point in the theorem.

To prove the second point, we assume the existence of a left split short exact sequence (7). We want to factor $g = hk$ for $h \in \mathrm{im}(\varphi_G)$ and k in some set of coset representatives. We find $h = \varphi_G \circ \varphi_s(g)$ and $k = h^{-1}g$, thus $k = \sigma(g)$ where

$$\sigma(g) := \left(\varphi_G \circ \varphi_s(g)\right)^{-1} g.$$

If φ_s is a homomorphism we can check that $\sigma(hg) = \sigma(g)$ and $H\sigma(g) = Hg$, hence σ picks a unique representative from each coset. We conclude that the mapping $\psi \colon G \to K \times H$, $g \mapsto (\varphi_K(g), \varphi_s(g))$ is an invertible set function. It is clearly a group homomorphism and hence also an isomorphism. We conclude that $G \simeq K \times H$. Conversely it is easy to check that any direct product $G = K \times H$ is left split.

Since a direct product is also a semi-direct product, we conclude the third point that left split implies right split. □

2.5 Domains in computational mathematics

It is time to be a bit more philosophical and less technical on the role of groups in computational mathematics. A fundamental question is what do we mean by a 'Domain'? More specifically, what abstract properties do we require from the 'domain' of a differential equation? Over the last century, mathematicians agree that the notion of a *(differentiable) manifold* is a powerful abstract setting for a general theory of differential equations. Manifolds are sets endowed with with a 'differentiable structure', we have *points* in the domain as well as *tangents* at a given point. Points have a position (coordinates), tangents are velocities, specifying a direction and a speed. The most important properties of manifolds is that they support scalar functions (real or complex), and derivations of these in directions specified by tangent vectors. Examples of manifolds are the familiar space $\mathbb{R}^n$, but also spaces like the surface of a sphere. In $\mathbb{R}^n$ both points and tangents are vectors in $\mathbb{R}^n$, but the spherical surface is a good example of a space where these two different things should not be confused!

The mathematical definition of a manifold does not have enough structure to be suitable as an abstraction of a computational domain. Manifolds have tangents, but there is no (practical) way of moving in the direction of a given tangent. In pure mathematics motions arise from the concept of the solution operator (flow map) of a tangent vector field (= solution of ordinary differential equations, ODEs), but in computations one cannot assume that differential equations can be solved exactly. For doing computations we need to specify a set of motions which we assume can be computed fast and accurately. Here groups and group actions come in handy! For the purpose of solving ODEs, it has turned out to be very useful to study algorithms on *homogeneous spaces*, which are domains together with a transitive action of a Lie group. One example is $\mathbb{R}^n$, which acts on itself by translations, the basic motions are obtained by adding vectors. An other example is the surface of a sphere, under the action of the orthogonal group. A substantial theory of numerical Lie group integration has been developed over the last two decades [14].

Among the homogeneous spaces, abelian Lie groups are the most important for practical computations. Most modelling in physics and engineering take place in $\mathbb{R}^n$ or subdomains of this. As a *domain* (set of positions), the important structure of $\mathbb{R}^n$ is its abelian Lie group structure, where we can move around using translations. As a tangent space (set of velocities), the important structure is the vector space structure (Lie algebra structure) of $\mathbb{R}^n$. These spaces play very different roles in theory and in computations and should not be confused!

The theory of abelian groups is *much* simpler than general groups and Fourier analysis is a ubiquitous tool which is tightly associated with the group structure of these spaces. The relationship between the continuous and the discrete is a fundamental aspect of computational algorithms, such as the re-

lationship between continuous groups such as $\mathbb{R}^n$ and discrete subgroups (lattices). The Fourier transform can be computed fast on finite abelian groups, but not on T nor $\mathbb{R}$. Without a good mathematical theory and supporting software, it is, however not trivial to relate the continuous and the discrete Fourier transforms. This is in particular the case for general sampling lattices in $\mathbb{R}^n$.

The general theory of subgroups, quotients and exact sequences turns out to be a very useful framework for developing and analysing computational algorithms. This is the topic of the next chapter.

3 Abelian groups, Fourier analysis, lattices and sampling

In this chapter we will introduce abelian groups as domains for computations. We will in particular present a general theory of discretisation lattices, sampling theory and the relationship between discrete and continuous Fourier analysis, and discuss a variety of computational algorithms. Circulant matrices and their multidimensional analogues is also a central theme.

3.1 Introduction to abelian groups

Definition and basic properties.

Using the additive notation with + and 0, we define abelian groups:

Definition 19 (Abelian group). *An abelian group is a set G with a binary operation $+\colon G \times G \to G$ called the* sum*, such that*

1. *The sum is associative, $x + (y + z) = (x + y) + z$ for all $x, y, z \in G$.*
2. *The sum is commutative, $x + y = y + x$.*
3. *There exits an identity element $\mathbf{0} \in G$ such that $x + \mathbf{0} = x$ for all $x \in G$.*
4. *Every element $x \in G$ has an inverse $-x \in G$ such that $x + (-x) = \mathbf{0}$.*

For abelian groups the direct product is the same as a *direct sum*[4], we write this as $H \oplus K \equiv H \times K$. This means, as before, that $H \oplus K = \{(k, h)\}$ with $(k, h) + (k', h') = (k + k', h + h')$.

Abelian groups are much simpler than general groups, since there is no difference between 'left' and 'right', they enjoy the following properties:

- Any subgroup $H < G$ is a normal subgroup.
- A short exact sequence is right split if and only if it is left split, thus a split short exact sequence is always of the form

$$\mathbf{0} \longrightarrow H \underset{\psi_H}{\overset{\varphi_G}{\rightleftarrows}} G \underset{\psi_G}{\overset{\varphi_K}{\rightleftarrows}} K \longrightarrow \mathbf{0} \tag{8}$$

  corresponding to the direct sum decomposition $G = H \oplus K$.

[4] The category theoretical definition of *products* and *coproducts* are dual of each other, but for abelian groups they coincide.

Topology.

In order to develop a mathematical theory of Fourier analysis, it is necessary to have some topology (notion of continuity) on the groups. The standard foundation of Fourier analysis on groups are so called *locally compact* groups.

Definition 20 (Locally compact group). *A group is called locally compact if it has a topology such that every point has a compact neighbourhood, and such that the product and inverse in the group are continuous operations.*

Definition 21 (LCA - Locally Compact Abelian). *LCA denotes the locally compact abelian groups.*

Between topological groups, homomorphisms are always assumed continuous, and when we talk about a subgroup $H < G$, we will always assume that H is a *closed subset*. For example, the rationals $(\mathbb{Q}, +)$ is algebraically a subgroup of $(\mathbb{R}, +)$, but it is not a topologically closed subset.

The elementary groups.

In these lectures we will mainly focus the *elementary* abelian groups, those that can be obtained from $\mathbb{R}$ and $\mathbb{Z}$ by taking direct sums, (closed) subgroups and quotients. The topology for these is what we are used to, e.g. $\mathbb{Z}$ has the discrete topology where every subset is an open set (and also closed!), and $\mathbb{R}$ has the familiar topology of the real line based on open intervals defining open subsets. The elementary LCAs are isomorphic to one of the following:

Definition 22. *The* elementary LCAs *are:*

- *The* reals $\mathbb{R}$ *under addition, with the standard definition of open sets.*
- *The* integers $\mathbb{Z}$ *under addition (with the discrete topology). This is also known as the infinite cyclic group.*
- *The* 1*-dimensional* torus*, or* circle $T = \mathbb{R}/\mathbb{Z}$ *defined as* $[0, 1) \subset \mathbb{R}$ *under addition modulo* 1*, with the circle topology.*
- *The* cyclic group *of order* k*,* $\mathbb{Z}_k = \mathbb{Z}/k\mathbb{Z}$*, which consists of the integers* $0, 1, \ldots, k-1$ *under addition modulo* k *(with the discrete topology).*
- Direct sums *of the above spaces,* $G \oplus H$*, in particular* $\mathbb{R}^n$ *(real* n*-space),* T^n *(the* n*-torus) and all finitely generated abelian groups.*

A set of *generators* for a group is a subset such that any element in the group can be written as a finite sum (or difference) of the generators. The *finitely generated abelian groups* are those having a finite set of generators. These are easy to describe, since they are always isomorphic to a direct sum of $\mathbb{Z}$ and $\mathbb{Z}_{n_i}$. We take this as a definition, but keep in mind that they may appear in disguise, as e.g. the multiplicative group C_n isomorphic to $\mathbb{Z}_n$.

Definition 23 (Finitely Generated Abelian group, FGA). *An FGA is a group isomorphic to*

$$\mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_k} \oplus \mathbb{Z}^d.$$

We represent this as the space of integer column vectors of length $k+d$ under addition mod n_i *in first k components and integer addition in the last d components. The canonical generators are $(1,0,\ldots,0)^T$, $(0,1,0,\ldots,0)^T$, $\ldots$, $(0,\ldots,0,1)^T$. Note that $\mathbb{Z}_1 = \mathbf{0}$ and $\mathbf{0} \oplus G \simeq G$, hence we can remove the terms $\mathbb{Z}_{n_i}$ whenever $n_i = 1$.*

We will in the sequel use the following compact notation for FGAs

$$\mathbb{Z}_{\mathbf{k}} \oplus \mathbb{Z}^d := \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_k} \oplus \mathbb{Z}^d,$$

where $\mathbf{k} = (n_1, n_2, \ldots, n_k)$ is a multi-index of length k. The number $k+d$ (number of generators) is called the *rank* of the FGA, but the rank is not an invariant under isomorphisms.

FGAs are similar to vector spaces, but where the 'scalars' are the integers $\mathbb{Z}$ instead of the usual fields $\mathbb{R}$ or $\mathbb{C}$. The generators of the FGA are similar to basis vectors for a vector space. Homomorphisms between FGAs can always be written as integer matrices representing how the homomorphism acts on the canonical generators. Note that we will always assume that the target space knows the periods of its components, e.g. the homomorphism $\cdot 2 \colon \mathbb{Z} \to \mathbb{Z}_3$ (multiplication by 2), sends $0 \mapsto 0$, $1 \mapsto 2$, $2 \mapsto 4 \equiv 1 (\mathrm{mod}\ 3)$, etc. We will not write the reduction modulo 3 explicitly.

Note that not every integer matrix (of appropriate dimensions) represents a homomorphism. The obstruction is that every integer vector congruent to $\mathbf{0}$ in the source group must be mapped to an integer vector congruent to $\mathbf{0}$ in the target group. For example $\cdot 2$ does not define a homomorphism from $\mathbb{Z}_2$ to $\mathbb{Z}_3$ since $2 \cdot 2 \neq 0 (\mathrm{mod}\ 3)$, however $\cdot 4 \colon \mathbb{Z}_3 \to \mathbb{Z}_{12}$ is a homomorphism since $4 \cdot 3k = 0 (\mathrm{mod}\ 12)$ for every $k \in \mathbb{Z}$.

The notion of the dimension is less clear in the theory of FGAs compared to standard linear algebra (where the size of a basis is invariant under basis change). In Example 12, we claimed that $\mathbb{Z}_{12} \simeq \mathbb{Z}_3 \oplus \mathbb{Z}_4$, so isomorphic groups can have different numbers of independent generators. In this case $(4,3) \colon \mathbb{Z}_3 \oplus \mathbb{Z}_4 \to \mathbb{Z}_{12}$ and $(1,-1)^T \colon \mathbb{Z}_{12} \to \mathbb{Z}_3 \oplus \mathbb{Z}_4$ define isomorphisms between the two groups. (Check this!)

In general we have that $\mathbb{Z}_p \oplus \mathbb{Z}_q \simeq \mathbb{Z}_{pq}$ if and only if p and q are relative prime numbers (i.e. if their greatest common divisor, gcd, is 1). To compute the isomorphism between these, we employ the *extended Euclidean algorithm* (matlab function 'gcd'), which given two positive integers p and q produces two integers a and b such that

$$ap + bq = \gcd(p,q).$$

If $\gcd(p,q)=1$, we have that $(q,p)\colon \mathbb{Z}_p \oplus \mathbb{Z}_q \to \mathbb{Z}_{pq}$ and $(b,a)^T\colon \mathbb{Z}_{pq} \to \mathbb{Z}_p \oplus \mathbb{Z}_q$ are isomorphisms. We can also illustrate the isomorphism by the split short exact sequence

$$\mathbf{0} \longrightarrow \mathbb{Z}_p \underset{\cdot b}{\overset{\cdot q}{\rightleftarrows}} \mathbb{Z}_{pq} \underset{\cdot p}{\overset{\cdot a}{\rightleftarrows}} \mathbb{Z}_q \longrightarrow \mathbf{0}$$

Check yourself that this is split exact!

We have two standard ways of representing FGAs uniquely (up to isomorphisms), the first of these has the largest possible number of generators and the latter the smallest possible:

Theorem 2 (Classification of FGA). *If G is an FGA, and $G \neq \mathbf{0}$, then G is isomorphic to a group of the form called the* primary factor decomposition

$$\mathbb{Z}_{p_1^{n_1}} \oplus \mathbb{Z}_{p_2^{n_2}} \oplus \cdots \oplus \mathbb{Z}_{p_\ell^{n_\ell}} \oplus \mathbb{Z}^n \tag{9}$$

where p_i are primes, $p_1 \leqslant p_2 \leqslant \cdots \leqslant p_\ell$, $n_i \in \mathbb{N}$ and $n_i \leqslant n_{i+1}$ whenever $p_i = p_{i+1}$. Furthermore G is also isomorphic to a group of the form called the invariant factor decomposition

$$\mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_k} \oplus \mathbb{Z}^n \tag{10}$$

where $n_i > 1$ and $n_i | n_{i+1}$, $1 \leqslant i < k$ (n_i divides n_{i+1}). In both forms the representation is unique, i.e. two FGA are isomorphic iff they can be transformed into the same canonical form.

3.2 Computing with FGAs

For applications in lattice sampling algorithms and computational Fourier analysis, it is important to do computations on FGAs and homomorphisms between these. It is our philosophy that software in computational mathematics should closely follow mathematical definitions. *Object oriented programming* is founded on the idea that programming is *'what' + 'how'*, i.e. the software is constructed from classes where there is a distinction between the (public) signature (what) of the class and the (private) implementation (how). The signature consists of the functions operating on the structure and the implementation consists of data structures and algorithms. To help finding useful abstractions for defining the 'what' part of program design, we have found mathematical category theory very useful. Categories consists of objects and arrows between the objects, just as we have seen in the discussion of exact sequences above. In category theory one does not explicitly describe what is 'inside' an object, the only mathematical properties one is interested in are those that can be described in terms of arrows into and out of the object. This philosophy fits very well with object oriented programming design, and a categorical definition of a mathematical object is a good starting point for object oriented software construction.

For example, a split exact sequence

$$\mathbf{0} \longrightarrow G_1 \underset{\text{proj}_1}{\overset{\text{inj}_1}{\rightleftarrows}} G_1 \oplus G_2 \underset{\text{inj}_2}{\overset{\text{proj}_2}{\rightleftarrows}} G_2 \longrightarrow \mathbf{0}$$

could be taken as the *definition* of the direct sum $G_1 \oplus G_2$. The 'object' $G_1 \oplus G_2$ is defined by the existence of the four morphisms inj_1, inj_2, proj_1 and proj_2 such that $\text{proj}_1 \circ \text{inj}_1 = \text{Id}_{G_1}$ (the identity homomorphism), $\text{proj}_2 \circ \text{inj}_2 = \text{Id}_{G_2}$ and exactness of the diagram in both directions. The usual *implementation* ('how') of the direct product $G_1 \oplus G_2$ is as pairs (g_1, g_2), where the arrows are $\text{inj}_1(g_1) = (g_1, 0)$, $\text{proj}_1((g_1, g_2)) = g_1$, and similarly for G_2. Could there possibly be any other implementation of the direct sum? Yes, for high dimensional n there are different ways of representing $\mathbb{R}^n$. The most common is just as vectors of length n, but if many of the vectors are sparse, on could instead use a sparse representation where only the non-zero components are stored. It is important to realise that these two implementations are isomorphic realisations of the 'specification' given by the split exact sequence.

Abelian categories.

Category theory gives an important hint on what are the most fundamental properties we should know about when designing software. The collection of all FGAs and homomorphisms between these form an *abelian category*, where each object is an FGA and each arrow is a homomorphism between two FGAs. Abelian categories have the following properties:

- There is a *zero object* $\mathbf{0}$. For any object G there is a unique $\mathbf{0}$ arrow $\mathbf{0} \to G$ and a unique arrow $G \to \mathbf{0}$.
- We can form the *product* and *coproduct* of any two objects. In the setting of FGAs, these two are the same, represented by the direct sum of two abelian groups $G_1 \oplus G_2$ and the arrows in and out of the sum.
- The set $\hom(H, G)$ of all morphisms from H to G is an object in the category, i.e. it contains the $\mathbf{0}$-arrow, any two arrows can be added, and furthermore the composition $\circ\colon\ \hom(H, G) \times \hom(G, K) \to \hom(H, K)$ is bilinear. In our case $\hom(H, G)$ is an FGA.
- Every homomorphism $\varphi \in \hom(H, G)$ has a kernel $\ker(\varphi) \in \hom(K, H)$ and a cokernel $\text{coker}(\varphi) \in \hom(G, C)$, such that the following is an exact sequence

$$\mathbf{0} \longrightarrow K \xrightarrow{\ker(\varphi)} H \xrightarrow{\ \varphi\ } G \xrightarrow{\text{coker}(\varphi)} C \longrightarrow \mathbf{0}\ .$$

- Every monomorphism is a kernel of some homomorphism and every epimorphism is the cokernel of some homomorphism.
- Every homomorphism $\varphi \in \hom(G_1, G_2)$ factors in the composition of an epimorphism followed by a monomorphism The epimorphism is called the coimage, and the monomorphism is called the image,

$$\varphi = \text{im}(\varphi) \circ \text{coim}(\varphi).$$

All these properties should be implemented in a software package for computing with FGAs. Furthermore, there are a set of operations which are derived from the addition and composition of homomorphisms. We introduce some short hand notation for these. First three operations which are related to direct sums. For homomorphisms represented as matrices, these operations correspond to creating new matrices from matrix blocks. The matrix interpretation is based on FGAs being column vectors and the sum $G_1 \oplus G_2$ interpreted as putting the two column vectors on top of each other. For Matlab users semicolon notation is familiar. If $x \in G_1$ and $y \in G_2$ are column vectors, then $(x; y) \in G_1 \oplus G_2$ means that we put x and y together in a long column with x on top of y.

Definition 24 (Block compositions of homomorphisms).

- *For* $\phi_1 \in \hom(G_1, H_1)$, $\phi_2 \in \hom(G_2, H_2)$ *we define*

$$\phi_1 \oplus \phi_2 \in \hom(G_1 \oplus G_2, H_1 \oplus H_2)$$

as

$$(\phi_1 \oplus \phi_2)(x; y) := (\phi_1(x); \phi_2(y)).$$

This corresponds to a diagonal 2×2 *block matrix with* ϕ_1 *in upper left and* ϕ_2 *in lower right block, or the diagram*

$$\begin{array}{ccccc} G_1 & \rightleftarrows & G_1 \oplus G_2 & \rightleftarrows & G_2 \\ \downarrow{\scriptstyle \phi_1} & & \downarrow{\scriptstyle \phi_1 \oplus \phi_2} & & \downarrow{\scriptstyle \phi_2} \\ H_1 & \rightleftarrows & H_1 \oplus H_2 & \rightleftarrows & H_2. \end{array}$$

- *For* $\phi_1 \in \hom(G_1, H)$, $\phi_2 \in \hom(G_2, H)$ *we define*

$$\phi_1 | \phi_2 \in \hom(G_1 \oplus G_2, H)$$

as

$$(\phi_1 | \phi_2)(x; y) := \phi_1(x) + \phi_2(y).$$

This corresponds to putting two matrices horizontally in a 1×2 *block matrix, or the diagram*

$$\begin{array}{ccccc} G_1 & \rightleftarrows & G_1 \oplus G_2 & \rightleftarrows & G_2 \\ & {\scriptstyle \phi_1} \searrow & \downarrow{\scriptstyle \phi_1 | \phi_2} & \swarrow {\scriptstyle \phi_2} & \\ & & H & & \end{array}.$$

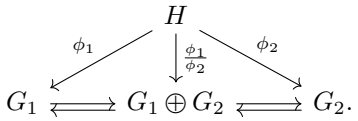
- *For* $\phi_1 \in \hom(H, G_1)$, $\phi_2 \in \hom(H, G_2)$ *we define*

$$\frac{\phi_1}{\phi_2} \in \hom(H, G_1 \oplus G_2)$$

as

$$\begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix}(x) := (\phi_1(x);\phi_2(x))\,.$$

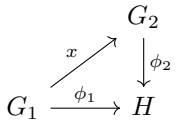
This corresponds to putting two matrices vertically in a 2×1 *block matrix, or the diagram*



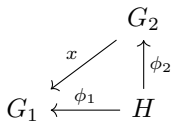
The next two operations are factorisations of a homomorphism through another, which is similar to solving linear equations.

Definition 25 (Factorisation of a homomorphism through another). *We define two ways of solving for an unknown homomorphism* x*. The solution may not exist, or may not be unique (conditions apply).*

- *For* $\phi_1 \in \hom(G_1, H)$ *and* $\phi_2 \in \hom(G_2, H)$ *we denote* $x = \phi_2\backslash\phi_1$ *a homomorphism* $x \in \hom(G_1, G_2)$ *such that* $\phi_2 \circ x = \phi_1$*.*



- *For* $\phi_1 \in \hom(H, G_1)$ *and* $\phi_2 \in \hom(H, G_2)$ *we denote* $x = \phi_1/\phi_2$ *a homomorphism* $x \in \hom(G_2, G_1)$ *such that* $x \circ \phi_2 = \phi_1$*.*



Free FGAs and Smith's normal form.

The free finitely generated abelian groups are those which have no relations between the generators, i.e. the abelian groups $\mathbb{Z}^n$ for $n \in \mathbb{N}$. These are particularly simple, since the set of integer matrices are in 1–1 correspondence with homomorphisms

$$\hom(\mathbb{Z}^n, \mathbb{Z}^m) \approx \mathbb{Z}^{m\times n}.$$

The set $\hom(\mathbb{Z}^n, \mathbb{Z}^m)$ is an FGA with addition defined as matrix addition and 0 being the zero matrix. The composition of homomorphisms is given by matrix products. From Cramers rule we realise that a matrix $A \in \mathbb{Z}^{n\times n}$ has an inverse in $\mathbb{Z}^{n\times n}$ if and only if $\det(A) = \pm 1$.

Definition 26 (Unimodular matrix). *A matrix* $A \in \mathbb{Z}^{n\times n}$ *with* $\det(A) = \pm 1$ *is called* unimodular *and represents an isomorphism in* $\mathrm{iso}(\mathbb{Z}^n, \mathbb{Z}^n)$*. The unimodular* $n \times n$ *integer matrices are denoted* $GL(n, \mathbb{Z})$*.*

Many fundamental properties of homomorphisms in $\hom(\mathbb{Z}^n, \mathbb{Z}^m)$ are computed from the *Smith normal form* of A, a decomposition quite similar to the SVD. An algorithm for computing this is given in Wikipedia [31].

Theorem 3. *An integer matrix $A \in \mathbb{Z}^{m\times n}$ can be decomposed in a product*

$$A = U\Sigma V$$

where $U \in GL(m, \mathbb{Z})$ and $V \in GL(n, \mathbb{Z})$ are unimodular and $\Sigma \in \mathbb{Z}^{m\times n}$ is diagonal with non-negative diagonal elements. The diagonal elements $n_i = \Sigma_{ii}$ satisfy $n_i | n_{i+1}$ $\forall$ $1 \leqslant i < k$ and $n_i = 0$ $\forall$ $k < i \leqslant \min(m, n)$.

Theorem 4. *Let $A \in \hom(\mathbb{Z}^n, \mathbb{Z}^m)$ with Smith decomposition $A = U\Sigma V$, with matrices partitioned as*

$$U = \begin{pmatrix} U_1 & U_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix},$$

where $U_1 \in \mathbb{Z}^{m\times k}$, $U_2 \in \mathbb{Z}^{m\times(m-k)}$, $\Sigma_{11} \in \mathbb{Z}^{k\times k}$, $\Sigma_{12} \in \mathbb{Z}^{k\times(n-k)}$, $\Sigma_{21} \in \mathbb{Z}^{(n-k)\times k}$, $\Sigma_{22} \in \mathbb{Z}^{(n-k)\times(n-k)}$, $V_1 \in \mathbb{Z}^{k\times n}$ and $V_2 \in \mathbb{Z}^{(n-k)\times n}$. The matrix Σ_{11} has diagonal $\mathbf{k} = (n_1, n_2, \ldots, n_k)$ such that $n_i | n_{i+1}$ and Σ_{12}, Σ_{21} and Σ_{22} are all zero. Partition U^{-1} and V^{-1} as

$$U^{-1} = \begin{pmatrix} U_1^{-1} \\ U_2^{-1} \end{pmatrix}, V^{-1} = \begin{pmatrix} V_1^{-1} & V_2^{-1} \end{pmatrix},$$

where $U_1^{-1} \in \mathbb{Z}^{k\times m}$, $U_2^{-1} \in \mathbb{Z}^{(m-k)\times m}$, $V_1^{-1} \in \mathbb{Z}^{n\times k}$ and $V_2^{-1} \in \mathbb{Z}^{n\times(n-k)}$. Then

$$\begin{aligned} \ker(A) &= V_2^{-1} \in \operatorname{mono}(\mathbb{Z}^{n-k}, \mathbb{Z}^n) \\ \operatorname{coker}(A) &= U^{-1} = \frac{U_1^{-1}}{U_2^{-1}} \in \operatorname{epi}(\mathbb{Z}^m, \mathbb{Z}_{\mathbf{k}} \oplus \mathbb{Z}^{m-k}) \\ \operatorname{coim}(A) &= V_1 \in \operatorname{epi}(\mathbb{Z}^n, \mathbb{Z}^k) \\ \operatorname{im}(A) &= U_1\Sigma_{11} \in \operatorname{mono}(\mathbb{Z}^k, \mathbb{Z}^m). \end{aligned}$$

Proof. Check that that the diagrams

$$\mathbf{0} \longrightarrow \mathbb{Z}^{n-k} \xrightarrow{V_2^{-1}} \mathbb{Z}^n \xrightarrow{A} \mathbb{Z}^m \xrightarrow{U^{-1}} \mathbb{Z}_{\mathbf{k}} \oplus \mathbb{Z}^{m-k} \longrightarrow \mathbf{0}$$

and

$$\begin{array}{ccccc} & & \mathbf{0} & & \\ & & \downarrow & & \\ \mathbb{Z}^n & \xrightarrow{V_1} & \mathbb{Z}^k & \longrightarrow & \mathbf{0} \\ & \searrow^{A} & \downarrow{\scriptstyle U_1\Sigma_{11}} & & \\ & & \mathbb{Z}^m & & \end{array}$$

are commutative with exact rows and columns. □

Example 13. Given $A \in \hom(\mathbb{Z}^3, \mathbb{Z}^4)$ with Smith normal form

$$A = \begin{pmatrix} -20 & 8 & 16 \\ -6 & 0 & 6 \\ 0 & -12 & 6 \\ 4 & -16 & 4 \end{pmatrix} = \begin{pmatrix} 8&6&3&0 \\ 3&2&1&0 \\ 3&1&0&0 \\ 2&0&0&1 \end{pmatrix} \begin{pmatrix} 2&0&0 \\ 0&6&0 \\ 0&0&0 \\ 0&0&0 \end{pmatrix} \begin{pmatrix} 1 & -4 & 1 \\ -1 & 2 & 0 \\ 0 & 1 & -1 \end{pmatrix} = U\Sigma V,$$

where

$$U^{-1} = \begin{pmatrix} -1 & 3 & 0&0 \\ 3 & -9 & 1&0 \\ -3 & 10 & -2&0 \\ 2 & -6 & 0&1 \end{pmatrix}, V^{-1} = \begin{pmatrix} -2 & -3 & -2 \\ -1 & -1 & -1 \\ -1 & -1 & -2 \end{pmatrix}.$$

From this we see that A generates a rank-2 subgroup of $\mathbb{Z}^4$, spanned by $2 \cdot (8,3,3,2)^T$ and $6 \cdot (6,2,1,0)^T$. The quotient of $\mathbb{Z}^4$ with this subgroup is $\mathbb{Z}_2 \oplus \mathbb{Z}_6 \oplus \mathbb{Z}$ and the matrix U^{-1} projects onto this quotient. The kernel of A is the rank-1 subgroup of $\mathbb{Z}^3$ spanned by $(-2,-1,-2)^T$

Example 14. Let $H < \mathbb{Z}^2$ be the subgroup spanned by $(-1,3)^T$ and $(2,2)$. Compute $\mathbb{Z}^2/H$ and the projection. We compute the Smith factorisation of the generators

$$A = \begin{pmatrix} -1 & 2 \\ 3 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 11 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 8 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} = U\Sigma V,$$

where $U = U^{-1}$. From this we see that $\mathbb{Z}^2/H = \mathbb{Z}_1 \oplus \mathbb{Z}_8 \simeq \mathbb{Z}_8$, and $\operatorname{coker}(A)$ is just the last row of U^{-1},

$$\operatorname{coker}(A) = \begin{pmatrix} 11 & 1 \end{pmatrix} \in \operatorname{epi}(\mathbb{Z}^2, \mathbb{Z}_8).$$

General homomorphisms.

We have seen that homomorphisms between the free finitely generated abelian groups are integer matrices. How can we represent and compute homomorphisms between general FGAs?

Lemma 1. *Any FGA G of rank m is given as the image of a projection of the free group $\mathbb{Z}^m$ onto G, $\pi_G \in \operatorname{epi}(\mathbb{Z}^m, G)$.*

Proof. For $G = \mathbb{Z}_{\mathbf{k}} \oplus \mathbb{Z}^d$ let $m = k + d$, $k = |\mathbf{k}|$ and set $\pi_G(z) = z \bmod \mathbf{k}$ in the first k components and $\pi_G(z) = z$ in the last d components. Since any FGA is isomorphic to such a G we can produce a projection on any FGA by composing π_G with an isomorphism. □

We call the above defined π_G *the canonical projection*. In many situations it is useful to represent G by an other projection, e.g. we can more generally choose some $A \in \mathbb{Z}^{m \times n}$ and let $\pi_G = \operatorname{coker}(A)$. In this case, if $A = U\Sigma V$ we have $\pi_G(z) = U^{-1}z \bmod \mathbf{k}$, where $\mathbf{k}$ is the diagonal of Σ.

Lemma 2. *Let G and H be arbitrary FGAs and let $\pi_G \in \mathrm{epi}(\mathbb{Z}^n, G)$ and $\pi_H \in \mathrm{epi}(\mathbb{Z}^m, H)$ be projections onto these. A matrix $A \in \mathbb{Z}^{m\times n} \approx \hom(\mathbb{Z}^n, \mathbb{Z}^m)$ represents a homomorphism $\phi = (\pi_H \circ A)/\pi_G \in \hom(G, H)$ if and only if $\ker(\pi_G) < \ker(\pi_H \circ A)$. Any $\phi \in \hom(G, H)$ can be written this way. The matrix A is generally not unique for a given ϕ.*

$$\begin{array}{ccc} \mathbb{Z}^n & \xrightarrow{A} & \mathbb{Z}^m \\ \downarrow{\scriptstyle \pi_G} & & \downarrow{\scriptstyle \pi_H} \\ G & \xrightarrow{\phi} & H \end{array}$$

Proof. First we start with a given $\phi \in \hom(G, H)$. Since $\phi \circ \pi_G = \pi_H \circ A$, we have $A = \pi_H \backslash (\phi \circ \pi_G)$. Since π_H is onto H, this equation can always be solved, but the solution is not unique since we can add something in the $\ker(\pi_H)$ to A without affecting the solution. From the diagram it is easy to check that $\ker(\pi_G) < \ker(\pi_H \circ A)$.

Now, assume we are given an A such that $\ker(\pi_G) < \ker(\pi_H \circ A)$. This means that for any $y \in \ker(\pi_G)$ we have $(\pi_H \circ A)(x + y) = (\pi_H \circ A)(x)$ for all x. Hence $\pi_H \circ A$ takes constant values on each of the cosets of $\ker(\pi_G) < \mathbb{Z}^n$, and it defines a function on $G \simeq \mathbb{Z}^n / \ker(\pi_G)$. In other words, this is the necessary condition for solving the equation $\phi = (\pi_h \circ A)/\pi_G \in \hom(G, H)$. ▫

Exercise 5. Check that $A = 4$ defines a homomorphism $\phi \in \hom(\mathbb{Z}_2, \mathbb{Z}_8)$. Find an other A representing the same homomorphism.

Definition 27 (Matrix representation of a general homomorphism). *The notation $A \in \hom(G, H)$ for some matrix $A \in \mathbb{Z}^{m\times n}$ means that A represents the homomorphism $(\pi_H \circ A)/\pi_G \in \hom(G, H)$. Unless otherwise specified, the projections π_G and π_H are the canonical projections (i.e. $\mathbb{Z}^m$ and $\mathbb{Z}^n$ mod the periods of G and H).*

Even if we can represent any homomorphism in terms of an integer matrix, it does not mean that all computations are trivial. Some care must be taken! We illustrate by an example.

Example 15. Let $G = \mathbb{Z}_8 \oplus \mathbb{Z}_8$ and $H = \langle (1; 5) \rangle < G$, meaning that H is the subgroup of G generated by the element $(1; 5) \in G$. We want to compute G/H. Let $\pi \in \mathrm{epi}(\mathbb{Z}^2, G)$ be the natural projection, $\pi(z) = z \bmod (8; 8)$ and $A = (1; 5) \in \hom(\mathbb{Z}, \mathbb{Z}^2)$. We have $H = \mathrm{im}(\phi)$ where $\phi = \pi \circ A$, and our task is to compute $\mathrm{coker}(\phi) \in \mathrm{epi}(G, G/H)$. Note that even if $\phi = \pi \circ A$, it is not so that the pre-image of H in $\mathbb{Z}^2$ is the image of A. The image of A is just the line of points $\langle (1; 5) \rangle < \mathbb{Z}^2$, while the pre-image of H contains all the points $j \cdot (1; 5) + y$ for all $y \in \ker(\pi)$. To find the pre-image of H we compute

$$\ker(\pi) = \begin{pmatrix} 8 & 0 \\ 0 & 8 \end{pmatrix}$$

and

$$\widetilde{A} = A|\ker(\pi) = \begin{pmatrix} 1 & 8 & 0 \\ 5 & 0 & 8 \end{pmatrix}.$$

Thus, the image of $\widetilde{A}$ is exactly the pre-image of H in $\mathbb{Z}^2$. Smith decomposition yields

$$\widetilde{A} = \begin{pmatrix} 1 & 0 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \end{pmatrix} \begin{pmatrix} 1 & 8 & 0 \\ -1 & -3 & -1 \\ 0 & 1 & 0 \end{pmatrix},$$

from which we see that $\mathbb{Z}^2/\operatorname{im}(\widetilde{A}) = \mathbb{Z}_1 \oplus \mathbb{Z}_8$ and

$$\operatorname{coker}(\widetilde{A}) = U^{-1} = \begin{pmatrix} 1 & 0 \\ -3 & -1 \end{pmatrix} \in \operatorname{epi}(\mathbb{Z}^2, \mathbb{Z}_1 \oplus \mathbb{Z}_8).$$

Of course $\mathbb{Z}_1 \oplus \mathbb{Z}_8 \simeq \mathbb{Z}_8$ and we have $\operatorname{coker}(\widetilde{A}) = (-3\ -1) \in \operatorname{epi}(\mathbb{Z}^2, \mathbb{Z}_8)$. Since $\ker(\pi) < \operatorname{im}(\widetilde{A})$ we have that $\operatorname{coker}(\widetilde{A})$ factors through π and we can compute

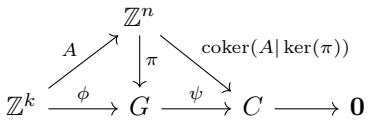
$$\operatorname{coker}(\phi) = \operatorname{coker}(\widetilde{A})/\pi = (-3\ -1) \in \operatorname{epi}(G, \mathbb{Z}_8).$$

We conclude that $G/H = \mathbb{Z}_8$ with this projection.

Example 16 (Computing the cokernel of a general homomorphism). The above example generalises to the general problem of computing G/H, where $H < G$ is a subgroup generated by k elements of G. Let $\pi \in \operatorname{epi}(\mathbb{Z}^n, G)$ be the natural projection and $A \in \hom(\mathbb{Z}^k, \mathbb{Z}^n)$ such that H is the image of $\phi = \pi \circ A$, i.e. the columns of A represent the generators of H. We claim

$$\operatorname{coker}(\phi) = \psi := \pi \backslash \operatorname{coker}(A|\ker(\pi)). \tag{11}$$

To prove this, we must show that the bottom row of



is exact. First we check that $\psi \circ \phi = 0$ by following the two top diagonal arrows (which by definition compose to zero). Next we see that ψ is an epimorphism (onto C), since $\operatorname{coker}(A|\ker(\pi))$ by definition is onto. Last we pick an $z \in G$ such that $\psi(z) = 0$. This must mean that $z = \pi(y)$ for some $y \in \operatorname{im}(A|\ker(\pi))$, hence $y = Ax + w$ for some $x \in \mathbb{Z}^k$ and $w \in \ker(\pi)$, from which it follows that $\phi(x) = z$. This proves that $\operatorname{im}(\phi) = \ker(\psi)$ and the bottom line is exact. We conclude that $C = G/\operatorname{im}(\phi) = G/H$ and that $\operatorname{coker}(\phi) = \psi \in \operatorname{epi}(G, G/H)$ is the projection.

Hermite's normal form

Smith's normal form is perfect for computing the structure of quotients. To compute images (and coimages) of maps into general FGAs, another normal form is sometimes more useful. Whereas Smith's normal form is the integer matrix version of SVD, the *Hermite normal form* is the integer version of LU factorisation. The basic idea is to factorise $A \in \mathbb{Z}^{m\times n}$ as $AV = H$, where $H \in \mathbb{Z}^{m\times k}$ is in *lower echelon form* and $V \in \mathrm{GL}(n,\mathbb{Z})$. If the columns of A are generators of some subgroup then the columns of H constitute a set of generators for the same subgroup. The details of Hermite's normal form differ among different authors. There are row and column versions and some other details that can be done differently. Since we interpret group elements as column vectors we prefer a column version.

We say that an element $h_{i,j}$ in H is a *pivot* if $h_{i,j} \neq 0$ and everything above and to the right of $h_{i,j}$ is zero, i.e. $h_{k,\ell} = 0$ whenever $k \leqslant i$ and $\ell \geqslant j$ and $(k,\ell) \neq (i,j)$. The matrix H is in *lower echelon form* if every column i has a pivot $h_{p(i),i}$ and furthermore $p(i) < p(i+1)$ for every $i \in \{1,\dots,k-1\}$.

Definition 28. *A matrix $H \in \mathbb{Z}^{m\times k}$ is in Hermite's normal form if*

1. *H is in lower echelon form.*
2. *Each pivot $h_{p(i),i} > 0$.*
3. *All elements to the left of a pivot are nonnegative and smaller than the pivot; for every every $j \in \{1,\dots,i-1\}$ we have $0 \leqslant h_{p(i),j} < h_{p(i),i}$.*

Lemma 3. *For every non-zero $A \in \mathbb{Z}^{m\times n}$ there exists a $V \in GL(n,\mathbb{Z})$ partitioned as $V = (V_1|V_2)$, where $V_1 \in \mathbb{Z}^{n\times k}$, $V_2 \in \mathbb{Z}^{n\times(n-k)}$ such that*

$$H = AV_1 \in \mathbb{Z}^{m\times k}$$

is in Hermite's normal form.

Proof. We sketch an algorithm for computing this factorisation by applying elementary unimodular matrices acting on A from the right. A matrix of the form

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

in the upper left, and the identity in the lower right swaps the two first columns of A. More generally, any permutation matrix having exactly one 1 in each column and in each row, and zeros elsewhere permutes the columns and is always unimodular. A less trivial unimodular matrix is obtained as follows. Consider two positive integers r and s. From the euclidean algorithm we compute integers a and b such that

$$ar + bs = g = \gcd(r,s).$$

Note that

$$\begin{pmatrix} r & s \end{pmatrix} \begin{pmatrix} a & -\frac{s}{g} \\ b & \frac{r}{g} \end{pmatrix} = \begin{pmatrix} g & 0 \end{pmatrix},$$

where the matrix is unimodular. Thus, if $A_{11} = r$ and $A_{12} = s$ we can multiply A by such a matrix from the right to obtain $(g, 0)$ in the first two positions of row 1. We can continue to eliminate all entries to the right of A_{11}, possibly swapping columns if some entries in the first row are 0. We proceed the algorithm by searching for a new pivot in position 2 to m in row 2. If there are no pivots here we go to the next row etc. Whenever we have eliminated everything to the right of a pivot, we subtract multiples of the pivot column from all columns to the left of the pivot to fulfil criterion 3. in the definition of the Hermite normal form. □

For $A \in \hom(\mathbb{Z}^n, \mathbb{Z}^m)$ it is clear that the columns of H are independent and span the image of A, hence we can find both the kernel, image and coimage of A from the Hermite normal form decomposition:

$$\begin{aligned} \ker(A) &= V_2 \in \mathrm{mono}(\mathbb{Z}^{n-k}, \mathbb{Z}^n) \\ \mathrm{im}(A) &= H \in \mathrm{mono}(\mathbb{Z}^k, \mathbb{Z}^m) \\ \mathrm{coim}(A) &= V_1^{-1} \in \mathrm{epi}(\mathbb{Z}^n, \mathbb{Z}^k), \end{aligned}$$

where V_1^{-1} denotes the upper $k \times n$ block of V^{-1}.

Example 17. The matrix A of Example 13 has Hermite factorisation

$$\begin{pmatrix} -20 & 8 & 16 \\ -6 & 0 & 6 \\ 0 & -12 & 6 \\ 4 & -16 & 4 \end{pmatrix} \begin{pmatrix} -1 & 0 & 2 \\ 0 & -2 & 1 \\ -1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 6 & 0 \\ -6 & 30 & 0 \\ -8 & 36 & 0 \end{pmatrix}$$

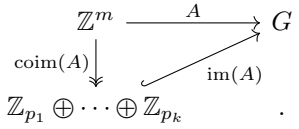
from which we find

$$A = \mathrm{im}(A) \circ \mathrm{coim}(A) = H V_1^{-1} = \begin{pmatrix} 4 & 0 \\ 0 & 6 \\ -6 & 30 \\ -8 & 36 \end{pmatrix} \begin{pmatrix} -5 & 2 & 4 \\ -1 & 0 & 1 \end{pmatrix}$$

and $\ker(A) = V_2 = (2; 1; 2)$.

If the matrix A represents a homomorphism in $\hom(\mathbb{Z}^n, G)$ for an arbitrary FGA G, things are less simple. The columns of H still span the image of A, but they need not be independent generators, in which case H is not a monomorphism, so in order to compute kernels and images of general homomorphisms, we must be a bit more sophisticated.

Example 18 (Computing the image/coimage of a general homomorphism). Given a homomorphism in terms of m generators, $A \in \hom(\mathbb{Z}^m, G)$, we want to compute $\mathrm{im}(A) \in \mathrm{mono}(\mathbb{Z}_{p_1} \oplus \mathbb{Z}_{p_2} \oplus \cdots \oplus \mathbb{Z}_{p_k}, G)$ such that



The image splits in components $\mathrm{im}(A) = \alpha_1|\alpha_2|\cdots|\alpha_k$, where the components $\alpha_i \in \mathrm{mono}(\mathbb{Z}_{p_i}, G)$ are the generators. The numbers p_i are called the *order* of the generator, i.e. the smallest positive integer such that $p_i\alpha_i = 0$. The idea of the algorithm is to compute the generators α_i by recursion in the rank m of A. First we show that we can compute one generator. Then we show that if one generator is known, we can reduce the computation to finding the image of a rank $m-1$ map.

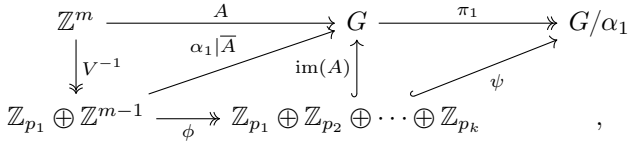
- **Computing one generator.** Start by eliminating the first row of A as in the Hermite normal form algorithm, obtaining

$$AV = (a_1|\overline{A}),$$

  where a_1 is the first column with a non-zero top element and $\overline{A} \in \mathrm{hom}(\mathbb{Z}^{m-1}, G)$ are the remaining columns eliminated to 0 on the top row. Clearly, a_1 is independent from $\overline{A}$, so it must be a generator. We compute p_1, the order of this generator, and find

$$\alpha_1 \in \mathrm{mono}(\mathbb{Z}_{p_1}, G),$$

  where a_1 and α_1 are represented by the same column vector.
- **Computing the rest by recursion in the rank.** The following diagram is of help in explaining the recursion step.



  where $\pi_1 = \mathrm{coker}(\alpha_1)$ and ψ is defined as

$$\psi = \pi_1 \circ \mathrm{im}\, A = 0|(\pi_1 \circ \alpha_2)|(\pi_1 \circ \alpha_3)|\cdots|(\pi_1 \circ \alpha_k),$$

  which is a monomorphism. Since ψ is mono, and ϕ is epi, we must have

$$\psi = \mathrm{im}(\pi_1 \circ (\alpha_1|\overline{A})) = \mathrm{im}(0|\pi_1 \circ \overline{A}) = 0|(\mathrm{im}(\pi_1 \circ \overline{A})),$$

  hence

$$\mathrm{im}(\pi_1 \circ \overline{A}) = (\pi_1 \circ \alpha_2)|(\pi_1 \circ \alpha_3)|\cdots|(\pi_1 \circ \alpha_k). \tag{12}$$

  Since $\pi_1 \circ \overline{A}$ is of lower rank, we get by recursion the image of this

$$\mathrm{im}(\pi_1 \circ \overline{A}) = \alpha'_2|\alpha'_3|\cdots|\alpha'_k,$$

and from (18) we obtain

$$\alpha_i = \pi_1 \backslash \alpha_i', \quad 2 \leqslant i \leqslant k,$$

from which we get the answer

$$\mathrm{im}(A) = \alpha_1 | \alpha_2 | \cdots | \alpha_k.$$

It will often happen that a column of A at some stage of the recursion becomes $\mathbf{0}$. In this case the column is swapped out to the right, and V^{-1} is replaced by the upper $(m-1) \times m$ block of V^{-1}.

- **Computing** $\mathrm{coim}(A)$**.** We have

$$\phi = \mathrm{coim}(\pi_1 \circ (\alpha_1 | \overline{A}))$$

Example 19. Let $G = \mathbb{Z}_4 \oplus \mathbb{Z}_{12}$. The matrix

$$A = \begin{pmatrix} 2 & 0 \\ 4 & 8 \end{pmatrix} \in \hom(\mathbb{Z}^2, \mathbb{Z}_4 \oplus \mathbb{Z}_{12})$$

is in Hermite normal form, but the columns are not independent, since

$$A \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in G.$$

Following the above procedure to compute $\mathrm{im}(A)$ we find

$$\alpha_1 = (2; 4) \in \mathrm{mono}(\mathbb{Z}_6, \mathbb{Z}_4 \oplus \mathbb{Z}_{12}),$$

and by the technique of Example 16 we compute

$$\pi_1 = \mathrm{coker}(\alpha_1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathrm{epi}(\mathbb{Z}_4 \oplus \mathbb{Z}_{12}, \mathbb{Z}_2 \oplus \mathbb{Z}_4).$$

Since $\pi_1(0; 8) = 0$, we are done (the projection discovered dependence between the generators), and we obtain

$$\mathrm{im}(A) = \alpha_1 = (2; 4) \in \mathrm{mono}(\mathbb{Z}_6, \mathbb{Z}_4 \oplus \mathbb{Z}_{12}).$$

Example 20. Compute $\mathrm{im}(A)$ for

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 5 \end{pmatrix} \in \hom(\mathbb{Z}^2, \mathbb{Z}_4 \oplus \mathbb{Z}_8).$$

- We eliminate first row

$$\begin{pmatrix} 1 & 1 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 3 & 2 \end{pmatrix},$$

yielding

$$\alpha_1 = (1; 3) \in \mathrm{mono}(\mathbb{Z}_8, \mathbb{Z}_4 \oplus \mathbb{Z}_8).$$

- From the Smith normal form decomposition

$$\begin{pmatrix} 1 & 4 & 0 \\ 3 & 0 & 8 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 4 & -1 & -1 \\ 0 & -2 & 1 \end{pmatrix}$$

  we find $\pi_1 = (-1, -1) \in \hom(\mathbb{Z}_4 \oplus \mathbb{Z}_8, \mathbb{Z}_4)$ and $\overline{A} = \pi_1(0; 2) = 2 \in \hom(\mathbb{Z}, \mathbb{Z}_4)$ and $\mathrm{im}(\overline{A}) = 2 \in \mathrm{mono}(\mathbb{Z}_2, \mathbb{Z}_4)$. Solving $\alpha_2 = \pi_1 \backslash \mathrm{im}(\overline{A})$ yields $\alpha_2 = (2; 4) \in \mathrm{mono}(\mathbb{Z}_2, \mathbb{Z}_4 \oplus \mathbb{Z}_8)$.
- We assemble and find

$$\mathrm{im}(A) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \hom(\mathbb{Z}_8 \oplus \mathbb{Z}_2, \mathbb{Z}_4 \oplus \mathbb{Z}_8).$$

Summary

In this section we have sketched the outline of a software system for doing general computations in the category of finitely generated abelian groups. We have introduced the main operations in such a package and indicated the algorithms behind the construction. In a forthcoming paper we will describe such a package in more detail.

3.3 Circulant matrices and the Discrete Fourier Transform (DFT)

To pave the road for later developments, we will start with a discussion of linear operators which are invariant under discrete circular shifts, and generalisations to finite abelian groups. This example has many of the properties of the general theory, but is simpler, since the spaces involved are finite dimensional vector spaces, and there are no questions of convergence. The classical notion of a *circulant matrix* is an $n \times n$ matrix

$$A = \begin{pmatrix} a_0 & a_{n-1} & \cdots & a_2 & a_1 \\ a_1 & a_0 & a_{n-1} & & a_2 \\ \vdots & a_1 & a_0 & a_{n-1} & \\ & & \ddots & \ddots & \ddots \\ a_{n-1} & & & a_1 & a_0 \end{pmatrix},$$

where the 'wrap-around' diagonals are constant, $A_{i,j} = a_{i-j \bmod n}$. The special circulant

$$S = \begin{pmatrix} & & & 1 \\ 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix},$$

where $a_1 = 1$ and $a_i = 0$ for $i \neq 1$ is called the *unit shift operator*. Let $\mathbb{C}[\mathbb{Z}_n]$ denote the vector space of all complex valued functions on $\mathbb{Z}_n$. The shift matrix S can be defined by its action on $\mathbb{C}[\mathbb{Z}_n]$

$$S\mathbf{x}(j) = \mathbf{x}(j-1) \quad \text{for all } j \in \mathbb{Z}_n.$$

Lemma 4. *For a matrix $A \in \mathbb{C}^{n\times n}$ the following are equivalent*

1. *A is circulant.*
2. *A is a polynomial in the shift operator*
$$A = \sum_{j\in\mathbb{Z}_n} a_j S^j.$$
3. *A acts on a vector $\mathbf{x} \in \mathbb{C}[\mathbb{Z}_n]$ through the* convolution product *$A\mathbf{x} = \mathbf{a} * \mathbf{x}$, defined as*
$$(\mathbf{a} * \mathbf{x})(j) := \sum_{\ell\in\mathbb{Z}_n} \mathbf{a}(\ell)\mathbf{x}(j-\ell).$$
4. *A is a linear translation invariant (LTI) operator on $\mathbb{C}[\mathbb{Z}_n]$, i.e. $AS = SA$.*
5. *The eigenvectors of A are $\{\chi_k\}_{\mathbf{k}\in\mathbb{Z}_n}$ given as*
$$\chi_k(j) = e^{2\pi ijk/n}.$$

The reader is encouraged to prove this result for this case of classical circulant matrices (over the cyclic group $\mathbb{Z}_n$). We generalise to the case of a general finite abelian group.

Definition 29 (Group ring). *Let G be a finite abelian group. The* group ring $\mathbb{C}[G]$ *is the vector space of all complex valued functions $a\colon G \to \mathbb{C}$.*

Alternatively (since G is finite), we can identify the group ring with the $\mathbb{C}$-linear combinations of elements of G,

$$\mathbb{C}[G] = \left\{ \sum_{j\in G} a(j)j \right\}.$$

The structure of the domain G being a group yields important additional structure of $\mathbb{C}[G]$. For any $t \in G$ we define the shift operator $S_t\colon \mathbb{C}[G] \to \mathbb{C}[G]$

$$(S_t a)(j) := a(j-t). \tag{13}$$

It is easy to check that the shifts define an *action* of G on $\mathbb{C}[G]$, i.e. we have $S_t S_u = S_{t+u}$, and furthermore this action is by linear transformations on a vector space. Such linear actions are called *group representations* and are fundamental objects in Fourier analysis (both on commutative and non-commutative groups).

Since G forms a basis for $\mathbb{C}[G]$, we can extend the product on G by linearity to a product $*\colon \mathbb{C}[G] \times \mathbb{C}[G] \to \mathbb{C}[G]$ called the *convolution*, given as

$$(a * b)(j) := \sum_{\ell \in G} a(\ell) b(j - \ell) = \sum_{\ell \in G} a(j - \ell) b(\ell). \tag{14}$$

The convolution product is associative and commutative, and the *delta-function* $\delta \in \mathbb{C}[G]$, defined such that

$$\delta(j) = \begin{cases} 1 \text{ if } j = 0 \\ 0 \text{ otherwise} \end{cases}$$

is the unit of the convolution product, satisfying $\mathbf{a} * \delta = \delta * \mathbf{a} = \mathbf{a}$ for all $a \in \mathbb{C}[G]$. onvolutions and translation invariant operators go hand-in-hand.

Definition 30 (Linear translation invariant operator (LTI)). *A mapping $A\colon \mathbb{C}[G] \to \mathbb{C}[G]$ is called LTI if it is linear and commutes with shifts*

$$AS_t = S_t A \quad \textit{for all } t \in G.$$

For a vector space V let $\mathrm{End}(V)$ denote the linear mappings $A\colon V \to V$ (endomorphisms). The elements of G form the natural basis for $\mathbb{C}[G]$, and with respect to this basis any $A \in \mathrm{End}(\mathbb{C}[G])$ is represented by a matrix $A_{i,j}$ for indices $i, j \in G$, such that $(A\mathbf{x})(i) = \sum_j A_{i,j} x(j)$. From this it is straightforward to verify the following result:

Lemma 5. *$A \in \mathrm{End}(\mathbb{C}[G])$ is LTI if and only if*

$$A_{i,j} = A_{i-t,j-t}$$

for all $i, j, t \in G$.

We can reconstruct an LTI A from its first column. Let $\mathbf{a} = A\delta \in \mathbb{C}[G]$, in coordinates $\mathbf{a}(i) = A_{i,0}$, then

$$A_{i,j} = \mathbf{a}(i - j).$$

We see that in the case $G = \mathbb{Z}_n$, the LTI operators are exactly the circulant matrices. Writing the matrix-vector product in terms of $\mathbf{a}$ we find

$$(A\mathbf{x})(i) = \sum_{\ell \in G} A_{i,\ell} \mathbf{x}(\ell) = \sum_{\ell \in G} \mathbf{a}(i - \ell)\mathbf{x}(\ell) = \mathbf{a} * \mathbf{x}.$$

Conversely, any linear operator defined in terms of a convolution must be LTI. Hence, we conclude

Lemma 6. *For a finite abelian group G a matrix $A \in \mathrm{End}(\mathbb{C}[G])$ is LTI if and only if it is given as a convolution*

$$A\mathbf{x} = \mathbf{a} * \mathbf{x}.$$

We want to understand the eigenvectors of convolutional operators. Recall that $\mathbb{T}$ is the multiplicative group of complex numbers on the unit circle.

Lemma 7. *The eigenvectors of a convolutional operator $A\mathbf{x} = \mathbf{a}*\mathbf{x}$ are exactly the non-zero homomorphisms $\chi \in \hom(G, \mathbb{T})$, i.e. the $\chi \in \mathbb{C}[G]$ such that*

$$\chi(j+k) = \chi(j)\chi(k) \quad \textit{for all } j, k \in G.$$

For $\chi \in \hom(G, \mathbb{T})$ we have $A\chi = \widehat{\mathbf{a}}(\chi) \cdot \chi$, where the eigenvalue $\widehat{\mathbf{a}}(\chi)$ is

$$\widehat{\mathbf{a}}(\chi) = \sum_{j \in G} \mathbf{a}(j)\overline{\chi(j)}.$$

Proof. We start by picking a $\chi \in \hom(G, \mathbb{T})$. Then, using $\chi(-j) = \overline{\chi(j)}$

$$(\mathbf{a} * \chi)(k) = \sum_{j \in G} \mathbf{a}(j)\chi(k-j) = \left(\sum_j \mathbf{a}(j)\chi(-j)\right)\chi(k) = \widehat{\mathbf{a}}(\chi)\chi(k).$$

When we compute $\hom(G, \mathbb{T})$ explicitly, we will see that $\hom(G, \mathbb{T}) \simeq G$. Since $\dim(\mathbb{C}[G]) = |G|$, this is a complete set of eigenvectors. □

The mapping $\mathbf{a} \mapsto \widehat{\mathbf{a}}$ is called the *discrete Fourier transform* (DFT), and can be understood as an expansion in the orthogonal basis for $\mathbb{C}[G]$ given by the eigenvectors $\hom(G, \mathbb{T})$, henceforth called the *characters* of G. Let $\langle \cdot, \cdot \rangle \colon \mathbb{C}[G] \times \mathbb{C}[G] \to \mathbb{C}$ denote the inner product on $\mathbb{C}[G]$

$$\langle f, g \rangle := \sum_{\ell \in G} \overline{f(\ell)} g(\ell).$$

Theorem 5 (Discrete Fourier Transform (DFT)).
Let $G = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_d}$ be a FAG. The characters $\hom(G, \mathbb{T})$ are in 1–1 correspondence with G; for every $k \in G$ there is a unique character $\chi_k \in \hom(G, \mathbb{T})$ given at $j \in G$ as

$$\chi_k(j) = \exp\left(2\pi i \left(\frac{k_1 j_1}{n_1} + \frac{k_2 j_2}{n_2} + \cdots + \frac{k_d j_d}{n_d}\right)\right).$$

The characters are orthogonal

$$\langle \chi_k, \chi_{k'} \rangle = \begin{cases} |G| & \textit{if } k = k' \\ 0 & \textit{otherwise} \end{cases}.$$

The discrete Fourier transform $\widehat{\ } \colon \mathbb{C}[G] \to \mathbb{C}[G]$ and its inverse are given as

$$\widehat{\mathbf{a}}(k) = \sum_{j \in G} \mathbf{a}(j)\overline{\chi_k(j)} = \langle \chi_k, \mathbf{a} \rangle \tag{15}$$

$$\mathbf{a}(j) = \sum_{k \in G} \widehat{\mathbf{a}}(k)\chi_k(j). \tag{16}$$

Proof. We first compute the characters on the group $\mathbb{Z}_n$. For any character χ we have that $\chi(0) = 1$. Now, since $n1 = 0$ in G, we find that $\chi(1)^n = \chi(0) = 1$, thus $\chi(1) = \exp(2\pi ik/n)$ for some $k \in \{0, 1, \dots n-1\}$. Let χ_k be the character with $\chi_k(1) = \exp(2\pi ik/n)$. Then $\chi_k(j) = \chi_k(1)^j = \exp(2\pi ijk/n)$. Thus the characters on $\mathbb{Z}_n$ are given as

$$\chi_k(j) = \exp(2\pi ijk/n) \text{ for } k \in \mathbb{Z}_n. \tag{17}$$

By the formula for a geometric sum, it is straightforward to verify the orthogonality

$$\langle \chi_k, \chi_{k'} \rangle = \begin{cases} |G| & \text{if } k = k' \\ 0 & \text{otherwise} \end{cases}.$$

(Orthogonality of characters is proven for general LCAs in the next section).

For $G = G_1 \oplus G_2$ we check that $\chi^1 \in \hom(G_1, \mathbb{T})$ and $\chi^2 \in \hom(G_2, \mathbb{T})$ produces a character $\chi = \chi^1 \oplus \chi^2 \in \hom(G, \mathbb{T})$, and furthermore

$$\langle \chi^1 \oplus \chi^2, \widetilde{\chi}^1 \oplus \widetilde{\chi}^2 \rangle = \langle \chi^1, \widetilde{\chi}^1 \rangle \cdot \langle \chi^2, \widetilde{\chi}^2 \rangle.$$

From this the characters on $G = \mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2} \oplus \cdots \oplus \mathbb{Z}_{n_d}$ and their orthogonality relations follow, thus $\hom(G, \mathbb{T})$ forms a complete orthogonal basis for $\mathbb{C}[G]$. The DFT and its inverse follow from the orthogonal expansion in $\mathbb{C}[G]$

$$\mathbf{a} = \sum_{\chi \in \hom(G,\mathbb{T})} \frac{\langle \chi, \mathbf{a} \rangle}{\langle \chi, \chi \rangle} \chi.$$

□

The basic facts that **LTI ⇔ convolutions**, that Fourier transforms *diagonalise convolutions*,

$$\widehat{(\mathbf{a} * \mathbf{b})}(\chi) = \hat{\mathbf{a}}(\chi) \cdot \hat{\mathbf{b}}(\chi)$$

and that the DFT can be computed blazingly fast using the Fast Fourier Transform (FFT) explains why the FFT is one of the most important algorithms in computational mathematics. In the sequel we discuss Fourier analysis on more general LCAs. It is only for finite G that we have a direct fast algorithms for computing the Fourier transform. Hence, a detailed understanding of the relationship between the continuous and the discrete Fourier analysis is crucial for computational Fourier analysis. We will detail these relationships using the language of group homomorphisms introduced above.

3.4 Fourier analysis on general LCAs

In this section we provide a quick survey of the general theory of Fourier analysis on general Locally Compact Abelian groups. The general theory has many of the properties of the finite case, but more care must be taken with respect to analytical properties of function spaces. the most important topic of these lectures.

Functions on G

For finite G the group ring $\mathbb{C}[G]$ is a well defined space of all functions on G. For infinite G we have to be more careful, due to convergence issues. We will use the following notation:

- $\mathbb{C}^G$**: Complex valued functions on** G**.** This space is too large to be useful for mathematical analysis and we use this notation when we want to convey an idea without being very accurate on convergence issues. Read this as *"an appropriate space of functions on G"*.
- $L^2(G)$**: Square integrable functions,**

$$L^2(G) = \left\{ f \in \mathbb{C}^G \colon \int_G |f(x)|^2 dx < \infty \right\},$$

  where $\int_G$ is defined below.
- $\mathcal{S}(G)$**: Schwartz functions**. Defined below, this space consists of rapidly decreasing, infinitely smooth functions.
- $\mathcal{S}'(G)$**: Tempered distributions**. Defined below, this is the dual space of $\mathcal{S}(G)$ and consists of generalised functions such as the Dirac δ (point mass).

Shifts, integrals and convolutions

Let G be an LCA. Shifts of functions are defined as

$$(S_y f)(x) = f(x - y) \text{ for } x, y \in G \text{ and } f \in \mathbb{C}^G. \tag{18}$$

In [24] it is shown that for any LCA there exists a non negative measure μ which is shift invariant, i.e.

$$\mu(E) = \mu(E + x) \geqslant 0$$

for all Borel sets E and all $x \in G$, and $\mu(E) > 0$ for some E. This is called the Haar measure, and is unique up to a scaling. From this we obtain a shift invariant integral on G which we will just write as $\int_G \cdot dx$. For any integrable function f it satisfies:

$$\int_G f(x)dx = \int_G S_y f(x) \text{ for all } y \in G. \tag{19}$$

Example 21. For the LCAs of Definition 22 the invariant integrals are are:

$$\mathbb{R} : \int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx \text{ (the standard integral)}$$
$$T : \int_{T} f(x)dx = \int_{0}^{1} f(x)dx \text{ (the standard integral)}$$
$$\mathbb{Z} : \int_{\mathbb{Z}} f(x)dx = \sum_{j=-\infty}^{\infty} f(j)$$
$$\mathbb{Z}_n : \int_{\mathbb{Z}_n} f(x)dx = \sum_{j=0}^{n-1} f(j)$$

For direct products $G = G_1 \times G_2$ it is obtained as a multiple integral

$$\int_G f(x)dx = \int_{G_1}\int_{G_2} f(x,y)dxdy.$$

For any finitely generated group G, the integral notation means the discrete sum

$$\int_G f(x)dx = \sum_{j\in G} f(j).$$

From the integral we get two important products on $\mathbb{C}^G$, the inner product and the convolution. The inner product $\langle\cdot,\cdot\rangle : \mathbb{C}^G \times \mathbb{C}^G \to \mathbb{C}$ is defined as

$$\langle f, g\rangle = \int_G \overline{f(x)}g(x)dx \tag{20}$$

where $\overline{f}$ denotes the complex conjugate. We will sometimes write $\langle f, g\rangle_{\mathbb{C}^G}$ to emphasize on which domain we consider the innerproduct.

The convolution product $*\ : \mathbb{C}^G \times \mathbb{C}^G \to \mathbb{C}^G$ is defined as

$$(f * g)(y) = \int_G f(x)g(y-x)dx. \tag{21}$$

Note that the convolution can be understood as a weighted linear combination of shifts. This is evident in the finite case, $(f*g)(y) = \sum_{x\in G} f(x)g(y-x)$ thus $f * g = \sum_{x\in G} f(x)S_x g$. The various shifts $S_x g$ are multiplied with the weights $f(x)$. By a change of variables we verify that $f * g = g * f$, so one may also think of g as being the weights and f the function that is shifted.

Convolutions are ubiquitous in computational mathematics, common examples being finite difference approximations and linear digital filters. As a rule of thumb, convolutions are important whenever a problem is invariant under shifts. To be more precise, we say that a *linear* operator $A : \mathbb{C}^G \to \mathbb{C}^G$ is *translation invariant* (LTI) if $AS_g = S_g A$ for all $g \in G$. Thus linear differential operators with constant coefficients such as d/dx and ∇^2 are examples of LTI operators on $\mathbb{C}^{\mathbb{R}^n}$.

In the case of finite G, we saw that LTI operators are the same as convolutions. This is generally not the case for infinite G. E.g. there exists no (classical) function $f \in \mathbb{C}^{\mathbb{R}}$ such that $f * g = dg/dx$ for all differentiable g, and there is no (classical) function being the identity of convolution, $\delta * g = g$. However, there are various ways of *approximating* LTI operators on $\mathbb{C}^{\mathbb{R}}$ by convolutions, and the convolutional identity exists as a distribution $\delta \in \mathcal{S}'(G)$. So, we think of LTI and convolutional operators as being *essentially the same*, also for infinite G.

The dual group

Since all shifts commute, they share a common set of eigenfunctions. Convolutions are linear combinations of shifts, and do hence also share the same eigenfunctions. These are called the *characters* of the group. We will see that the characters form an orthogonal basis for $L^2(G)$, the square integrable functions on G. The Fourier transform is an expansion of functions in this basis. In the Fourier basis all convolutions become diagonal matrices. This diagonalizing property is the most important property of the Fourier transform. We will in this section see that the space of all Fourier coefficients also has the structure of an abelian group. It is called the dual group.

As above, let $\mathbb{T}$ denote the unitary complex numbers

$$\mathbb{T} = \{\, z \in \mathbb{C} \mid |z| = 1 \,\}. \tag{22}$$

As a multiplicative abelian group $\mathbb{T}$ is isomorphic with T, via the map $T \ni x \mapsto \exp(2\pi i x) \in \mathbb{T}$.

Definition 31 (Group character). *A character on a group G is a (continuous[5]) homomorphism $\chi \in \hom(G, \mathbb{T})$ i.e.*

$$\chi(x + y) = \chi(x)\chi(y) \text{ for all } x, y \in G. \tag{23}$$

Note that $(S_y\chi)(x) = \chi(x - y) = \chi(-y)\chi(x)$, which shows that the characters are eigenfunctions of shifts. In Theorem 5 we have found that for G finite, the characters are in 1–1 correspondence with G itself.

Example 22. We want to find the characters on $\mathbb{R}$. We have $\chi(x + t) = \chi(t)\chi(x)$, differentiation with respect to t at $t = 0$ yields

$$\chi'(x) = \chi'(0)\chi(x).$$

Since $|\chi(x)| = 1$ we must have $\chi'(0) = i\omega$ for some $\omega \in \mathbb{R}$. Combined with $\chi(0) = 1$ this yields the complete family of continuous characters

$$\chi_\omega(x) = \exp(i\omega x) \text{ for } \omega \in \mathbb{R}. \tag{24}$$

[5] For topological groups $\hom(G, H)$ denotes the continuous homomorphisms

To complete the argument we have to show that every continuous $\chi(x)$ is differentiable. We can always choose a small $\delta > 0$ such that $\int_0^\delta \chi(t)dt = \alpha > 0$. Thus

$$\alpha \cdot \chi(x) = \chi(x) \int_0^\delta \chi(t)dt = \int_0^\delta \chi(x+t)dt = \int_x^{x+\delta} \chi(t)dt.$$

The right hand side is the integral of a continuous function, and is thus differentiable. Hence $\chi(x)$ is differentiable.

Example 23. Let us compute the continuous characters on the unit circle $T = \mathbb{R}/\mathbb{Z}$. Let $x \in T$ be an irrational number, thus the sequence $x, 2x, 3x, \ldots$ fills a dense subset of T. Once we have fixed the value of a character at x, we can derive the value of the character on this dense subset, $\chi(jx) = \chi(x)^j$. If we require $\chi(x)$ to be continuous, we can extend it uniquely to the whole of T. We leave it to the reader to verify that the resulting continuous characters on T are given as

$$\chi_k(x) = \exp(2\pi i k x) \text{ for } k \in \mathbb{Z}. \tag{25}$$

One may alternatively arrive at the same result using the technique of the previous example. Note that if we did not have the condition that the characters should be continuous functions, we would get an awful lot of them, since we could make a separate choice of χ on each coset of $\mathbb{Q}$ in $\mathbb{R}$.

We define the product of two characters as

$$(\chi_k \cdot \chi_l)(x) = \chi_k(x) \cdot \chi_l(x). \tag{26}$$

The product is obviously commutative. A simple computation shows that $(\chi_k \cdot \chi_l)(x+y) = (\chi_k \cdot \chi_l)(x) \cdot (\chi_k \cdot \chi_l)(y)$, thus also the product is a character.

Definition 32 (Dual group). *Let G be an LCA. The dual group Γ is defined as the collection of all (continuous) characters on G with the product (26). Γ has a natural topology turning it into an LCA*[6].

A natural question to ask is *what is the dual of the dual group?* For a given $x \in G$ and $\chi \in \Gamma$, let $\psi_x(\chi) = \chi(x)$. Since

$$\psi_x(\chi_k \cdot \chi_l) = \chi_k(x)\chi_l(x) = \psi_x(\chi_k) \cdot \psi_x(\chi_l),$$

we see that ψ_x is a character on Γ. It is also easy to verify that $\psi_x \cdot \psi_y = \psi_{x+y}$, thus G can at least be identified with a subgroup of the group of characters on Γ. If topology is taken into the picture it can be shown that G and the dual of Γ are isomorphic as LCAs, see [24].

Theorem 6 (Pontryagin duality). *The identification of $x \in G$ with the character $\psi_x(\chi) = \chi(x)$ is an isomorphism between G and the dual of Γ.*

[6] It is given the weakest topology such that for any $x \in G$, the map $k \mapsto \chi_k(x) \colon \Gamma \to \mathbb{T}$ is continuous in k, see [24].

Example 23 showed that the dual of T is isomorphic with $\mathbb{Z}$ under the map $\mathbb{Z} \ni k \mapsto \chi_k(\cdot) = \exp(2\pi i k \cdot)$ and Theorem 6 implies that the dual of $\mathbb{Z}$ is naturally isomorphic to T. In order to recover the characters on T and on $\mathbb{Z}$, we define the function $(\cdot,\cdot)\colon \mathbb{Z}\times T \to \mathbb{T}$ as

$$(k,x) \equiv \chi_k(x) = \exp(2\pi i k x).$$

If we fix k then $(k,\cdot)$ gives us all the characters on T, and when x is fixed we get all the characters $(\cdot,x)$ on $\mathbb{Z}$. Thus we may simply say that T and $\mathbb{Z}$ are dual spaces, where the characters are recovered by *the dual pairing* $(\cdot,\cdot)$.

Definition 33 (Dual pair). *Two LCAs G and $\widehat{G}$ are called a* dual pair of LCAs *if there exists a continuous function $(\cdot,\cdot) : \widehat{G}\times G \to \mathbb{T}$ such that the map*

$$\widehat{G} \ni k \mapsto (k,\cdot) \in \mathbb{C}^G$$

is an LCA isomorphism between $\widehat{G}$ and the dual of G, and the map

$$G \ni x \mapsto (\cdot,x) \in \mathbb{C}^{\widehat{G}}$$

is an LCA isomorphism between G and the dual of $\widehat{G}$.

In particular the reader is encouraged to verify the following identities:

$$(k+k',x) = (k,x)\cdot(k',x) \tag{27}$$
$$(k,x+x') = (k,x)\cdot(k,x') \tag{28}$$
$$(0,x) = (k,0) = 1 \tag{29}$$
$$\overline{(k,x)} = (-k,x). \tag{30}$$

Furthermore, if $(k,x) = 1$ for all k then $x = 0$, and if $(k,x) = 1$ for all x then $k = 0$.

Since $T = \mathbb{R}/\mathbb{Z}$ is isomorhpic to $\mathbb{T}$ through the exponential mapping $x \mapsto \exp(2\pi i x)$, we will often present the dual pairing in its *bi-additive form* $\langle\cdot,\cdot\rangle\colon \widehat{G} \times G \to T$ such that

$$(k,x) = \exp\left(2\pi i \langle k,x\rangle\right).$$

This satisfies the following equations

$$\langle k+k',x\rangle = \langle k,x\rangle + \langle k',x\rangle \tag{31}$$
$$\langle k,x+x'\rangle = \langle k,x\rangle + \langle k,x'\rangle \tag{32}$$
$$\langle 0,x\rangle = \langle k,0\rangle = 0 \tag{33}$$
$$\langle k,x\rangle = 0 \quad \forall k \Leftrightarrow x = 0 \tag{34}$$
$$\langle k,x\rangle = 0 \quad \forall x \Leftrightarrow k = 0, \tag{35}$$

thus we can think of $\langle\cdot,\cdot\rangle$ as an abelian group version of a non-degenerate bilinear pairing between vector spaces.

The Fourier transform

The main goal of this section is to study the expansion of functions $f \in \mathbb{C}^G$ in terms of characters (the Fourier basis),

$$f(x) = \int_{\widehat{G}} \widehat{f}(k)(k,x)dx = \int_{\widehat{G}} \widehat{f}(k)e^{2\pi i\langle k,x\rangle}dx \quad \text{for some } \widehat{f}(k) \in \mathbb{C}^{\widehat{G}}. \tag{36}$$

If G is finite, than *any* $f \in \mathbb{C}^G$ can be expanded in this basis, but this is not generally true for infinite G. Necessary and sufficient conditions for functions to be expressible in terms of Fourier series is discussed in many textbooks on Fourier analysis, see e.g. [11].

The following lemma shows that the Fourier basis is orthogonal.

Lemma 8. *The characters are orthogonal under the inner product defined in (20)*

$$\langle (k,\cdot),(l,\cdot)\rangle = \int_G \overline{(k,x)} \cdot (l,x)dx = 0 \text{ when } k \neq l. \tag{37}$$

Proof. Assume $k \neq l$.

$$\int_G (-k,x)\cdot(l,x)dx = \int_G (l-k,x)dx = \int_G (m,x)dx$$

where $m \neq 0$. Pick a point $x_0 \in G$ such that $(m,x_0) \neq 1$. Using the invariance of the integral, we find

$$\int_G (m,x)dx = (m,x_0)\int_G (m,x-x_0)dx = (m,x_0)\int_G (m,x)dxdx.$$

Hence, $\int_G (m,x)dx = 0$. □

Definition 34 (Fourier transform). *The Fourier transform is a linear map* $\widehat{\ }: \mathbb{C}^G \to \mathbb{C}^{\widehat{G}}$ *given as*

$$\widehat{f}(k) = \langle (k,\cdot), f(x)\rangle_G = \int_G (-k,x)f(x)dx. \tag{38}$$

We also use the alternative notation

$$\mathcal{F}_G[f] := \widehat{f} \tag{39}$$

to specify the domain G explicitly.

Inversion of the Fourier transform is simple due to orthogonality of the characters. If G is compact then $\widehat{G}$ is discrete, Theorem 12, and the integral over $\widehat{G}$ is given as a sum $\int_{\widehat{G}} g(k)dk = \sum_{k\in\widehat{G}} g(k)$.

Lemma 9. *If G is compact, then*

$$f(x) = \frac{1}{C}\int_G \widehat{f}(k)(k,x)dk = \frac{1}{C}\sum_{k\in\widehat{G}} \widehat{f}(k)(k,x)dk, \tag{40}$$

where $C = \int_G 1dx$.

Proof. Given f as in (36). Using the orthogonality of the characters we find

$$\begin{aligned}\langle (k,\cdot), f(\cdot)\rangle_G &= \int_{x\in G}(-k,x)\sum_{\ell\in\widehat{G}} g(\ell)(\ell,x)dx \\ &= \sum_{\ell\in\widehat{G}} g(\ell)\int_{x\in G}(-k,x)(\ell,x)dx = g(k)\int_G 1dx,\end{aligned}$$

thus $g(k) = \frac{1}{C}\widehat{f}(k)$. □

A similar result holds also in the general case, see [24] for a proof:

Theorem 7 (Fourier reconstruction). *Given any LCA G there exists a constant C so that Fourier reconstruction is given as*

$$f(x) = \frac{1}{C}\int_G \widehat{f}(k)(k,x)dk. \tag{41}$$

As stated in the beginning of this section, the most fundamental property of the Fourier transform is the diagonalization of convolutions:

Theorem 8 (Convolution theorem).

$$\widehat{(f*g)}(k) = \widehat{f}(k)\widehat{g}(k). \tag{42}$$

Proof.

$$\begin{aligned}\widehat{(f*g)}(k) &= \int_G (f*g)(x)(-k,x)dx = \int_G\int_G f(x-y)g(y)(k,-x)dxdy \\ &= \int_G f(x-y)(k,-x+y)dx\int_G g(y)(k,-y)dy = \widehat{f}(k)\widehat{g}(k).\end{aligned}$$

□

A very related result is the following, which states that a shift of a function $f\in\mathbb{C}^G$ corresponds to a multiplication of $\widehat{f}$ by a character on $\widehat{G}$, while a multiplication of f by a character on G corresponds to a shift of $\widehat{f}$. The proof is a straight forward computation left as an exercise.

Theorem 9 (Shift formulas). *Let $\chi_k = (k,\cdot)$ and $\chi_x = (\cdot,x)$ be characters on G and $\widehat{G}$. Let S_x and S_k be shifts on $F(G)$ and $F(\widehat{G})$ defined in (18). Then*

$$\widehat{S_x f} = \chi_{-x}\cdot\widehat{f} \tag{43}$$

$$S_k\widehat{f} = \widehat{\chi_k\cdot f}. \tag{44}$$

The final fundamental result of this section states that the Fourier transform $\widehat{\ }: \mathbb{C}^G \to \mathbb{C}^{\widehat{G}}$ preserves the inner product on the two spaces. It bears the name of Parceval or Plancherel, depending on whether or not f and g are equal.

Theorem 10 (Parseval – Plancherel). *Let C be the constant of the Fourier inversion (41). Then*

$$\int_G \overline{f}(x)g(x)dx = \frac{1}{C}\int_{\widehat{G}} \overline{\widehat{f}}(k)\widehat{g}(k)dk. \tag{45}$$

Proof.

$$\begin{aligned}\int_G \overline{f}(x)g(x)dx &= \int_G \overline{f}(x)\frac{1}{C}\int_{\widehat{G}} \widehat{g}(k)(k,x)dkdx \\ &= \frac{1}{C}\int_{\widehat{G}}\int_G \overline{f}(x)(k,x)dx\widehat{g}(k)dk = \frac{1}{C}\int_{\widehat{G}} \overline{\widehat{f}}(k)\widehat{g}(k)dk.\end{aligned}$$

□

Example 24. **Fourier analysis on the classical groups.** The following table presents the group and dual groups, the dual pairing, the Fourier transform and reconstruction for the basic groups $\mathbb{R}$, T, $\mathbb{Z}$ and $\mathbb{Z}_n$.

G	$\widehat{G}$	$(\cdot,\cdot)$	$\widehat{f}(\cdot)$	$f(\cdot)$
$x \in \mathbb{R}$	$\omega \in \mathbb{R}$	$e^{2\pi i\omega x}$	$\int_{-\infty}^{\infty} e^{-2\pi i\omega x} f(x)dx$	$\int_{-\infty}^{\infty} e^{2\pi i\omega x}\widehat{f}(\omega)d\omega$
$x \in T$	$k \in \mathbb{Z}$	$e^{2\pi ikx}$	$\int_0^1 e^{-2\pi ikx} f(x)dx$	$\sum_{k=-\infty}^{\infty} e^{2\pi ikx}\widehat{f}(k)$
$j \in \mathbb{Z}_n$	$k \in \mathbb{Z}_n$	$e^{\frac{2\pi ikj}{n}}$	$\sum_{j=0}^{n-1} e^{\frac{-2\pi ikj}{n}} f(j)$	$\frac{1}{n}\sum_{k=0}^{n-1} e^{\frac{2\pi ikj}{n}}\widehat{f}(k)$

Multidimensional versions are given by the componentwise formulae:

$$\begin{aligned} x &= (x_1,x_2) \in G = G_1 \oplus G_2 \\ k &= (k_1,k_2) \in \widehat{G} = \widehat{G_1} \oplus \widehat{G_2} \\ (k,x) &= (k_1,x_1)\cdot(k_2,x_2) \\ \widehat{f}(k_1,k_2) &= \int_{G_1}\int_{G_2} (-k_1,x_1)(-k_2,x_2)f(x_1,x_2)dx_1dx_2 \\ f(x_1,x_2) &= \frac{1}{C_1C_2}\int_{\widehat{G_1}}\int_{\widehat{G_2}} (k_1,x_1)(k_2,x_2)f(k_1,k_2)dk_1dk_2. \end{aligned}$$

This gives the explicit form of the multidimensional transforms

G	$\widehat{G}$	$\langle\cdot,\cdot\rangle$	$\widehat{f}(\cdot)$	$f(\cdot)$
$x \in \mathbb{R}^n$	$\omega \in \mathbb{R}^n$	$\sum_\ell x_\ell\omega_\ell$	$\int_{-\infty}^{\infty} e^{-2\pi i\langle\omega,x\rangle} f(x)dx$	$\int_{-\infty}^{\infty} e^{2\pi i\langle\omega,x\rangle}\widehat{f}(\omega)d\omega$
$x \in T^n$	$k \in \mathbb{Z}^n$	$\sum_\ell x_\ell k_\ell$	$\int_{T^n} e^{-2\pi i\langle k,x\rangle} f(x)dx$	$\sum_{k\in\mathbb{Z}^n} e^{2\pi i\langle k,x\rangle}\widehat{f}(k)$
$j \in \mathbb{Z}_{\mathbf{m}}$	$k \in \mathbb{Z}_{\mathbf{m}}$	$\sum_{\ell=1}^n \frac{j_\ell k_\ell}{m_\ell}$	$\sum_{j\in\mathbb{Z}_{\mathbf{m}}} e^{-2\pi i\langle k,j\rangle} f(j)$	$\frac{1}{M}\sum_{k\in\mathbb{Z}_{\mathbf{m}}} e^{2\pi i\langle k,j\rangle}\widehat{f}(k)$

where $\mathbb{Z}_{\mathbf{m}} = \mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2} \oplus \cdots \oplus \mathbb{Z}_{m_n}$ and $M = \prod_{\ell=1}^n m_i$ is the number of grid points in $\mathbb{Z}_{\mathbf{m}}$.

Schwartz space and tempered distributions

For a proper discussion of sampling we need to introduce Schwartz functions and tempered distributions. The set $\mathcal{S}(G)$ of Schwartz functions on the elementary groups are defined as follows:

- On a finite G the Schwartz functions are all functions in $\mathbb{C}[G]$.
- On $\mathbb{Z}^n$ the Schwartz functions are the functions that decrease faster than polynomially towards infinity.
- On T^n, the Schwartz functions are $C^\infty(T)$, the smooth functions.
- On $\mathbb{R}^n$ the Schwartz functions are those $f \in C^\infty(\mathbb{R})$ such that both $f(x)$ and $\widehat{f}(\xi)$ decrease fast (faster than polynomially) as $x, \xi \to \infty$.

On a general LCA G there is also a notion of such functions called *Schwartz–Bruhat functions*. These can be defined as the functions f on G, such that both f and $\widehat{f}$ have rapidly decreasing L^∞-norms [23]. Important for introducing these functions are the following properties:

- $\mathcal{S}(G)$ is closed under sums, products, translations and convolutions.
- $\mathcal{S}(G)$ is dense in $L^p(G)$ for all $1 \leqslant p \leqslant \infty$.
- The space of *bump functions* $C_c^\infty(G)$ (smooth functions with compact support) is dense in $\mathcal{S}(G)$.
- The Fourier transform is a linear isomorphism between $\mathcal{S}(G)$ and $\mathcal{S}(\widehat{G})$.
- If $\phi \in \mathrm{mono}(H, G)$ and $f \in \mathcal{S}(G)$ then $f \circ \phi \in \mathcal{S}(H)$.

The *tempered distributions* $\mathcal{S}'(G)$ is the set of all linear functionals on $\mathcal{S}(G)$, linear mappings from $\mathcal{S}(G)$ to $\mathbb{C}$. For $T \in \mathcal{S}'(G)$ and $\phi \in \mathcal{S}(G)$ it is convenient to use the notation $\langle T, \phi \rangle$ for the evaluatoin of T at ϕ. Every measurable function $f \colon G \to \mathbb{C}$ growing slowly (not faster than polynomial) defines a distribution $T_f \in \mathcal{S}'(G)$ via the integral

$$\langle T_f, \phi \rangle := \int_{x \in G} f(x)\phi(x)dx \quad \text{for all } \phi \in \mathcal{S}(G).$$

These are called the *regular distributions*. There are, however, also other (singular) distributions which are not given by classical functions, such as the Dirac function $\delta(x)$, which physicists interpret as a unit mass in 0 such that $\int_{x \in G} \delta(x)\phi(x)dx = \phi(0)$. Such a function $\delta(x)$ is not a classical function, and the correct way to think of this is as the functional $\delta \in \mathcal{S}'(G)$ defined such that

$$\langle \delta, \phi \rangle := \phi(0) \quad \text{for all } \phi \in \mathcal{S}(G).$$

We can define Fourier transforms and derivatives of distributions by dualisation. The regular distributions give the hint on the correct definitions. For smooth slowly growing functions $f(x)$ on $G = \mathbb{R}$, integration by parts yields

$$\int f(x)\frac{d\phi(x)}{dx}dx = \int -\frac{df}{dx}\phi(x)dx.$$

For this reason, we *define* the derivative of $T \in \mathcal{S}'(\mathbb{R})$ as

$$\langle \frac{dT}{dx}, \phi(x) \rangle := \langle T, -\frac{d\phi(x)}{dx} \rangle. \tag{46}$$

Similarly, for a nice function $f \in \mathbb{C}^G$ we have, by Plancherel's theorem

$$\langle f, \check{\phi} \rangle_G = \frac{1}{C} \langle \widehat{f}, \phi \rangle_{\widehat{G}},$$

where $\check{}: \mathbb{C}^{\widehat{G}} \to \mathbb{C}^G$ denotes the inverse Fourier transform. Thus, we *define* the Fourier transform of a distribution $T \in \mathcal{S}'(G)$ as

$$\langle \widehat{T}, \phi \rangle := C \langle T, \check{\phi} \rangle, \tag{47}$$

where C is the constant from Plancherel's theorem.

Exercise 6. Check that this definition implies $\widehat{\delta} = 1$.

We refer to [11] for more details on tempered distributions.

Pullback and pushforward of functions on groups

The central topic of our lectures are the relationship between functions defined on a group and related functions on a subgroup and the quotient.

Definition 35 (Pullback and pushforward of functions).
For $\phi \in \hom(H, G)$ we define pullback $\phi^*: \mathbb{C}^G \to \mathbb{C}^H$ and pushforward $\phi_*: \mathbb{C}^H \to \mathbb{C}^G$ as adjoint operators with respect to the inner products

$$\phi^*(f) := f \circ \phi \tag{48}$$

$$\langle \phi_*(g), f \rangle_{\mathbb{C}^G} := \langle g, \phi^*(f) \rangle_{\mathbb{C}^H} \tag{49}$$

for $f \in \mathbb{C}^G$ and $g \in \mathbb{C}^H$.

Exercise 7. Show that for $\phi \in \hom(G_1, G_2)$ where G_1, G_2 are finite, we have

$$\phi_* f(\phi(j)) = \sum_{k \in \ker(\phi)} f(j + k)$$

$$\phi_* f(\ell) = 0 \quad \text{for } \ell \notin \operatorname{im}(\phi).$$

The pullback is easy to understand, e.g. if $\phi \in \operatorname{mono}(H, G)$ defines a discrete lattice, then $\phi^* f$ is the sampling of f in the lattice points $\phi(H) < G$. The push forward $\phi_* f$ along $\phi \in \operatorname{epi}(G, K)$ is summing up the values of f in all points of G mapping to the same point in K. It can be shown that for $\phi \in \operatorname{epi}(G, K)$ there always exists a constant C (depending on the choice of Haar measures on G and K) such that

$$\phi_* f(\phi(x)) = C \int_{y \in \ker(\phi)} f(x + y) dy.$$

It follows that:

Lemma 10. *For $\psi \in \mathrm{mono}(H,G)$ pullback is well-defined for Schwartz functions, $\psi^*\colon \mathcal{S}(G) \to \mathcal{S}(H)$. For $\phi \in \mathrm{epi}(G,K)$ pushforward is well-defined for Schwartz functions $\phi_*\colon \mathcal{S}(G) \to \mathcal{S}(K)$.*

Example 25. Pullback along epimorphisms does not in general send Schwartz functions to Schwartz functions, for example take $\phi \in \mathrm{epi}(\mathbb{R},T)$ as $\phi(x) = x$. The constant function $f(x) = 1 \in \mathcal{S}(T)$, but $\phi^*(1) = 1 \notin \mathcal{S}(\mathbb{R})$. The result is, however, a distribution in $\mathcal{S}'(\mathbb{R})$. Similarly, pushforward along monomorphisms is in general well-defined for distributions and not for Schwartz functions.

Definition 36 (Pullback and pushforward of distributions).
For $\psi \in \mathrm{mono}(H,G)$ we define pushforward of distributions $\psi_*\colon \mathcal{S}'(H) \to \mathcal{S}'(G)$ as

$$\langle \psi_* T, f\rangle := \langle T, \psi^* f\rangle \quad \text{for all } f \in \mathcal{S}(G).$$

For $\phi \in \mathrm{epi}(G,K)$ we define pullback of distributions $\phi^*\colon \mathcal{S}'(K) \to \mathcal{S}'(G)$ as

$$\langle \phi^* T, f\rangle := \langle T, \phi_* f\rangle \quad \text{for all } f \in \mathcal{S}(G).$$

Example 26. Let $\phi \in \mathrm{mono}(\mathbf{0},\mathbb{R})$ (the $\mathbf{0}$-arrow). Since

$$\langle \phi_* 1, f\rangle = \langle 1, \phi^* f\rangle = f(0),$$

we have $\phi_* 1 = \delta$, the Dirac distribution on $\mathbb{R}$. There is nothing particular about $\mathbb{R}$ here, indeed for any LCA G we have the following equivalent characterisations of the δ-distribution:

$$\delta = \mathbf{0}_* 1 \in \mathcal{S}'(G) \Leftrightarrow \langle \delta, f\rangle = f(0). \tag{50}$$

3.5 Duality of subgroups and quotients

Dual homomorphisms

Recall the discussion above, for an LCA G, the dual group $\widehat{G}$ is isomorphic to $\mathrm{hom}(G,\mathbb{T})$, which contains all eigen functions of the shift operators S_t acting on $\mathbb{C}^G$, or as mappings to the additive group $T = \mathbb{R}/\mathbb{Z}$ we have $\widehat{G} \simeq \mathrm{hom}(G,T)$. Similar to the adjoint of a linear mapping, we define the adjoint of a LCA homomorphism

Definition 37 (Dual homomorphism). *Given $\phi \in \mathrm{hom}(H,G)$ the dual homomorphism $\widehat{\phi} \in \mathrm{hom}(\widehat{G},\widehat{H})$ is defined for $\widehat{H} = \mathrm{hom}(H,T)$ and $\widehat{G} = \mathrm{hom}(G,T)$, acting on an element $\alpha \in \mathrm{hom}(G,T)$ as*

$$\widehat{\phi}(\alpha) = \alpha \circ \phi \in \mathrm{hom}(H,T).$$

Equivalently, for dual pairs G, $\widehat{G}$ and H, $\widehat{H}$ with pairings $\langle\cdot,\cdot\rangle_G\colon \widehat{G} \times G \to T$ and $\langle\cdot,\cdot\rangle_H\colon \widehat{H} \times H \to T$, we define

$$\langle \widehat{\phi}(\alpha), h\rangle_H = \langle \alpha, \phi(h)\rangle_G,$$

for all $\alpha \in \widehat{G}$ and $h \in H$.

Example 27. Let $H = \mathbb{Z}^n$ and $\widehat{H} = T^n$ with pairings $\langle \xi, \mathbf{j} \rangle_{\mathbb{Z}^n} = \xi^T \mathbf{j} \bmod 1$ and $\langle \xi, \mathbf{x} \rangle_{\mathbb{R}^n} = \xi^T \mathbf{x} \bmod 1$. A non-singular matrix $A \in \mathbb{R}^{n \times n}$ defines a homomorphism $\phi \in \hom(\mathbb{Z}^n, \mathbb{R}^n)$ as $\phi(\mathbf{j}) = A\mathbf{j}$. The dual homomorphism $\widehat{\phi} \in \hom(\mathbb{R}^n, T^n)$ is given as $\widehat{\phi}(\xi) = A^T \xi \bmod 1$. If the columns of A are linearly independent then ϕ is a monomorphism and $\widehat{\phi}$ an epimorphism.

The fundamental duality theorem

The main topic of this section is a theorem relating subgroup and quotient decompositions of a group to decompositions of the dual spaces. To prepare for this we discuss duality of sequences of homomorphisms in general. A *chain complex* is a sequence of groups G_i and homomorphisms $\phi_i \in \hom(G_i, G_{i+1})$ such that $\phi_{i+1} \circ \phi_i = 0$ for all i. A *co-chain complex* is similarly defined, where the indices decrease rather than increase. Recall that the the chain complex is *exact* if $\mathrm{im}(\phi_i) = \ker(\phi_{i+1})$ for all i, and similarly for the co-chain. An equivalent way of defining exactness is to say that whenever $x \in G_{i+1}$ such that $\phi_{i+1}(x) = 0$, there exists an $y \in G_i$ such that $x = \phi_i(y)$.

Lemma 11. *If (ϕ_i, G_i) is a chain complex, then the dual $(\widehat{\phi}_i, \widehat{G_i})$ is a co-chain complex, and if one of them is exact, then also the other is exact.*

Proof. Let $\widehat{G_i} = \hom(G_i, T)$. For $\alpha \in G_{i+2}$, we see $(\widehat{\phi}_i \circ \widehat{\phi_{i+1}})(\alpha) = \alpha \circ \phi_{i+1} \circ \phi_i = 0$, hence $(\widehat{\phi}_i, \widehat{G_i})$ is a co-chain complex.

To prove the statement about exactness, assume (ϕ_i, G_i) exact. We pick a $\chi \in \widehat{G_{i+1}}$ such that $\phi_i^*(\chi) = \chi \circ \phi_i = 0$ and want to show that there exists a $\chi' \in G_{i+2}^*$ such that $\widehat{\phi_{i+1}}(\chi') = \chi$. Pick an $x \in G_{i+1}$ such that $\phi_{i+1}(x) = 0$. Exactness implies that $x = \phi_i(y)$ for some $y \in G_i$, hence $\chi(x) = (\chi \circ \phi_i)(y) = 0$. Thus $\chi \circ \ker(\phi_{i+1}) = 0$, and since it is zero on the kernel we can solve the equation $\chi' = \chi / \phi_{i+1}$ for $\chi' \in \widehat{G_{i+2}}$. This proves exactness of $(\widehat{\phi}_i, \widehat{G_i})$.

If $(\widehat{\phi}_i, \widehat{G_i})$ is exact then (ϕ_i, G_i) must be exact because of Pontryagin duality. □

A very important consequence of this lemma is the following theorem, which is fundamental for the understanding of sampling theory and computational Fourier transforms:

Theorem 11 (Fundamental Duality Theorem of LCAs).
A short sequence

$$\mathbf{0} \longrightarrow H \xrightarrow{\phi_1} G \xrightarrow{\phi_2} K \longrightarrow \mathbf{0} \tag{51}$$

is exact if and only if the dual sequence

$$\mathbf{0} \longleftarrow \widehat{H} \xleftarrow{\widehat{\phi_1}} \widehat{G} \xleftarrow{\widehat{\phi_2}} \widehat{K} \longleftarrow \mathbf{0} \tag{52}$$

is exact. Furthermore, $\phi_1(H) < G$ is a closed subgroup if and only if $\widehat{\phi_2}(\widehat{K}) < \widehat{G}$ is closed.

A proof of the final statement about closed subgroups is found in [24].

Corollary 1. *Let H be a closed subgroup of G and $K = G/H$. Then $\widehat{K}$ is a closed subgroup of $\widehat{G}$ and $\widehat{H} \approx \widehat{G}/\widehat{K}$.*

Corollary 2. *If $\phi \in \mathrm{mono}(H, G)$ then $\widehat{\phi} \in \mathrm{epi}(\widehat{G}, \widehat{H})$, and if $\phi \in \mathrm{epi}(G, K)$ then $\widehat{\phi} \in \mathrm{mono}(\widehat{K}, \widehat{G})$.*

Definition 38 (Annihilator subgroup). *For a closed subgroup $H < G$ the closed subgroup $\widehat{G/H} < \widehat{G}$ is called* the annihilator subgroup *of $\widehat{G}$, denoted*

$$H^\perp := \widehat{G/H}.$$

The annihilator $H^\perp$ consists exactly of exactly those characters in $\hom(G, \mathbb{T}) \approx \widehat{G}$ (a.k.a. Fourier basis functions) which evaluate to 1 at all points $h \in H$:

Lemma 12. *Referring to diagrams (51)-(52) we have that*

$$(\xi, \phi_1(H))_G \equiv 1$$

if and only if $\xi \in \widehat{\phi_2}(\widehat{K})$.

Proof. For $x = \phi_1(h)$ and $\xi = \widehat{\phi_2}(k)$ we get

$$(\xi, x)_G = (\widehat{\phi_2}(k), \phi_1(h))_G = (k, \phi_2 \circ \phi_1(h))_K = (k, 0)_K = 1.$$

On the other hand, if $(\xi, \phi_1(h))_G = 1$ for all $h \in H$, then $(\widehat{\phi_1}(\xi), h)_H = 0)$, and hence $\widehat{\phi_1}(\xi) = 0$. Exactness implies the existence of a $k \in \widehat{K}$ such that $\widehat{\phi_2}(k) = \xi$. □

In terms of the bi-additive pairing $\langle \cdot, \cdot \rangle_G$, we have

$$\phi_1(H)^\perp = \left\{ \xi \in \widehat{G} \colon \langle \xi, \phi_1(H) \rangle_G \equiv 0 \in T \right\},$$

so the annihilator is the abelian group version of orthogonal complement in linear algebra. We also have $\left(H^\perp\right)^\perp = H$.

3.6 Lattices and sampling

An LCA K is called *compact* if

$$\mathrm{vol}(K) := \int_K dx < \infty.$$

An LCA H is called *discrete* if every point in H (and every subset of H) are open sets. We say that H is *continuous* if it is not discrete. The following result is proven in [24]. We have not discussed enough topology to reproduce the proof.

Theorem 12. *An LCA G is compact if and only if the dual group $\widehat{G}$ is discrete, and G is discrete if and only if the dual $\widehat{G}$ is compact.*

Example 28.

- $\mathbb{R} \leftrightarrow \widehat{\mathbb{R}} \approx \mathbb{R}$ (continuous, non-compact $\leftrightarrow$ continuous, non-compact).
- $\mathbb{Z} \leftrightarrow \widehat{\mathbb{Z}} \approx T$ (discrete, non-compact $\leftrightarrow$ compact, continuous).
- $\mathbb{Z}_n \leftrightarrow \widehat{\mathbb{Z}_n} \approx \mathbb{Z}_n$ (discrete, compact $\leftrightarrow$ discrete, compact).

Perhaps it is worth noting that we could choose $G = \mathbb{R}$ with a discrete topology. In this case $\widehat{G}$ is a compact space which is called the Bohr compactification of $\mathbb{R}$, after Harald Bohr, the brother of Niels Bohr, who studied the Fourier analysis of so-called *almost periodic functions*. A discussion of this topic is interesting, but brings us beyond the scope of these notes.

Definition 39 (Lattice). *A lattice is a* discrete *and closed subgroup $H < G$ such that G/H is compact.*

Recall that $H^\perp \approx \widehat{G/H}$ hence if H is a lattice, then the annihilator $H^\perp < \widehat{G}$ is also a lattice, called the *reciprocal lattice*. In the sequel we will study sampling theory as movements of functions between the domains in the diagram

$$\begin{array}{ccccccccc} \mathbf{0} & \longleftarrow & \widehat{H} & \xleftarrow{\widehat{\phi_1}} & \widehat{G} & \xleftarrow{\widehat{\phi_2}} & H^\perp & \longleftarrow & \mathbf{0} \\ & & \big| & & \big| & & \big| & & \\ \mathbf{0} & \longrightarrow & H & \xrightarrow{\phi_1} & G & \xrightarrow{\phi_2} & G/H & \longrightarrow & \mathbf{0}, \end{array} \tag{53}$$

where the vertical lines indicate dual pairs of groups, $\phi_1(H) < G$ and $\widehat{\phi_2}(H^\perp) < \widehat{G}$ are the reciprocal lattices and both rows are exact. In particular we study the relationship between Fourier transforms on G and on H, and we will even see that there is a relationship between functions on the spaces $\widehat{H}$ and G/H, which explains the Fast Fourier Transform. First, let us give a few concrete examples of this diagram.

Example 29 (Sound sampling). The classical setting of sampling of sound is the case where $G = \mathbb{R}$, $H = \mathbb{Z}$, $\phi_1(j) = j \cdot h$, where h is the sampling interval. We can set $G/H = h$ with $\phi_2(t) = t/h \bmod 1$, and $H^\perp = \mathbb{Z}$ with $\widehat{\phi_2}(k) = k/h$ and $\widehat{H} = T$ with $\widehat{\phi_1}(\xi) = \xi \cdot h$. The pairings are $\langle \xi, t \rangle_G = \xi \cdot t$, $\langle \xi, j \rangle_H = \xi \cdot j$ and $\langle k, t \rangle_{G/H} = k \cdot t$.

Example 30 (Multidimensional sampling of $\mathbb{R}^n$). Let $G = \mathbb{R}^n$ and $H = \mathbb{Z}^n$. A nonsingular matrix $A \in \mathbb{R}^{n \times n}$ defines $\phi_1(\mathbf{j}) = A\mathbf{j}$, where $G/H = T^n$ and $\phi_2(\mathbf{x}) = A^{-1}\mathbf{x}$. On the dual side we have $\widehat{H} = T^n$, $\widehat{G} = \mathbb{R}^n$ and $H^\perp = T^n$ with pairings $\langle \xi, \mathbf{x} \rangle_G = \xi^T \mathbf{x}$, $\langle \xi, \mathbf{j} \rangle_H = \xi^T \mathbf{j}$ and $\langle \mathbf{k}, \mathbf{x} \rangle_{G/H} = \mathbf{k}^T \mathbf{x}$. This yields the dual homomorphisms $\widehat{\phi_1}(\xi) = A^T \xi$ and $\widehat{\phi_2}(\mathbf{j}) = A^{-T}\mathbf{j}$. Note that a matrix of rank lower than n does *not* define a lattice in $\mathbb{R}^n$, since the quotient G/H in that case is non-compact.

Example 31 (Splitting for the FFT). Let $G = \mathbb{Z}_{mn}$ and $H = \mathbb{Z}_m$ with $\phi_1(j) = jn$. We have $G/H = \mathbb{Z}_n$ and $\phi_2(j) = j$. On the dual side we have $\widehat{H} = \mathbb{Z}_m$, $\widehat{G} = \mathbb{Z}_{mn}$, $H^\perp = \mathbb{Z}_n$ with pairings $\langle k, j\rangle_G = kj/mn$, $\langle k, j\rangle_H = kj/m$ and $\langle k, j\rangle_{G/H} = kj/n$. This yields the dual homomorphisms $\widehat{\phi_1}(k) = k$, since $\langle\widehat{\phi_1}(k), j\rangle_H = \langle k, j\rangle_H = kj/m = \langle k, nj\rangle_G = \langle k, \phi_1(j)\rangle_G$. Similarly we find $\widehat{\phi_2}(k) = km$.

Pullback and pushforward on lattices

For $\phi_1 \in \mathrm{mono}(H, G)$, where H is discrete, we call the operation $\phi_2^*\colon \mathbb{C}^G \to \mathbb{C}^H$ (down) *sampling*. For $\phi_2 \in \mathrm{epi}(G, K)$, where $\ker(\phi_2)$ is discrete, we call $\phi_{2*}\colon \mathcal{S}(G) \to \mathcal{S}(K)$ *periodisation*, given as

$$\phi_{2*}f(\phi_2(x)) = \sum_{k\in\ker(\phi_2)} f(x + k).$$

The name 'periodisation' reminds us that if we compute $g = {\phi_2}^* \circ \phi_{2*}f$, we obtain $g \in \mathcal{S}'(G)$ as a function periodic $g(x + k) = g(x)$ for all $k \in \ker(\phi_2)$. If we have a lattice $H < G$ and $K = G/H = \{g + H\}$, as the cosets, we have

$$\phi_{2*}f(g + H) = \sum_{h\in H} f(g + h),$$

which can be interpreted as a H-periodic function in $\mathbb{C}^G$.

Distributions can be moved in the opposite direction by ϕ_1 and ϕ_2. We call $\phi_{1*}\colon \mathcal{S}'(H) \to \mathcal{S}'(G)$ *up sampling*. This is given as

$$\phi_{1*}f = \sum_{h\in H} f(h)\delta_{\phi_1(h)},$$

where $\delta_{\phi_1(h)} = \delta(x - \phi_1(h))$ is the shifted δ-distribution. Up sampling of a discrete function on a lattice yields a set of point masses in the lattice points. The operation $\phi_{2*}\colon \mathcal{S}'(K) \to \mathcal{S}'(G)$ yields an H-periodic distribution on G.

Many important dual relationships can be derived from the following result, which is proven in many texts, see e.g. [22]:

Theorem 13 (Poisson summation formula). *Let $\phi_1 \in \mathrm{mono}(H, G)$ be a lattice with dual lattice $\widehat{\phi_2} \in \mathrm{mono}(H^\perp, G)$, as in (53). For $f \in \mathcal{S}(G)$ we have*

$$\sum_{h\in H} f(\phi_1(h)) = \frac{1}{C}\sum_{k\in H^\perp} \widehat{f}(\widehat{\phi_2}(k)),$$

where the constant $C = \mathrm{vol}(G/\phi_1(H))$ is the volume of the unit-cell of the lattice (if G is discrete C is the number of points in the unit-cell).

Choosing coset representatives

For many computational problems it is necessary to choose representative elements from each of the cosets in the quotient groups G/H and $\widehat{H} = \widehat{G}/H^\perp$. E.g. in sampling theory on a lattice $H < G = \mathbb{R}^n$, all characters in a coset $H^\perp + \xi \subset \widehat{G}$ *alias* on H (i.e. they evaluate to the same on H), but physical relevance is usually given to the character $\xi' \in H^\perp + \xi$ which is closest to 0 (the lowest frequency mode). Similarly, we often represent $K = G/H$ by picking a representative from each coset (e.g. $\mathbb{R}/\mathbb{Z}$ can be represented by $[0, 1) \subset \mathbb{R}$). The projection map $\phi \in \mathrm{epi}(G, K)$ assigns each coset to a unique element in K, and we need to decide on a right inverse of this map.

Definition 40 (Transversal of quotient $K = G/H$). *Given a quotient projection $\phi \in \mathrm{epi}(G, K)$, a function $\sigma\colon K \to G$ is called a* transversal *of ϕ if $\phi_1 \circ \sigma = \mathrm{Id}_K$ (this is often also called a* section *of the projection).*

Note that in general we cannot choose σ as a group homomorphism (only if $G = H \oplus K$), but it can be chosen as a continuous function. In many applications G has a natural norm (e.g. Euclidean distance on $\mathbb{R}^n$) and we can choose σ such that the coset representatives are as close to the origin as possible, i.e. such that $||\sigma(k)|| \leqslant ||\sigma(k) - h||$ for all $h \in H$.

Definition 41 (Voronoi transversal). *Let $G = \mathbb{R}^n$ or $G = \mathbb{T}^n$, and let $H < G$ be a lattice. The transversal $\sigma\colon G/H \to G$ such that $||\sigma(k)|| \leqslant ||\sigma(k) - h||$ for all $h \in H$ is called the* Voronoi transversal. *The image of the Voronoi transversal is a polyhedron around the origin in G, limited by hyperplanes orthogonal to the lines between the origin and the closest lattice points, and dividing these in the middle.*

Sampling and aliasing

Shannon's theory of sampling and reconstruction is a classical topic discussed in any textbook on signal processing, usually presented in the setting of Example 29. We review this in our setting of abelian groups, referring to the general lattice decomposition in (53). Periodisation and sampling are dual operations, in the sense that a function $f \in \mathbb{C}^G$ can be moved to $\mathbb{C}^{\widehat{H}}$ in two different ways, we can first sample f down to H and then compute the Fourier transform on H, or we can Fourier transform f on G and then periodise $\widehat{f}$ down to $\widehat{H}$. The result of these two operations is the same!

Theorem 14. *For a lattice $\phi_1 \in \mathrm{mono}(H, G)$*

$$\mathcal{F}_H \left[{\phi_1}^* f \right] = \widehat{\phi_1}_* \mathcal{F}_G[f] \quad \forall f \in \mathcal{S}(G), \tag{54}$$

where $\mathcal{F}_H[\cdot]$ and $\mathcal{F}_G[\cdot]$ denotes the Fourier transforms on H and G.

Proof. Pick an arbitrary $\xi \in \widehat{G}$ and let $\chi_\xi(x) := (\xi, x)_G$ be the corresponding character on G. Using the shift property of the Fourier transform and the Poisson summation formula, we find

$$\begin{aligned}\widehat{\phi_1}_* \mathcal{F}_G[f]\left(\widehat{\phi_1}(\xi)\right) &= \sum_{k\in\ker(\widehat{\phi_1})} \mathcal{F}_G[f](\xi+k) = \sum_{k\in\ker(\widehat{\phi_1})} \mathcal{F}_G\left[\chi_{-\xi} f\right](k) \\ &= \sum_{h\in H} (-\xi, \phi_1(h))_G f(\phi_1(h)) = \sum_{h\in H} (-\widehat{\phi_1}(\xi), h)_H f(\phi_1(h)) \\ &= \mathcal{F}_H\left[{\phi_1}^* f\right]\left(\widehat{\phi_1}(\xi)\right).\end{aligned}$$

□

Let $f \in \mathbb{C}^G$, $\widehat{f} \in \mathbb{C}^{\widehat{G}}$, $f_H := \phi_1^* f$ and $\widehat{f_H} := \mathcal{F}_H(f_H)$. Theorem 14 says:

$$\widehat{f_H}(\widehat{\phi_1}(\xi)) = \sum_{k\in\ker(\widehat{\phi_1})} \widehat{f}(\xi+k).$$

The *aliasing phenomenon* is the fact that Fourier components of $\widehat{f}$ which belong to the same coset of the reciprocal lattice add up to the same component of $\widehat{f_H}$. To reconstruct $\widehat{f}$ from $\widehat{f_H}$ we must decide on which of the aliasing components in the coset is the best representative for the coset. Reconstruction of f_H is based on choosing $\sigma\colon \widehat{H} \to \widehat{G}$ a transversal of $\widehat{\phi_1} \in \mathrm{epi}(\widehat{G}, \widehat{H})$. The standard choice if $G = \mathbb{R}^n$ or a $G = T^n$ is the Voronoi transversal, where σ picks points closest possible to 0 in the Euclidean norm. In the standard setting of Shannon sampling of Example 29 we choose $\sigma\colon T \to \mathbb{R}$ as $\sigma(x) = x/h$ for $x \in [0, \frac{1}{2})$ and $\sigma(x) = (x-1)/h$ for $x \in [\frac{1}{2}, 1)$, but other choices are also used in particular applications, where we we want to reconstruct particular parts of the spectrum (e.g. sideband coding). We will always assume that σ is chosen such that the closure of $\mathrm{im}(\sigma) \subset \widehat{G}$ is compact.

Definition 42 (Bandlimited function). *A function $f \in \mathbb{C}^G$ is* bandlimited *with respect to a transversal $\sigma\colon \widehat{H} \to \widehat{G}$ if $\mathrm{supp}(\widehat{f}) \subset \mathrm{im}(\sigma)$, where .$\mathrm{supp}(\widehat{f})$ is the support of $\widehat{f}$ i.e. the points where it takes non-zero values.*

For a given transversal σ we define a corresponding (low-pass) filter $\alpha_\sigma \in \mathbb{C}^{\widehat{G}}$ as the indicator function on the image of σ,

$$\alpha_\sigma(x) = \begin{cases} 1 \text{ for } x \in \mathrm{im}(\sigma) \\ 0 \text{ else.} \end{cases}$$

Thus, f is band-limited if and only if $\widehat{f} \cdot \alpha_\sigma = \widehat{f}$. Hence we have:

Lemma 13 (Shannon–Nyquist). *A band-limited function $f \in \mathcal{S}(G)$ can be reconstructed from its down-sample f_H as*

$$\widehat{f} = \alpha_\sigma \cdot \left(\widehat{\phi_1}^* \widehat{f_H}\right). \tag{55}$$

Polyhedral Dirichlet kernels.

We henceforth assume that $G = \mathbb{R}^n$ or $G = \mathbb{T}^n$ and the transversal $\sigma\colon \widehat{H} \to \widehat{G}$ is the Voronoi transversal, with an image being a polyhedron centered at $\mathbf{0} \in \widehat{G}$. The corresponding low-pass filter is 1 inside this polyhedron and on the boundary (in particular if $\widehat{G}$ is discrete) we give weight $1/n$ on all n points which tie-break on the distance criterion.

Definition 43 (Polyhedral Dirichlet kernel). *Let*

$$\Omega = \left\{ \xi \in \widehat{G} \;:\; ||\xi|| < ||\xi - k|| \text{ for all } k \in \widehat{\phi_2}(H^\perp) \setminus \mathbf{0} \right\}$$
$$\partial\Omega = \left\{ \xi \in \widehat{G} \;:\; ||\xi|| = ||\xi - k|| \text{ for some } k \in \widehat{\phi_2}(H^\perp) \setminus \mathbf{0} \right\}$$

We define the low-pass filter $\widehat{\mathcal{D}}_H \in \mathcal{S}'(\widehat{G})$ *as*

$$\widehat{\mathcal{D}}_H(\xi) = \begin{cases} 1 & \text{for } \xi \in \Omega \\ \frac{1}{N} & \text{for } \xi \in \partial\Omega, \\ 0 & \text{otherwise} \end{cases},$$

where $N = \#\{k \in \widehat{\phi_2}(H^\perp) \;:\; ||\xi|| = ||\xi - k||\}$. *The polyhedral Dirichlet kernel* $\mathcal{D}_H \in \mathbb{C}^\infty(G) \cap \mathcal{S}'(G)$ *is defined*[7] *as*

$$\mathcal{D}_H = \mathcal{F}_G^{-1}(\widehat{\mathcal{D}}_H).$$

Example 32. Continuing Example 29, where $G = \widehat{G} = \mathbb{R}$, and $\phi_1(j) = hj \in \mathrm{mono}(H, G)$, we find

$$D_H(x) = \int_{-\frac{1}{2h}}^{\frac{1}{2h}} e^{2\pi i \xi x} d\xi = \frac{\sin(\pi x/h)}{\pi x} = \frac{1}{h}\,\mathrm{sinc}(\pi x/h).$$

Example 33. For $G = T$, $H = \mathbb{Z}$ and $\phi_1(j) = j/N$, we have $\widehat{G} = \mathbb{Z}$, $H^\perp = \mathbb{Z}$ and $\widehat{\phi_2}(k) = Nk$, which gives

$$D_H(x) = \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} e^{2\pi i k x} = \frac{\sin(N\pi x)}{\sin(\pi x)} \quad \text{if } N \text{ is odd}$$
$$D_H(x) = \sum_{k=-\frac{N}{2}-1}^{\frac{N}{2}-1} e^{2\pi i k x} + \frac{1}{2}(e^{\pi i N x} + e^{-\pi i N x})$$
$$= \frac{\sin((N-1)\pi x)}{\sin(\pi x)} + \cos(N\pi x) \quad \text{if } N \text{ is even}$$

[7] Since $\widehat{\mathcal{D}}_H$ has compact support, its inverse Fourier transform is smooth.

We want to reproduce the classical convolutional formula for band-limited reconstruction of a sampled function in our setting. Let $f \in \mathcal{S}(G)$ and let the Shannon–Nyquist low-pass reconstruction[8] be given as

$$f \approx \mathcal{F}_G^{-1}\left[\widehat{\mathcal{D}}_H \cdot \left(\widehat{\phi_1}^* \widehat{f_H}\right)\right].$$

We have $\widehat{\phi_1}^* \widehat{f_H} \in \mathcal{S}'(\widehat{G})$. Recall Theorem 14, for Schwartz functions sampling and periodisation are dual operations. Tempered distributions belong to the dual space and move in the opposite direction, so we have in particular

$$\mathcal{F}_G^{-1}\left[\widehat{\phi_1}^* \widehat{f_H}\right] = \phi_{1*} f_H = \sum_{j\in H} f(\phi_1(j))\delta_{\phi_1(j)},$$

where $\delta_{\phi_1(j)}(x) = \delta(x - \phi_1(j))$. Since $\widehat{\mathcal{D}}_H$ has compact support, there is a convolutional formula for distributions leading to the reconstruction

$$\mathcal{F}_G^{-1}\left[\widehat{\mathcal{D}}_H \cdot \left(\widehat{\phi_1}^* \widehat{f_H}\right)\right] = \frac{1}{C}\mathcal{D}_H * \left(\sum_{j\in H} f(\phi_1(j))\delta_{\phi_1(j)}\right),$$

where the constant $C = D_H(0)$. This yields:

Theorem 15 (Shannon–Nyquist convolution formula). *The band-limited reconstruction of f from $f_H = \phi_1^* f = f\circ\phi_1$ can be computed as*

$$f(x) \approx \frac{1}{D_H(0)} \sum_{j\in H} D_H(x - \phi_1(j)) f(\phi_1(j)). \tag{56}$$

This is an exact reconstruction for band-limited f.

We see from band limited f that the formula is interpolating in the lattice points, and we conclude:

Lemma 14. *The normalised polyhedral Dirichlet kernel satisfies for $j \in H$*

$$\frac{D_H(\phi_1(j))}{D_H(0)} = \begin{cases} 1 \text{ for } j = \mathbf{0} \\ 0 \text{ else.} \end{cases}$$

The translates $S_{\phi_1(j)} D_H(x) = D_H(x - \phi_1(j))$ for all $j \in H$ form a complete set of Lagrangian basis functions for band limited trigonometric interpolation in the lattice points.

Analytical properties of polyhedral Dirichlet kernels are important for understanding sampling theory on general lattices. Detailed analysis of these functions is done in [27, 30]. In particular it is important that they in the case $G = T^n$ the interpolation operator has a Lebesque constant scaling like $\mathcal{O}(\log^n(N))$, where N is the number of sampling points in H.

[8] For band limited f this is exact, for other functions it is an interpolating formula.

3.7 The Fast Fourier Transform (FFT)

We return once more to the basic splitting diagram (53), but in this section we assume that all involved groups are finite. The aim is to compute the discrete Fourier transform (DFT) on G by expressing $\mathcal{F}_G$ in terms of the DFTs $\mathcal{F}_H$ and $\mathcal{F}_K$, where $K = G/H$. The simplest situation is when the diagram (53) splits, i.e. the case when $G = H \oplus K$. Then there exists homomorphisms $\sigma_1 \in \mathrm{epi}(G, H)$ and $\sigma_2 \in \mathrm{mono}(K, G)$ such that $\sigma_1 \circ \phi_1 = \mathrm{Id}_H$ and $\phi_2 \circ \sigma_2 = \mathrm{Id}_K$, and an isomorphism $\psi = \frac{\sigma_1}{\phi_2} \in \mathrm{iso}(G, H \oplus K)$. On $\mathbb{C}[H \oplus K]$ the DFT is $\mathcal{F}_H \oplus \mathcal{F}_K$, thus the whole DFT on G factorises as

$$\mathcal{F}_G = \widehat{\psi} \circ (\mathcal{F}_H \oplus \mathcal{F}_K) \circ \psi.$$

We can think of $\mathbb{C}[H \oplus K]$ as a 2D table. The isomorphisms ψ and $\widehat{\psi}$ are just permutations of the data, so the factorisation has three stages; first we use ψ_1 to arrange the data in a 2D table, then we use $\mathcal{F}_H$ on each colum, and $\mathcal{F}_K$ on each row of the table, and finally we collect the data back into $\widehat{G}$. The computation is facilitated by a software package for doing computations of homomorphisms between finite abelian groups. This factorisation of the DFT in the case where $G = H \oplus K$ is in FFT literature called *twiddle-free* FFT decomposition.

In the more general situation we have that $H \oplus K$ is not isomorphic to G. In this case we still try to use H and K as coordinates on G, but we cannot do this in a canonical way. We choose two transversals $\sigma_K \colon K \to G$ and $\sigma_H \colon \widehat{H} \to \widehat{G}$ and write

$$\begin{aligned} j &= \phi_1(m) + \sigma_K(\ell) \quad \text{for } m \in H,\ \ell \in K,\ j \in G \\ k &= \widehat{\phi_2}(p) + \sigma_H(n) \quad \text{for } p \in \widehat{K},\ n \in \widehat{H},\ k \in \widehat{G}. \end{aligned}$$

Using the properties we have derived for dual pairings we find (exercise!)

$$(k, j)_G = (n, m)_H (p, \ell)_K (\sigma_H(n), \sigma_K(\ell))_G.$$

The last factor $(\sigma_H(n), \sigma_K(\ell))_G$ is called a 'twiddle factor' and it reflects the fact that $G \neq H \oplus K$. We find that the Fourier transform on G factorises as

$$\mathcal{F}_G[f](k) = \mathcal{F}_G[f](\widehat{\phi_2}(p) + \sigma_H(n)) = \tag{57}$$
$$\sum_{\ell \in K} \left((\sigma_H(n), \sigma_K(\ell))_G \sum_{m \in \widehat{H}} (-n, m)_H f(\phi_1(m) + \sigma_k(\ell)) \right) (-p, \ell)_K.$$

Again, interpreting f as data in a 2D array, indexed by $m \in H$ and $\ell \in K$, we see that the DFT on G factorises in applying $\mathcal{F}_H$ on each column, then multiplying by the twiddle factors and finally $\mathcal{F}_K$ on the rows. This is the basis for the Cooley–Tukey algorithm, where this factorisation is done recursively to obtain the Fast Fourier Transform. The fact that this can be done with respect

to any subgroup $H < G$ is of theoretical importance, and practical importance if we want to design versions of FFTs taking account of symmetries in the data f, see [18].

However, this factorisation is not canonical, there is a choice of transversals and twiddle factors involved. So aesthetically this factorisation formula is not optimal. It is possible to obtain a canonical factorisation of a similar nature. For completeness, I would like to explain also this factorisation. This involves the lifting of f to a larger space than $H \oplus K$, called the Heisenberg group (originating from quantum mechanics). The last part of this section may be skipped without loss of continuity, but not without loss of insight!

Heisenberg groups and the Weil–Brezin map

More material on topics related to this section is found in [5, 29].

We can act upon $f \in \mathbb{C}[G]$ with a time-shift $S_x f(t) := f(t+x)$ and with a frequency shift $\chi_\xi f(t) := (\xi, t) f(t)$. These two operations are dual under the Fourier transform, but do not commute:

$$\widehat{S_x f}(\xi) = \chi_x \widehat{f}(\eta) \tag{58}$$

$$\widehat{\chi_\xi f}(\eta) = S_{-\xi} \widehat{f}(\eta) \tag{59}$$

$$\left(S_x \chi_\xi f\right)(t) = (\xi, x) \cdot \left(\chi_\xi S_x f\right)(t). \tag{60}$$

The full (non-commutative) group generated by time and frequency shifts on $\mathbb{C}[G]$ is called the *Heisenberg group* of G.

The Heisenberg group of $\mathbb{R}^n$ is commonly defined as the multiplicative group of matrices of the form

$$\begin{pmatrix} 1 & x^T & s \\ 0 & I_n & \xi \\ 0 & 0 & 1 \end{pmatrix},$$

where $\xi, x \in \mathbb{R}^n$, $s \in \mathbb{R}$. This group is isomorphic to the semidirect product $\mathbb{R}^n \times \mathbb{R}^n \rtimes \mathbb{R}$ where

$$(\xi', x', s') \cdot (\xi, x, s) = (\xi' + \xi, x' + x, s' + s + x'^T \xi).$$

We prefer to instead consider $\mathbb{R}^n \times \mathbb{R}^n \rtimes \mathbb{T}$ (where $\mathbb{T}$ is the multiplicative group consisting of $z \in \mathbb{C}$ such that $|z| = 1$) with product

$$(\xi', x', z') \cdot (\xi, x, z) = (\xi' + \xi, x' + x, z' z e^{2\pi i x'^T \xi}).$$

More generally:

Definition 44. *For an LCA G we define the Heisenberg group $\mathcal{H}_G = \widehat{G} \times G \rtimes \mathbb{T}$ with the semidirect product*

$$(\xi', x', z') \cdot (\xi, x, z) = (\xi' + \xi, x' + x, z' \cdot z \cdot (\xi, x')).$$

We define a *left action* $\mathcal{H}_G \times \mathbb{C}[G] \to \mathbb{C}[G]$ as follows

$$(\xi, x, z) \cdot f = z \cdot \chi_\xi S_x f. \tag{61}$$

To see that this defines a left action, we check that $(0, 0, 1) \cdot f = f$ and

$$(\xi', x', z') \cdot ((\xi, x, z) \cdot f) = \big((\xi', x', z') \cdot (\xi, x, z)\big) \cdot f.$$

Lemma 15. *Let $\mathcal{H}_G = \widehat{G} \times G \rtimes \mathbb{T}$ and $\mathcal{H}_{\widehat{G}} = G \times \widehat{G} \rtimes \mathbb{T}$ act upon $f \in \mathbb{C}G$ and $\widehat{f} \in \mathbb{C}\widehat{G}$ as in (61). Then*

$$\mathcal{F}((\xi, x, z) \cdot f) = z \cdot (-\xi, x) \cdot \chi_x S_{-\xi} \widehat{f} = (x, -\xi, z \cdot (-\xi, x)) \cdot \widehat{f}$$

Proof. This follows from (58)–(60).

We will henceforth assume that H, G and K form a short exact sequence as in (53), with H discrete and $K = G/H$ compact.

Definition 45 (Weil–Brezin map). *The* Weil–Brezin map $\mathcal{W}_G^H$ *is defined for $f \in \mathbb{C}[G]$ and $(\xi, x, z) \in \mathcal{H}_G$ as*

$$\mathcal{W}_G^H f(\xi, x, z) = \sum_{j \in H} ((\xi, x, z) \cdot f)_H (j),$$

where $f_H := f \circ \phi_1$ denotes downsampling along $\phi_1 \in \mathrm{mono}(H, G)$.

A direct computation shows that the Weil–Brezin map satisfies the following symmetries for all $(h', h, 1) \in H^\perp \times H \times 1 \subset \mathcal{H}_G$ and all $z \in \mathbb{T}$:

$$\mathcal{W}_G^H f\left((h', h, 1) \cdot (\xi, x, s)\right) = \mathcal{W}_G^H f(\xi, x, s) \tag{62}$$

$$\mathcal{W}_G^H f(\xi, x, z) = z \cdot \mathcal{W}_G^H f(\xi, x, 1). \tag{63}$$

Lemma 16. *$\Gamma = H^\perp \times H \times 1$ is a subgroup of $\mathcal{H}_G$. It is not a normal subgroup, so we cannot form the quotient group. However, as a manifold the set of right cosets is*

$$\Gamma \backslash \mathcal{H}_G = \widehat{H} \times K \times \mathbb{T}.$$

The Heisenberg group has a right and left invariant volume measure given by the direct product of the invariant measures of $\widehat{G}$, G and $\mathbb{T}$. Thus we can define the Hilbert spaces $L^2(\mathcal{H}_G^H)$ and $L^2(\widehat{H} \times K \times \mathbb{T})$. By Fourier decomposition in the last variable (z-transform), $L^2(\widehat{H} \times K \times \mathbb{T})$ splits into an orthogonal sum of subspaces $\mathcal{V}_k$ for $k \in \mathbb{Z}$, consisting of those $g \in L^2(\widehat{H} \times K \times \mathbb{T})$ such that

$$g(\xi, x, z) = z^k g(\xi, x, 1) \quad \text{for all } z = e^{2\pi i \theta}.$$

It can be verified that $\mathcal{W}_G^H$ is unitary with respect to the L^2 inner product. Together with (62)–(63) this implies:

Lemma 17. *The Weil–Brezin map is a unitary transform*

$$\mathcal{W}_G^H \colon L^2(G) \to \mathcal{V}_1 \subset L^2(\widehat{H} \times K \times \mathbb{T}).$$

Note that the Weil–Bezin map on $\widehat{G}$, with respect to the reciprocal lattice $H^\perp$, is

$$\mathcal{W}_{\widehat{G}}^{H^\perp} \colon L^2(G) \to \mathcal{V}_1 \subset L^2(K \times \widehat{H} \times \mathbb{T}).$$

The Poisson summation formula (Theorem 13) together with Lemma 15 implies that these two maps are related via

$$\mathcal{W}_G^H f(\xi, x, z) = \mathcal{W}_{\widehat{G}}^{H^\perp} \widehat{f}(x, -\xi, z \cdot (\xi, x)).$$

Defining the unitary map $J \colon L^2 \subset L^2(\widehat{H} \times K \rtimes \mathbb{T}) \to L^2(K \times \widehat{H} \times \mathbb{T})$ as

$$Jf(x, -\xi, z \cdot (\xi, x)) = f(\xi, x, z), \tag{64}$$

we obtain the following fundamental theorem.

Theorem 16 (Weil–Brezin factorization). *Given an LCA G and a lattice $H < G$. The Fourier transform on G factorizes in a product of three unitary maps*

$$\mathcal{F}_G = \left(\mathcal{W}_{\widehat{G}}^{H^\perp}\right)^{-1} \circ J \circ \mathcal{W}_G^H. \tag{65}$$

The Zak transform.

We want to explain (57) in terms of the Weil–Brezin map. Given a lattice $H < G$ and transversals $\sigma \colon K \to G$ and $\widehat{\sigma} \colon \widehat{H} \to \widehat{G}$. The *Zak transform* is defined as

$$\mathcal{Z}_G^H f(\xi, x) := \mathcal{W}_G^H f(\xi, x, 1) \quad \text{for } \xi \in \widehat{\sigma}(\widehat{H}),\ x \in \sigma(K). \tag{66}$$

The Zak transform can be computed as a collection of Fourier transforms on H of f shifted by x, for all $x \in \sigma(K)$. The definition of the Fourier transform yields:

$$\mathcal{Z}_G^H f(-\xi, x) = \mathcal{F}_H\left((S_x f)_H\right)(\widehat{\phi}_0(\xi)). \tag{67}$$

We see that the Zak transform is invertible when $\mathcal{Z}_G^H f(-\xi, x)$ is computed for all $\xi \in \widehat{\sigma}(\widehat{H})$ and all $x \in \sigma(K)$. Written in terms of the Zak transform, the Weil–Brezin factorization (65) becomes

$$\mathcal{Z}_{\widehat{G}}^{H^\perp} \widehat{f}(x, \xi) = (\xi, x)_G \mathcal{Z}_G^H f(-\xi, x). \tag{68}$$

This is essentially the same formula as (57), where $(\xi, x)_G$ is the *twiddle factor*.

Due to the symmetries (62)–(63), the Weil–Brezin map is trivially recovered from the Zak transform. The Zak transform is the practical way of computing the Weil–Brezin map and its inverse. However, since the invertible Zak

transform cannot be defined canonically, independently of the transversals σ and $\widehat{\sigma}$, the Weil–Brezin formulation is more fundamental.

We end our discussion of the FFT at this point with the remark that the DFT on a finite abelian group G can always be computed with a complexity of $\mathcal{O}(|G| \log |G|)$ floating point operations, although for some cases such as $\mathbb{Z}_p$, where p is a large prime we must use other techniques than those discussed here. The underlying principles for computing the DFT are based on group theory. The details of state of the art FFT-software is involved, but for most applications in computational science it is sufficient to know that excellent FFT libraries exists. The practical question is then how the Fourier transforms on more general LCAs can be related to the finite groups.

3.8 Lattice rules

Lattice rules are numerical algorithms for computing in continuous groups by sampling in regular lattices and reducing to computations on finite groups. Most commonly the term refers to numerical integration of multivariate periodic functions in $\mathbb{C}^{T^n}$. The solution of PDEs by lattice sampling rules is discussed in [20]. In the present general setting, we discuss lattice rules for functions on $G = \mathbb{R}^n$, in which case we must introduce sampling lattices both in G and in $\widehat{G}$ to obtain a finite group where computations reduce to the FFT.

For general functions $f \in \mathbb{C}^G$, the error between the Fourier transform of the true and the sampled function follows from Theorem 14

$$\mathcal{F}_H(f_H)(\widehat{\phi}_1(\xi)) - \mathcal{F}_G(f)(\xi) = \sum_{k \in \widehat{\phi}_2(\widehat{K}) \setminus \{0\}} \mathcal{F}_G(f)(k + \xi).$$

The game of Lattice rules is, given f with specific properties, to find a lattice $H < G$ such that the error is minimised. We first assume (as is commonly done in the lattice-rule literature) that the original domain is periodic $G = T^n$. Lattice rules are designed such that the nonzero points in $H^\perp$ neighbouring 0 are pushed as far out as possible with respect to a given norm, depending on the properties of f. If f is spherically symmetric, H should be chosen as a *densest lattice packing* (with respect to the 2-norm) [8], e.g. hexagonal lattice in $\mathbb{R}^2$ and face centred cubic packing in $\mathbb{R}^3$ (as the orange farmers know well). In dimensions up to 8, these are given by certain root lattices [21]. The savings, compared to standard tensor product lattices, are given by the factors 1.15, 1.4, 2.0, 2.8, 4.6, 8.0 and 16.0 in dimensions $n = 2, 3, \ldots, 8$. This is important, but not dramatic, e.g. a camera with 8.7 megapixels arranged in a hexagonal lattice has approximately the same sampling error as a 10 megapixel camera with a standard square pixel distribution. However, these alternative lattices have other attractive features, such as larger spatial symmetry groups, yielding more isotropic discretizations. A hexagonal lattice picture can be rotated more uniformly than a square lattice picture.

Dramatic savings can be obtained for functions belonging to the *Korobov spaces*. This is a common assumption in much work on high dimensional approximation theory. Korobov functions are functions whose Fourier transforms have energy concentrated along the axis directions in $\widehat{G}$, the so-called hyperbolic cross energy distribution. Whereas the tensor product lattice with $2d$ points in each direction contains $(2d)^n$ lattice points in T^n, the optimal lattice with respect to the Korobov norm contains only $\mathcal{O}(2^n d(\log(d))^{n-1})$ points, removing exponential dependence on d.

The group theoretical understanding of lattice rules makes software implementation very clean and straightforward. In [20], numerical experiments are reported on lattice rules for FFT-based spectral methods for PDEs. Note that whereas the choice of transversal $\widehat{\sigma}\colon \widehat{H} \to \widehat{G}$ is irrelevant for lattice integration rules, it is essential for pseudospectral derivation. The Laplacian $\nabla^2 f$ is computed on $\widehat{G}$ as $\widehat{f}(\xi) \mapsto c|\xi|^2 \widehat{f}(\xi)$, whereas the corresponding computation on $\widehat{H}$ must be done as $\mathcal{F}_H(f_H)(\eta) \mapsto c|\widehat{\sigma}(\eta)|^2 \mathcal{F}_H(f_H)(\eta)$ for $\eta \in \widehat{H}$, and we must choose the Voronoi transversal to minimise aliasing errors.

Computational Fourier analysis on $\mathbb{R}^d$

A topic which in our opinion has not been fully addressed in the Lattice-rule literature is the computation of Fourier transforms on the non-compact continuous groups $\mathbb{R}^d$. The problem here is that there are no homomorphisms of a finite abelian group into $\mathbb{R}^d$, since any lattice in $\mathbb{R}^d$ must be non-compact. In order to move the computation to a finite group $\mathbb{Z}_{\mathbf{n}}$, $\mathbf{n} = (n_1, \cdots, n_k)$, we must use *two* homomorphisms

$$\mathbb{Z}_{\mathbf{n}} \xhookrightarrow{\phi_s} \mathbb{T}^d \xleftarrow{\phi_p} \mathbb{R}^d, \tag{69}$$

where ϕ_s is a sampling lattice and ϕ_p defines periodisation of a function with respect to the *periodisation lattice* $\ker(\phi_p) < \mathbb{R}^d$. A function $f \in S(\mathbb{R}^n)$ can be mapped down to a finite $f_{\mathbf{n}} \in \mathcal{S}(\mathbb{Z}_{\mathbf{n}})$ as

$$f_{\mathbf{n}} = \left({\phi_s}^* \circ \phi_{p*}\right)(f). \tag{70}$$

Since $\widehat{\mathbb{Z}_{\mathbf{n}}} = \mathbb{Z}_{\mathbf{n}}$, $\widehat{T^d} = \mathbb{Z}^d$ and $\widehat{\mathbb{R}^d} = \mathbb{R}^d$, the dual of (69) is

$$\mathbb{Z}_{\mathbf{n}} \xleftarrow{\widehat{\phi_s}} \mathbb{Z}^d \xhookrightarrow{\widehat{\phi_p}} \mathbb{R}^d. \tag{71}$$

The relationship between the discrete and the continuous Fourier transform follows from Theorem 14 (applied twice)

$$\widehat{f_{\mathbf{n}}} = \left(\widehat{\phi_s}_* \circ \widehat{\phi_p}^*\right)(\widehat{f}). \tag{72}$$

Note that sampling in the primary domain becomes periodisation in the Fourier domain and vice versa. The reconstruction of $\widehat{f}$ from $\widehat{f_{\mathbf{n}}}$ can be done as

a 2-stage process involving band limited approximation in the Voronoi domain of the dual sampling lattice $\ker(\widehat{\phi_s}) < \mathbb{Z}^d$ and a space limited approximation in the Voronoi domain of the periodisation lattice $\ker(\phi_p) < \mathbb{R}^d$, using the Shannon–Nyquist reconstruction formula twice. In this process no function (except $f = 0$) is perfectly reconstructed, since $f = 0$ is the only function which is both band limited and with compact support.

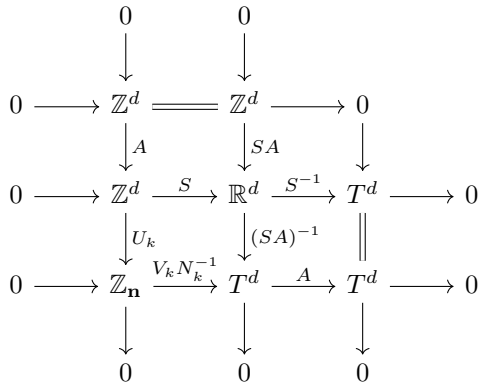
We want to understand the mappings involved in this two-stage sampling process in more detail. We have two lattices, the periodisation lattice $\ker(\phi_p) < \mathbb{R}^n$ and sampling lattice $\phi_s(\mathbb{Z}_\mathbf{n}) < T^d$. Considering the sampling lattice lifted to $\mathbb{R}^d$ as the subgroup $\phi_p^{-1}(\phi_s(\mathbb{Z}_\mathbf{n})) < \mathbb{R}^d$, we realise that the periodisation lattice is a sub-lattice of the sampling lattice in $\mathbb{R}^d$. The two lattices are described by two matrices $S \in \mathbb{R}^{d\times d}$ and $A \in \mathbb{Z}^{d\times d}$ with non-vanishing determinants. These define a sampling lattice $S\colon \mathbb{Z}^d \hookrightarrow \mathbb{R}^d$ and a periodisation sub lattice defined by $SA\colon \mathbb{Z}^d \hookrightarrow \mathbb{R}^d$. Consider the following commutative diagram where all rows and all columns are exact and where $\mathbf{n} = (n_1, \ldots, n_k)$ such that $n_i | n_{i+1}$ for $i = 1, \ldots, k-1$. The second row describes the sampling lattice and the second column the periodisation lattice in $\mathbb{R}^n$.

$$\begin{array}{ccccccccc}
 & & 0 & & 0 & & & & \\
 & & \downarrow & & \downarrow & & & & \\
0 & \longrightarrow & \mathbb{Z}^d & = & \mathbb{Z}^d & \longrightarrow & 0 & & \\
 & & \downarrow{\scriptstyle A} & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & \mathbb{Z}^d & \xrightarrow{S} & \mathbb{R}^d & \longrightarrow & T^d & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \| & & \\
0 & \longrightarrow & \mathbb{Z}_\mathbf{n} & \longrightarrow & T^d & \longrightarrow & T^d & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

Given A and S as above, there is a unique (up to isomorphisms) way of completing this diagram such that all rows and columns are exact. We will explicitly compute all the arrows. Let the Smith normal form of A be

$$A = UNV,$$

where $U \in \mathbb{Z}^{d\times d}$ and $V \in \mathbb{Z}^{d\times d}$ are unimodular and $N \in \mathbb{Z}^{d\times d}$ is diagonal, where the diagonal $n_i = N_{i,i}$ contains positive integers such that $n_i | n_{i+1}$ for all i. Since A has nonzero determinant, none of the n_i are zero, but the first ones could be 1. Let k denote the number of n_i such that $n_i > 1$, and let $\mathbf{n}$ be the vector containing these last k diagonal elements, defining the FAG $\mathbb{Z}_\mathbf{n}$. Let $U_k \in \mathbb{Z}^{k\times d}$ denote the last k rows of U^{-1} and let $V_k \in \mathbb{Z}^{d\times k}$ denote the last k columns of V^{-1}. Finally, let $N_k = \operatorname{diag}(\mathbf{n}) \in \mathbb{Z}^{k\times k}$ and $N_k^{-1} \in \mathbb{R}^{k\times k}$. Then the diagram is completed as follows.

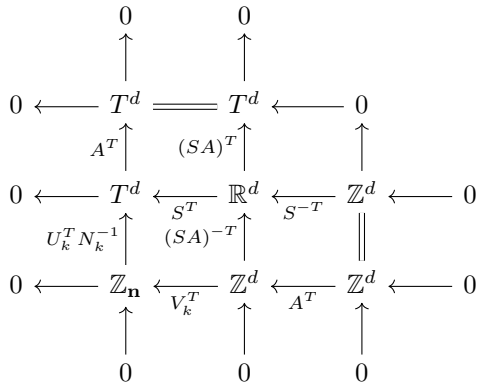


It is straightforward to check that the first two rows and last two columns are short and exact sequences. The first column is a short exact sequence because $U_k A \bmod \mathbf{n} = 0$ and U_k has maximal rank. It is straightforward to check the commutativity of the NW, NE and SE squares. The SW square commutes because $A^{-1} - V_k N_k^{-1} U_k$ is an integer matrix. The last row is exact by the 9-lemma of homological algebra [17].

This defines the periodisation lattice $SA\colon \mathbb{Z}^d \hookrightarrow \mathbb{R}^d$ with quotient mapping $(SA)^{-1}\colon \mathbb{R}^d \twoheadrightarrow T^d$. A function $f \in \mathcal{S}(\mathbb{R}^d)$ can be approximated by a function $f_{\mathbf{n}} \in \mathcal{S}(\mathbb{Z}_{\mathbf{n}})$ as

$$f_{\mathbf{n}} = (U_k)_* S^* f = (V_k N_k^{-1})^* (SA)_*^{-1} f.$$

We say that $V_k N_k^{-1}$ is a *rank k lattice rule*. Dualising the above sampling-perioisation diagram yields:



If we flip the diagram around the SW-NE diagonal, we see that this is nearly identical to the original sampling–periodisation diagram. But here the sampling lattice is given by $\widehat{SA} = (SA)^{-T}\colon \mathbb{Z}^d \to \mathbb{R}^d$, while the periodisation lattice is given by $\widehat{S} = S^{-T}\colon \mathbb{Z}^d \to \mathbb{R}^d$. Thus, the reciprocal of the primal sampling lattice is the dual periodisation lattice and the reciprocal of the primal periodisation lattice is the dual sampling lattice.

Complete symmetry between primal and dual spaces is obtained by letting the primal sampling lattice be obtained by down-scaling the reciprocal of the

primal periodisation lattice in $\mathbb{R}^d$. Specifically, given a non-singular matrix $L \in \mathbb{R}^{d\times d}$ and an integer m, we let the primal sampling lattice be $S = \frac{1}{m}L$ and $A = mL^{-1}L^{-T}$. The primal periodisation lattice is $SA = L^{-T}$. The dual sampling lattice is $(SA)^{-T} = L = mS$ and the dual periodisation lattice $S^{-T} = mSA$.

Eigenfunctions of the continuous and discrete Fourier transforms

We end this section with a brief remark showing a beautiful and perhaps unexpected property of discretising $\mathbb{R}^n$ in a completely symmetric fashion as discussed above. To understand the analytic properties of the discrete and continuous Fourier transforms, it is of importance to know the eigenfunctions of the Fourier operator. The eigenvectors of discrete Fourier transforms is a topic of interest both in pure and applied mathematics [5]. Both on $\mathbb{R}^n$ and on $\mathbb{Z}_{\mathbf{n}}$, the Fourier transform is a linear operator from a space to itself, so we can talk about eigenfunctions η with the property that $\widehat{\eta} = \lambda \cdot \eta$ for some $\lambda \in \mathbb{C}$. The Fourier transform satisfies $\mathcal{F}^4 = I$, hence we have that $\lambda \in \{1, i, -1, -i\}$. Since there are only 4 invariant subspaces, the eigenfunctions are not uniquely defined. However, for $\mathcal{F}_{\mathbb{R}}$ a particular complete set of eigenfunctions is known, The most famous eigenfunction is an appropriate scaling of the Gaussian $\exp(-x^2/\sigma^2)$, which is the ground state of the quantum harmonic oscillator. The set of all the eigenstates of the quantum harmonic oscillator form a complete set of eigenfunctions of $\mathcal{F}_{\mathbb{R}}$. These are of the form of a Hermite polynomial times a Gaussian. Similarly, we can take the eigenstates of the d-dimensional quantum harmonic oscillator as a basis for the eigen spaces of the d-dimensional Fourier transform $\mathcal{F}^d_{\mathbb{R}}$.

Theorem 17. *If we have a complete symmetry between the primal and dual sampling and periodisation lattices on $\mathbb{R}^d$, then the discretisation of eigenfunction η of the continuous Fourier transform $\mathcal{F}_{\mathbb{R}^d}$*

$$\eta_{\mathbf{n}} = (U_{k*} \circ S^*)\eta$$

is an eigenfunction of the discrete Fourier transform $\mathcal{F}_{\mathbb{Z}_{\mathbf{n}}}$.

Proof. In the symmetric situation we have that the primal and dual discretisations are the same

$$\begin{aligned} \eta_{\mathbf{n}} &= (U_{k*} \circ S^*)\eta \\ \mathcal{F}_{\mathbb{Z}_{\mathbf{n}}}[\eta_{\mathbf{n}}] &= (U_{k*} \circ S^*)\mathcal{F}_{\mathbb{R}^d}[\eta], \end{aligned}$$

hence

$$\mathcal{F}_{\mathbb{Z}_{\mathbf{n}}}[\eta_{\mathbf{n}}] = (U_{k*} \circ S^*)\mathcal{F}_{\mathbb{R}^d}[\eta] = \lambda \cdot (U_{k*} \circ S^*)\eta = \lambda \cdot \eta_{\mathbf{n}}.$$

It might be an interesting research topic to investigate computational reconstruction algorithms which aims at being accurate for down sampled eigenstates of the quantum harmonic oscillator.

3.9 Boundaries, mirrors and kaleidoscopes

The Fourier theory is a perfect tool for computing with shift invariant linear operators. This is, however, a very ideal world. Practical computational problems usually involve operators with coefficients varying over space and problems with boundaries. What can we do with such problems? A crucial technique is preconditioning, where real-life computational problems are approximated by problems in the ideal world. E.g. operators with variable coefficients can be approximated by operators where the coefficients are averaged over the domain. This is discussed in Section 3.11. For boundaries, it is worth knowing that *certain* special boundaries can be treated exactly within Fourier theory. The classification of such domains is of importance to computational science. One technique, which we will not pursue here, is based on separation of variables for PDEs. This has lead to fast solvers for Poisson problems on domains such as rectangles and circles.

We will instead discuss boundary problems which can be solved by Fourier techniques using mirrors on the boundaries of the domain, leading to fast computational techniques for *certain* triangles (2D) and simplexes in higher dimensions. These techniques also relates to beautiful topics in pure mathematics, such as the classification of reflection groups (kaleidoscopes), and the classification of semi-simple Lie groups.

We provide the basic idea with a well-known example derived in an unusual manner.

Example 34. Find eigenvectors and eigenvalues of the discrete 1-D Laplacian with Dirichlet conditions, the $(n-1)\times(n-1)$ matrix

$$A = \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{pmatrix}.$$

Apart from the boundaries, this is a convolution on $\mathbb{C}^{\mathbb{Z}}$ with $\mathbf{a} \in \mathbb{C}^{\mathbb{Z}}$ given as $\mathbf{a}(0) = 2$, $\mathbf{a}(1) = -1$, $\mathbf{a}(-1) = -1$. Now we fix the boundaries by setting up mirrors on the edges $j = 0$ and $j = n$. Since we have Dirichlet conditions, we set up two mirrors which act on a function by flipping it around a boundary point and changing the sign, i.e. we have two reflections acting on $f \in \mathbb{C}^{\mathbb{Z}}$ as

$$\begin{aligned} F_1 f(j) &= -f(-j) \\ F_2 f(j) &= -f(2n-j). \end{aligned}$$

Note that the convolution operator $\mathbf{a}$ commutes with the reflections, $F_i(\mathbf{a} * f) = \mathbf{a} * (F_i f)$ for every $f \in \mathbb{C}^{\mathbb{Z}}$. We seek a subspace of $\mathbb{C}^{\mathbb{Z}}$ of functions invariant under the action of F_i, the linear subspace $V \subset \mathbb{C}^{\mathbb{Z}}$ such that for

all $f_s \in V$ we have $F_1 f = F_2 f = f$. Note that $F_2 \circ F_1 = S_{2n}$, the shift operator $S_{2n} f(j) = f(j - 2n)$. Hence we see that $V \subset \mathbb{C}^{\mathbb{Z}/2n\mathbb{Z}} = \mathbb{C}[\mathbb{Z}_{2n}]$, and furthermore

$$V = \{f \in \mathbb{C}[\mathbb{Z}_{2n}] \,:\, F_1 f = f\}.$$

The other symmetry $F_2 f = f$ follows because of 2n-periodicity.
Let $\Pi = \frac{1}{2}(I + F_1)\colon \mathbb{C}[\mathbb{Z}_{2n}] \to V$ be the projection onto the symmetric subspace. V contains all $2n$ periodic functions of the form

$$(\ldots, -f_2, -f_1, 0, f_1, f_2, \ldots, f_{n-1}, 0, -f_{n-1}, \ldots).$$

Let $\Omega = \{1, 2, \ldots, n-1\}$ be the fundamental domain of the symmetric subspace, i.e. any $f \in V$ can be reconstructed from its restriction $f|_\Omega$. Note that A acts on the fundamental domain just like the convolution with $\mathbf{a}$ on V,

$$(\alpha * f)|_\Omega = A \cdot f|_\Omega \quad \text{for all } f \in V.$$

Since the convolution $\mathbf{a}*$ commutes with F_i, it also commutes with the projection Π and hence for any eigenvector $\alpha * \eta = \lambda\eta$ we have

$$\alpha * (\Pi\eta) = \Pi(\alpha * \eta) = \lambda \Pi\eta,$$

so $\Pi\eta$ is also an eigenvector with the same eigenvalue. Hence

$$A \cdot \Pi\eta|_\Omega = (\alpha * \Pi\eta)|_\Omega = \lambda\, \Pi\eta|_\Omega\,,$$

so $\Pi\eta|_\Omega$ is an eigenvector of A. On $\mathbb{C}[\mathbb{Z}_{2n}]$ the eigenvectors of the convolution $\mathbf{a}*$ are the characters $\chi_k = \exp(\pi i j k/n)$, which yelds the eigenvector for A:

$$(\Pi\chi_k)(j) = \frac{1}{2}(e^{\pi i j k/n} + e^{-\pi i j k/n}) = \cos(\pi j k/n).$$

The corresponding eigenvalue is $\lambda_k = \hat{\mathbf{a}}(k) = 2 - 2\cos(\pi k/n)$.

The trick in this example works for the following reasons:

- The matrix A acts like a convolution $\mathbf{a}*$ inside of the domain Ω.
- On the boundary of Ω, the boundary conditions can be satisfied by reflection operators, which commute with the convolution.
- A acts on the domain Ω as the convolution $\mathbf{a}*$ acts on the symmetrized extended functions in V.
- The reflections generate translations, so that we can obtain the eigenfunctions of A from the eigenfunctions on the larger periodic domain.

By similar techniques, we can find eigenvectors for a number of different tri-diagonal matrices with combinations of Dirichlet or Neumann conditions at lattice points or between lattice points. More generally, we can ask: On which domains in $\mathbb{R}^d$ can we define boundary conditions by similar mirror techniques and employ symmetric versions of Fourier expansions as a computational tool? To answer this question, we ask first what are the polytopes

in $\mathbb{R}^d$ with the property that if we reflect the domain at its boundaries, we eventually generate finite translations in all d directions? When these domains are understood, we can by reflection techniques find the eigenfunctions of the Laplacian ∇^2 on these domains, with various combinations of Dirichlet and Neumann conditions, and finally we can seek lattices which are invariant under the boundary reflections to obtain discretisations which can be computed by FFT techniques.

We will not go into the detail of this topic in these lectures. The interested reader is referred to [7]. We summarise the main results. In $\mathbb{R}^2$ the only domains with the property that reflections about the boundaries generate (finite) translations in both directions are:

- Any rectangle.
- The equilateral triangle. Reflections of the triangle produces six rotated and reflected triangles inside a hexagon, and continued reflections produce a tiling of $\mathbb{R}^2$ where this hexagon is shifted in two different directions.
- The $45^o - 45^o - 90^o$ triangle. Reflections of the triangle produces eight rotated/reflected triangles inside a square, and continued reflections produce a tiling of $\mathbb{R}^2$ where this square is shifted in two different directions.
- The $30^o - 60^o - 90^o$ triangle. Reflections of the triangle produces 12 rotated and reflected triangles inside a hexagon, and continued reflections produce a tiling of $\mathbb{R}^2$ where this hexagon is shifted in two different directions.

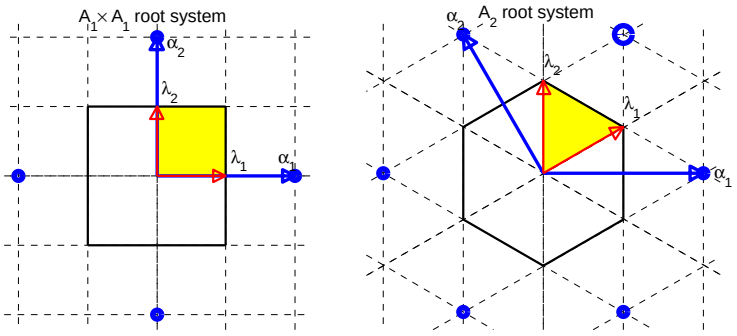


Fig. 1. Reducible root system $A_1 \times A_1$ and irreducible system A_2

The classification of such 'kaleidoscopic mirror systems', called *'root systems'*, in all dimensions was completed in the 1890s by Wilhelm Killing and Elie Cartan. They needed this to classify all semisimple Lie groups. There are some domains which decompose in orthogonal directions, such as a rectangle,

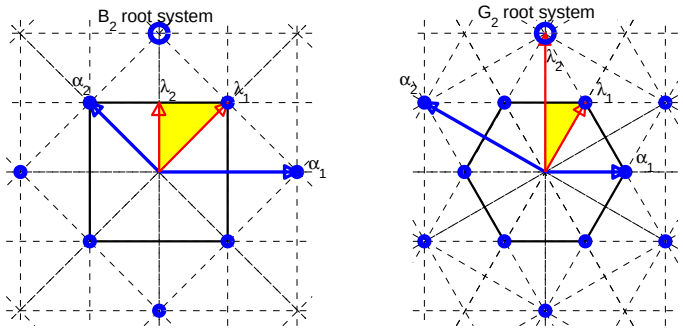


Fig. 2. The irreducible root systems B_2 and G_2

which decomposes in two mirrors in the horizontal direction and two mirrors in the vertical direction, and there are some irreducible domains which cannot be decomposed into orthogonal directions, such as the three triangles in $\mathbb{R}^2$ listed above. The fundamental domains for the irreducible cases are always special simplexes (triangles and their higher dimensional analogues). The fundamental domains for the reducible cases are cartesian products of irreducible domains, such as a rectangle in $2D$, which ls the cartesian product of two orthogonal lines (1-D simplexes) and an equilateral prism in $3D$, which is the cartesian product of an equilateral triangle with a line. We summarise the theory:

- There are four infinite families of irreducible root systems, $A_n, n > 0$, $B_n, n > 1$, $C_n, n > 2$ and $D_n, n > 3$. Here n is the dimension of the space. A_2 is the equilateral triangle and B_2 is the $45^o - 45^o - 90^o$ triangle.
- There are five exceptional root systems which only exist in particular dimensions, E_6, E_7, E_8, F_4 and G_2, where G_2 is the $30^o - 60^o - 90^o$ triangle.
- For each root system there corresponds two particularly nice families of lattices which are preserved under the reflection symmetries. These are called the roots lattice and the weights lattice and are straightforward to compute.

We refer the readers to [13, 25] for a discussion of these topics. We also mention that for each root system there corresponds a system of multivariate Chebyshev polynomials which has a number of remarkable approximation properties.

3.10 Cyclic reduction

The topic of this section is discussed in more detail in [19]. Cyclic reduction is a classical computational technique which has been employed in the design of fast Poisson solvers [28], among other things. As a simple example consider the solution of a linear system on $\mathbb{Z}_6$ with coefficient matrix A is given as

the convolution with $\mathbf{a}$ where $a(0) = 2$, $a(1) = a(-1) = -1$ and the rest is 0. Let us pre-multiply A with B being convolution with $\mathbf{b}$ where $b(0) = 2$, $b(1) = b(-1) = 1$ and the rest is 0. After re-arranging the nodes in even-odd order $0, 2, 4, 1, 3, 5$, we have the following matrix presentation of this 'odd-even' cyclic reduction step:

$$BA = \left(\begin{array}{cc|cc} 2 & & 1 & 1 \\ & 2 & 1 & 1 \\ \hline 1 & 1 & 2 & \\ 1 & 1 & & 2 \end{array}\right) \left(\begin{array}{cc|cc} 2 & & -1 & -1 \\ & 2 & -1 & -1 \\ \hline -1 & -1 & 2 & \\ -1 & -1 & & 2 \end{array}\right)$$

$$= \left(\begin{array}{ccc|ccc} 2 & -1 & -1 & & & \\ -1 & 2 & -1 & & & \\ -1 & -1 & 2 & & & \\ \hline & & & 2 & -1 & -1 \\ & & & -1 & 2 & -1 \\ & & & -1 & -1 & 2 \end{array}\right).$$

So, we have decoupled odd and even nodes, and can continue with a problem of half the size. As a convolution of 'stencils', the reduction is

$$[1, 2, 1] * [-1, 2, -1] = [-1, 0, 2, 0, -1].$$

Remarks:

- If $n = 2^k$ we can apply the procedure recursively for solving $Ax = b$.
- For solving $Ax = b$ it is only necessary to do the reduction to the even points, and back substitute the solution afterwards.
- The procedure also works on those domains with boundaries, that can be represented in terms of symmetric functions as in Section 3.9. This includes 1D problems on an interval (Dirichlet or Neumann boundaries), rectangles in 2D, 7-point Laplacian stencil on a hexagonal lattice on an equilateral triangle with Dirichlet or Neumann boundaries, etc.
- In 2D and higher dimensions, classical (1-way) cyclic reduction schemes are unstable, and special caution must be exercised.

Our question is now how this can be generalised to the reduction of a convolutional operator $\mathbf{a} \in \mathbb{C}[G]$ to $\mathbf{b}*\mathbf{a}$ with support on an arbitrary subgroup $H < G$? The answer is a nice exercise applying the duality theory of the Fourier transform.

Theorem 18. *Given a convolutional operator $\mathbf{a} \in \mathbb{C}[G]$ a subgroup $H < G$. Let $\chi_k(j) = (k, j)_G$ and let $\mathbf{b} \in \mathbb{C}[G]$ be given as the repeated convolution*

$$\mathbf{b} = \mathop{\ast}_{\substack{k \in H^\perp \\ k \neq \mathbf{0}}} (\chi_k \mathbf{a}) \tag{73}$$

then $\operatorname{supp}(\mathbf{b} * \mathbf{a}) \subset H$*, thus the operator* $\rho = \mathbf{b} * \mathbf{a}$ *is reduced to* H*. The eigenvalues of* ρ *are*

$$\widehat{\rho}(\xi) = \prod_{k \in H^\perp} \widehat{\mathbf{a}}(\xi + k). \tag{74}$$

Proof. We have

$$\rho = \mathbf{b} * \mathbf{a} = \underset{k \in H^\perp}{\ast} (\chi_k \mathbf{a}).$$

The shift formula $\widehat{\chi_k \mathbf{a}} = S_k \widehat{\mathbf{a}}$ yields

$$\widehat{\rho} = \prod_{k \in H^\perp} S_k \widehat{\mathbf{a}},$$

which proves (74). Since $\rho(\xi + k) = \rho(\xi)$ for all $k \in H^\perp$, it follows that $\operatorname{supp}(\rho) \subset H$, because function is supported on a lattice H if and only if its Fourier transform is periodic with respect to $H^\perp$, cfr. Theorem 14.

Note that for $k \in H^\perp$, the characters χ_k takes constant values on the cosets of H, hence the cyclic reduction is obtained by changing the values of $\mathbf{a}$ on the cosets of H according to χ_k.

Example 35. Consider the standard 5-point Laplacian stencil on the square lattice $\mathbb{Z}_{2n} \oplus \mathbb{Z}_{2n}$. Let $H = \mathbb{Z}_n \oplus \mathbb{Z}m$ and $K = G/H = \mathbb{Z}_2 \oplus \mathbb{Z}_2$. The characters on K take the following values on the cosets of H:

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

thus the (2-way) cyclic reduction to H is

$$\begin{bmatrix} & 1 & \\ 1 & 4 & 1 \\ & 1 & \end{bmatrix} * \begin{bmatrix} & 1 & \\ -1 & 4 & -1 \\ & 1 & \end{bmatrix} * \begin{bmatrix} & -1 & \\ 1 & 4 & 1 \\ & -1 & \end{bmatrix} * \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}$$

$$= \begin{bmatrix} & & & 1 & & & \\ & & 0 & 0 & 0 & & \\ & -2 & 0 & -32 & 0 & -2 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -32 & 0 & 132 & 0 & -32 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & -2 & 0 & -32 & 0 & -2 & \\ & & 0 & 0 & 0 & & \\ & & & 1 & & & \end{bmatrix}.$$

The generalised cyclic reduction algorithm presented here is of importance to boundary problems of the form discussed in Section 3.9, on triangle and simplexes. Furthermore it turns out to have some other advantages compared to the classical cyclic reduction, which is based on consecutive odd-even deductions in the same direction (1-way cyclic reduction). Whereas the classical

reduction is known to be unstable for Poisson problems, the condition number of the reduced operator grows exponentially in the number of reduction steps, on the other hand, with the two way (multi-) we have discussed here, the condition number of the reduced operator is *decreasing*.

Example 36. We consider the 5-point Laplacian stencil on a 32×32 square lattice with Dirichlet boundary conditions. A 1-way reduction to 1/16 of the original size increases the condition number of the reduced operator from 415 to $3 \cdot 10^{11}$. A 2-way reduction decreases the condition number to 20.9. The explanation comes from (74). The 2-way reduction is sampling the eigenvalues of the Laplacian in a much more uniform manner than the 1-way reduction. In the 1-way case small eigenvalues are multiplied by small, and large by large, so the condition number explodes.

3.11 Preconditioning with convolutional operators

So far we have discussed algorithms for special matrices; convolution operators on periodic domains and convolution operators on special domains with boundary conditions satisfied by reflection symmetries. In this section we will briefly discuss some applications of group theory to more general matrices.

Matrix multiplication by diagonals

Our first topic is the technique of *matrix multiplication by diagonals*. This is a computational technique which has been popular for matrix-matrix and matrix-vector multiplication with sparse structured matrices, such as matrices arising from 5-point or 7-point stencils on rectangular grids, but where (unlike our previous discussions) the coefficients vary over space. The matrices are stored as diagonals together with an off-set indicating the position of the diagonal. We revisit this technique and describe it in the language of groups and want to show that book keeping in matrices stored by diagonals is simplified by the notation of finite abelian groups.

Let G be a finite abelian group and $f \in \mathbb{C}[G]$. Let $D(f)$ denote the diagonal matrix corresponding to f, i.e. the matrix such that $(D(f)g)(j) = f(j)g(j)$ for all $g \in \mathbb{C}[G]$, and as before, for $\ell \in G$, we let S_ℓ is the shift matrix $(S_\ell g)(j) = g(j-\ell)$ for $g \in \mathbb{C}[G]$. We want to develop matrix algebra expressed in terms of sums of products of shift matrices and diagonal matrices, $S_\ell D(f)$. We ask the reader to verify the following result as an exercise:

Lemma 18. *Shift matrices and diagonals can be swapped according to the rule*

$$S_\ell D(f) = D(S_\ell f)\mathcal{S}_\ell. \tag{75}$$

Lemma 19 (Matrix representation by shifted diagonals). *Any matrix $A \in \mathrm{End}(\mathbb{C}[G])$ can be written as a sum of products of shift matrices with diagonal matrices,*

$$A = \sum_{\ell \in G} S_\ell D(a_\ell), \tag{76}$$

where and $a_\ell \in \mathbb{C}[G]$ *for all* $\ell \in G$.

Proof. Let δ be the delta-function, which is 1 in $\mathbf{0}$ and 0 elsewhere. Let $\delta_j = S_j\delta$. For a matrix A, and $i, j \in G$ let $A_{i,j} = (A\delta_j)(i)$ be its entries by classical index notation. For A defined in (76) we compute, using Lemma 19 that

$$(A\delta_j)(i) = f_{i-j}(j).$$

(The reader should check this for a simple example such as $G = \mathbb{Z}_4$). From this we see that we can represent any matrix $A \in \mathrm{End}(\mathbb{C}[G])$ by putting $A_{i,j}$ into $f_{i-j}(j)$, thus any A can be written in this form. □

The following is very easily verified (check an example with $G = \mathbb{Z}_n$):

Lemma 20 (Matrix–vector multiplication by diagonals). *Let* $A \in \mathrm{End}(\mathbb{C}[G])$ *be represented as in (76), then for* $x \in \mathbb{C}[G]$ *we have*

$$Ax = \sum_{\ell \in G} S_\ell(a_\ell \bullet x),$$

where $\bullet\colon \mathbb{C}[G] \times \mathbb{C}[G] \to \mathbb{C}[G]$ *denotes point wise product of vectors.*

Theorem 19 (Matrix–matrix multiplication by diagonals). *Let* $A, B, C \in \mathrm{End}(\mathbb{C}[G])$ *be represented as in (76) and let* $C = AB$. *Then*

$$c_r = \sum_{\ell \in G} (S_{-\ell} a_{r-\ell}) \bullet b_\ell.$$

Proof. We compute, setting $r = k + \ell$:

$$\begin{aligned} AB &= \sum_k S_k D(a_k) \sum_\ell S_\ell D(b_\ell) = \sum_{k,\ell} S_k S_\ell S_{-\ell} D(a_k) S_\ell D(b_\ell) \\ &= \sum_{k,\ell} S_{k+\ell} D((S_{-\ell} a_k) \bullet b_\ell) = \sum_r S_r D\left(\sum_\ell (S_{-\ell} a_{r-\ell}) \bullet b_\ell \right). \end{aligned}$$

□

Multiplication by diagonals is especially attractive for matrices such as 5-point stencils etc, where the number of diagonals is small. In this case we, of course, have to compute only those c_r for which $r = j + k$, where j and k are non-zero diagonals in A and B.

Preconditioning

We end this section with a brief result about matrix approximation using convolutional operators. The goal of preconditioning a linear system $Ax = b$ is to find an approximation $C \approx A$ such that $Cx = b$ can be easily solved. If C is a convolutional operator, we know that this can be easily solved by Fourier analysis or cyclic reduction (or a combination of these). What is the 'best' approximation of a general matrix by a convolution? This does depend on the norm we use to measure closeness. The Frobenius norm gives a particularly simple answer.

Definition 46 (Frobenius norm). *For $A, B \in \mathrm{End}(G)$, we define the Frobenius inner product*

$$(A, B)_F := \mathrm{trace}\,(A^h B),$$

where A^h denotes the complex conjugate and transpose of A. The Frobenius norm is

$$||A, B||_F = (A, B)_F^{\frac{1}{2}}.$$

Lemma 21. *The shift matrices S_j are orthogonal in the Frobenius inner-product*

$$(S_j, S_k)_F = \begin{cases} |G| & \text{if } j = k \\ 0 & \text{else} \end{cases}.$$

Proof. $S_j^h S_k = S_{k-j}$, which has diagonal entries all 1 if k=j and all 0 else. □

Theorem 20. *Let $A = \sum_j S_j a_j \in \mathrm{End}(\mathbb{C}[G])$ be represented by its diagonals $a_j \in \mathbb{C}[G]$. The best Frobenius norm approximation to A by a convolutional operator is given as*

$$C = \sum_{j \in G} c_j S_j,$$

where

$$c_j = \frac{1}{|G|} \sum_{k \in G} a_j(k) \in \mathbb{C}.$$

Proof. The shift matrices $\{S_j\}_{j \in G}$ form an orthogonal basis (with coefficients in $\mathbb{C}$) for the convolutional operators as a subspace of $\mathrm{End}(\mathbb{C}[G])$. Hence, we find the best approximation of $A \in \mathrm{End}(\mathbb{C}[G])$ in the subspace of convolutional operators by projecting A orthogonally onto the subspace in Frobenius inner product:

$$C = \sum_{j \in G} \frac{(S_j, A)_F}{(S_j, S_j)_F} S_j.$$

The result follows by noting that $(S_j, A)_F = \sum_k a_j(k)$. □

A lot more could have been said about preconditioning with convolutional operators, but time and space is limited so we leave this interesting topic at this point.

4 Domain symmetries and non-commutative groups

The topic of this chapter is applications of Fourier analysis on non-commutative groups in linear algebra. In particular we will as an example discuss the computation of matrix exponentials for physical problems being symmetric with respect to a discrete non-commutative group acting upon the domain. Assuming that the domain is discretized with a symmetry respecting discretization, we will show that by a change of basis derived from the irreducible representations of the group, the operator is block diagonalized. This simplifies the computation of matrix exponentials, eigenvalue problems and the solution of linear equations. The basic mathematics behind this Chapter is *representation theory of finite groups* [15, 16, 26]. Applications of this theory in scientific computing is discussed by a number of authors, see e.g. [2, 4, 6, 9, 12]. Our exposition, based on the *group algebra* is explained in detail in [1], which is intended to be a self contained introduction to the subject.

4.1 $\mathcal{G}$-equivariant matrices

A *group* is a set $\mathcal{G}$ with a binary operation $g, h \mapsto gh$, inverse $g \mapsto g^{-1}$ and identity element e, such that $g(ht) = (gh)t$, $eg = ge = g$ and $gg^{-1} = g^{-1}g = e$ for all $g, h, t \in \mathcal{G}$. We let $|\mathcal{G}|$ denote the number of elements in the group. Let $\mathcal{I}$ denote the set of indices used to enumerate the nodes in the discretization of a computational domain. We say that a group $\mathcal{G}$ *acts on* a set $\mathcal{I}$ (from the right) if there exists a product $(i, g) \mapsto ig : \mathcal{I} \times \mathcal{G} \to \mathcal{I}$ such that

$$ie = i \quad \text{for all } i \in \mathcal{I}, \tag{77}$$

$$i(gh) = (ig)h \quad \text{for all } g, h \in \mathcal{G} \text{ and } i \in \mathcal{I}. \tag{78}$$

The map $i \mapsto ig$ is a permutation of the set $\mathcal{I}$, with the inverse permutation being $i \mapsto ig^{-1}$. An action partitions $\mathcal{I}$ into disjoint *orbits*

$$\mathcal{O}_i = \{j \in \mathcal{I} \,:\, j = ig \text{ for some } g \in \mathcal{G}\},\ i \in \mathcal{I}.$$

We let $\mathcal{S} \subset \mathcal{I}$ denote a selection of *orbit representatives*, i.e. one element from each orbit. The action is called *transitive* if $\mathcal{I}$ consists of just a single orbit, $|\mathcal{S}| = 1$. For any $i \in \mathcal{I}$ we let the *isotropy subgroup at* i, $\mathcal{G}_i$ be defined as

$$\mathcal{G}_i = \{g \in \mathcal{G} \,:\, ig = i\}.$$

The action is *free* if $\mathcal{G}_i = \{e\}$ for every $i \in \mathcal{I}$, i.e., there are no fixed points under the action of $\mathcal{G}$.

Definition 47. *A matrix $\boldsymbol{A} \in \mathbb{C}^{\mathcal{I}\times\mathcal{I}}$, is $\mathcal{G}$-equivariant if*

$$\boldsymbol{A}_{i,j} = \boldsymbol{A}_{ig,jg} \quad \textit{for all } i, j \in \mathcal{I} \textit{ and all } g \in \mathcal{G}. \tag{79}$$

The definition is motivated by the result that if $\mathcal{L}$ is a linear differential operator commuting with a group of domain symmetries $\mathcal{G}$, and if we can find a set of discretization nodes $\mathcal{I}$ such that every $g \in \mathcal{G}$ acts on $\mathcal{I}$ as a permutation $i \mapsto ig$, then $\mathcal{L}$ can be discretized as a $\mathcal{G}$-equivariant matrix A, see [4, 6].

4.2 The group algebra

We will establish that $\mathcal{G}$ equivariant matrices are associated with (scalar or block) convolutional operators in the *group algebra*.

Definition 48. *The* group algebra $\mathbb{C}\mathcal{G}$ *is the complex vectorspace* $\mathbb{C}^{\mathcal{G}}$ *where each* $g \in \mathcal{G}$ *corresponds to a basis vector* $\boldsymbol{g} \in \mathbb{C}\mathcal{G}$. *A vector* $a \in \mathbb{C}\mathcal{G}$ *can be written as*

$$a = \sum_{g\in\mathcal{G}} a(g)\boldsymbol{g} \qquad \text{where } a(g) \in \mathbb{C}.$$

The convolution product $* : \mathbb{C}\mathcal{G} \times \mathbb{C}\mathcal{G} \to \mathbb{C}\mathcal{G}$ is induced from the product in $\mathcal{G}$ as follows. For basis vectors $\boldsymbol{g}, \boldsymbol{h}$, we set $\boldsymbol{g} * \boldsymbol{h} \equiv \boldsymbol{gh}$, and in general if $a = \sum_{g\in\mathcal{G}} a(g)\boldsymbol{g}$ and $b = \sum_{h\in\mathcal{G}} b(h)\boldsymbol{h}$, then

$$a * b = \left(\sum_{g\in\mathcal{G}} a(g)\boldsymbol{g}\right) * \left(\sum_{h\in\mathcal{G}} b(h)\boldsymbol{h}\right) = \sum_{g,h\in\mathcal{G}} a(g)b(h)(\boldsymbol{gh}) = \sum_{g\in\mathcal{G}} (a*b)(g)\boldsymbol{g},$$

where

$$(a*b)(g) = \sum_{h\in\mathcal{G}} a(gh^{-1})b(h) = \sum_{h\in\mathcal{G}} a(h)b(h^{-1}g). \tag{80}$$

Consider a $\mathcal{G}$-equivariant $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ in the case where $\mathcal{G}$ acts freely and transitively on $\mathcal{I}$. In this case there is only one orbit of size $|\mathcal{G}|$ and hence $\mathcal{I}$ may be identified with $\mathcal{G}$. Corresponding to $\boldsymbol{A}$ there is a unique $A \in \mathbb{C}\mathcal{G}$, given as $A = \sum_{g\in\mathcal{G}} A(g)g$, where A is the first column of $\boldsymbol{A}$, i.e.,

$$A(gh^{-1}) = \boldsymbol{A}_{gh^{-1},e} = \boldsymbol{A}_{g,h}. \tag{81}$$

Similarly, any vector $\boldsymbol{x} \in \mathbb{C}^n$ corresponds uniquely to $x = \sum_{g\in\mathcal{G}} x(g)g \in \mathbb{C}\mathcal{G}$, where $x(g) = \boldsymbol{x}_g$ for all $g \in \mathcal{G}$. Consider the matrix vector product:

$$(\boldsymbol{Ax})_g = \sum_{h\in\mathcal{G}} \boldsymbol{A}_{g,h}\boldsymbol{x}_h = \sum_{h\in\mathcal{G}} A(gh^{-1})x(h) = (A*x)(g).$$

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are two equivariant matrices, then $\boldsymbol{AB}$ is the equivariant matrix where the first column is given as

$$(\boldsymbol{AB})_{g,e} = \sum_{h\in\mathcal{G}} \boldsymbol{A}_{g,h}\boldsymbol{B}_{h,e} = \sum_{h\in\mathcal{G}} A(gh^{-1})B(h) = (A*B)(g).$$

We have shown that *if* $\mathcal{G}$ *acts freely and transitively, then the algebra of* $\mathcal{G}$*-equivariant matrices acting on* $\mathbb{C}^n$ *is isomorphic to the group algebra* $\mathbb{C}\mathcal{G}$ *acting on itself by convolutions from the left.*

In the case where $\boldsymbol{A}$ is $\mathcal{G}$-equivariant w.r.t. a free, but not transitive, action of $\mathcal{G}$ on $\mathcal{I}$, we need a block version of the above theory. Let $\mathbb{C}^{m\times\ell}\mathcal{G} \equiv \mathbb{C}^{m\times\ell} \otimes \mathbb{C}\mathcal{G}$ denote the space of vectors consisting of $|\mathcal{G}|$ matrix blocks, each block of size $m \times \ell$, thus $A \in \mathbb{C}^{m\times\ell}\mathcal{G}$ can be written as

$$A = \sum_{g\in\mathcal{G}} A(g)\otimes g \quad \text{where } A(g)\in\mathbb{C}^{m\times\ell}. \tag{82}$$

The convolution product (80) generalizes to a block convolution $*$: $\mathbb{C}^{m\times\ell}\mathcal{G}\times\mathbb{C}^{\ell\times k}\mathcal{G}\to\mathbb{C}^{m\times k}\mathcal{G}$ given as

$$A*B=\left(\sum_{g\in\mathcal{G}}A(g)\otimes g\right)*\left(\sum_{h\in\mathcal{G}}B(h)\otimes h\right)=\sum_{g\in\mathcal{G}}(A*B)(g)\otimes g,$$

where

$$(A*B)(g)=\sum_{h\in\mathcal{G}}A(gh^{-1})B(h)=\sum_{h\in\mathcal{G}}A(h)B(h^{-1}g), \tag{83}$$

and $A(h)B(h^{-1}g)$ denotes a matrix product.

If the action of $\mathcal{G}$ on $\mathcal{I}$ is free, but not transitive, then $\mathcal{I}$ split in m orbits, each of size $|\mathcal{G}|$. We let $\mathcal{S}$ denote a selection of one representative from each orbit. We will establish an isomorphism between the algebra of $\mathcal{G}$-equivariant matrices acting on $\mathbb{C}^n$ and the block-convolution algebra $\mathbb{C}^{m\times m}\mathcal{G}$ acting on $\mathbb{C}^m\mathcal{G}$. We define the mappings $\mu:\mathbb{C}^n\to\mathbb{C}^m\mathcal{G}$, $\nu:\mathbb{C}^{n\times n}\to\mathbb{C}^{m\times m}\mathcal{G}$ as:

$$\mu(\boldsymbol{y})_i(g)=y_i(g)=\boldsymbol{y}_{ig}\ \forall i\in\mathcal{S},\ g\in\mathcal{G}, \tag{84}$$

$$\nu(\boldsymbol{A})_{i,j}(g)=A_{i,j}(g)=\boldsymbol{A}_{ig,j}\ \forall i,j\in\mathcal{S}\ g\in\mathcal{G}. \tag{85}$$

In [1] we show:

Proposition 1. *Let $\mathcal{G}$ act freely on $\mathcal{I}$. Then μ is invertible and ν is invertible on the subspace of $\mathcal{G}$-equivariant matrices. Furthermore, if $\boldsymbol{A},\boldsymbol{B}\in\mathbb{C}^{n\times n}$ are $\mathcal{G}$-equivariant, and $\boldsymbol{y}\in\mathbb{C}^n$, then*

$$\mu(\boldsymbol{A}\boldsymbol{y})=\nu(\boldsymbol{A})*\mu(\boldsymbol{y}), \tag{86}$$

$$\nu(\boldsymbol{A}\boldsymbol{B})=\nu(\boldsymbol{A})*\nu(\boldsymbol{B}). \tag{87}$$

To complete the connection between $\mathcal{G}$-equivariance and block convolutions, we need to address the general case where the action is not free, hence some of the orbits in $\mathcal{I}$ have reduced size. One way to treat this case is to duplicate the nodes with non-trivial isotropy subgroups, thus a point $j\in\mathcal{I}$ is considered to be $|\mathcal{G}_j|$ identical points, and the action is extended to a free action on this extended space. Equivariant matrices on the original space is extended by duplicating the matrix entries, and scaled according to the size of the isotropy. We define

$$\mu(\boldsymbol{x})_i(g)=x_i(g)=\boldsymbol{x}_{ig}\ \forall i\in\mathcal{S},\ g\in\mathcal{G}, \tag{88}$$

$$\nu(\boldsymbol{A})_{i,j}(g)=A_{i,j}(g)=\frac{1}{|\mathcal{G}_j|}\boldsymbol{A}_{ig,j}\ \forall i,j\in\mathcal{S}\ g\in\mathcal{G}. \tag{89}$$

With these definitions it can be shown that (86)–(87) still hold. It should be noted that μ and ν are no longer invertible, and the extended block convolutional operator $\nu(\boldsymbol{A})$ becomes singular. This poses no problems for the

computation of exponentials since this is a forward computation. Thus we just exponentiate the block convolutional operator and restrict the result back to the original space. However, for inverse computations such as solving linear systems, the characterization of the image of μ and ν as a subspaces of $\mathbb{C}^m\mathcal{G}$ and $\mathbb{C}^{m\times m}\mathcal{G}$ is an important issue for finding the correct solution [1, 3].

4.3 The Generalized Fourier Transform (GFT)

So far we have argued that a differential operator with spatial symmetries becomes a $\mathcal{G}$-equivariant matrix under discretization, which again can be represented as a block convolutional operator. In this section we will show how convolutional operators are block diagonalized by a Fourier transform on $\mathcal{G}$. This is the central part of Frobenius' theory of group representations from 1897–1899. We recommend the monographs [10, 15, 16, 26] as introductions to representation theory with applications.

Definition 49. *A d-dimensional group representation is a map $R : \mathcal{G} \to \mathbb{C}^{d\times d}$ such that*

$$R(gh) = R(g)R(h) \quad \textit{for all } g, h \in \mathcal{G}. \tag{90}$$

Generalizing the definition of *Fourier coefficients* we define for any $A \in \mathbb{C}^{m\times k}\mathcal{G}$ and any d-dimensional representation R a matrix $\hat{A}(R) \in \mathbb{C}^{m\times k}\otimes\mathbb{C}^{d\times d}$ as:

$$\hat{A}(R) = \sum_{g\in\mathcal{G}} A(g)\otimes R(g). \tag{91}$$

Proposition 2 (The convolution theorem). *For any $A \in \mathbb{C}^{m\times k}\mathcal{G}$, $B \in \mathbb{C}^{k\times \ell}\mathcal{G}$ and any representation R we have*

$$\widehat{(A * B)}(R) = \hat{A}(R)\hat{B}(R). \tag{92}$$

Proof. The statement follows from

$$\begin{aligned}
\hat{A}(R)\hat{B}(R) &= \left(\sum_{g\in\mathcal{G}} A(g)\otimes R(g)\right)\left(\sum_{h\in\mathcal{G}} B(h)\otimes R(h)\right)\\
&= \sum_{g,h\in\mathcal{G}} A(g)B(h)\otimes R(g)R(h) = \sum_{g,h\in\mathcal{G}} A(g)B(h)\otimes R(gh)\\
&= \sum_{g,h\in\mathcal{G}} A(gh^{-1})B(h)\otimes R(g) = \widehat{(A * B)}(R).
\end{aligned}$$

Let d_R denote the dimension of the representation. For use in practical computations, it is important that $A * B$ can be recovered by knowing $\widehat{(A * B)}(R)$ for a suitable selection of representations, and furthermore that their dimensions d_R are as small as possible. Note that if R is a representation and $X \in \mathbb{C}^{d_R\times d_R}$ is non-singular, then also $\tilde{R}(g) = XR(g)X^{-1}$ is a representation. We say that R and $\tilde{R}$ are equivalent representations. If there

exists a similarity transform $\tilde{R}(g) = XR(g)X^{-1}$ such that $\tilde{R}(g)$ has a block diagonal structure, independent of $g \in \mathcal{G}$, then R is called *reducible*, otherwise it is *irreducible*.

Theorem 21 (Frobenius). *For any finite group $\mathcal{G}$ there exists a complete list $\mathcal{R}$ of non-equivalent irreducible representations such that*

$$\sum_{R\in\mathcal{R}} d_R^2 = |\mathcal{G}|.$$

Defining the GFT for $a \in \mathbb{C}\mathcal{G}$ as

$$\hat{a}(R) = \sum_{g\in\mathcal{G}} a(g)R(g) \quad \textit{for every } R \in \mathcal{R}, \tag{93}$$

*we may recover a by the inverse GFT (*IGFT*):*

$$a(g) = \frac{1}{|\mathcal{G}|} \sum_{R\in\mathcal{R}} d_R \mathrm{trace}\,(R(g^{-1})\hat{a}(R)). \tag{94}$$

For the block transform of $A \in \mathbb{C}^{m\times k}\mathcal{G}$ given in (91), the GFT and the IGFT are given componentwise as

$$\hat{A}_{i,j}(R) = \sum_{g\in\mathcal{G}} A_{i,j}(g)R(g) \in \mathbb{C}^{d_R\times d_R}, \tag{95}$$

$$A_{i,j}(g) = \frac{1}{|\mathcal{G}|} \sum_{R\in\mathcal{R}} d_R \mathrm{trace}\,(R(g^{-1})\hat{A}_{i,j}(R)). \tag{96}$$

Complete lists of irreducible representations for a selection of common groups are found in [16].

4.4 Applications to the matrix exponential

We have seen that via the GFT, any $\mathcal{G}$-equivariant matrix is block diagonalized. Corresponding to an irreducible representation R, we obtain a matrix block $\hat{A}(R)$ of size $md_R \times md_R$, where m is the number of orbits in $\mathcal{I}$ and d_R the size of the representation. Let W_{direct} denote the computational work, in terms of floating point operations, for computing the matrix exponential on the original data A, and let W_{fspace} be the cost of doing the same algorithm on the corresponding block diagonal GFT transformed data $\hat{A}$. Thus $W_{\text{direct}} = c(m|\mathcal{G}|)^3 = cm^3\left(\sum_{R\in\mathcal{R}} d_R^2\right)^3$, $W_{\text{fspace}} = cm^3\sum_{R\in\mathcal{R}} d_R^3$ and the ratio becomes

$$\mathcal{O}(n^3) \;:\quad W_{\text{direct}}/W_{\text{fspace}} = \left(\sum_{R\in\mathcal{R}} d_R^2\right)^3 \Big/ \sum_{R\in\mathcal{R}} d_R^3.$$

Table 1. Gain in computational complexity for matrix exponential via GFT.

Domain	$\mathcal{G}$	$\lvert\mathcal{G}\rvert$	$\{d_R\}_{R\in\mathcal{R}}$	$W_{\text{direct}}/W_{\text{fspace}}$
triangle	$\mathcal{D}_3$	6	$\{1,1,2\}$	21.6
tetrahedron	$\mathcal{S}_4$	24	$\{1,1,2,3,3\}$	216
cube	$\mathcal{S}_4\times\mathcal{C}_2$	48	$\{1,1,1,1,2,2,3,3,3,3\}$	864
icosahedron	$\mathcal{A}_5\times\mathcal{C}_2$	120	$\{1,1,3,3,3,3,4,4,5,5\}$	3541

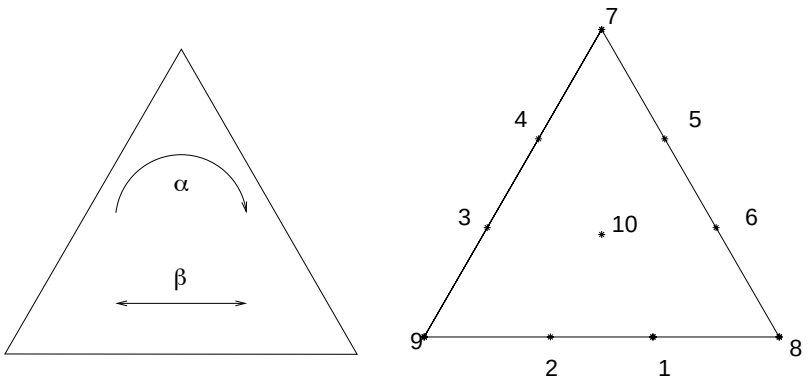


Fig. 3. Equilateral triangle with a symmetry preserving set of 10 nodes.

Table 1 lists this factor for the symmetries of the triangle, the tetrahedron, the 3D cube and the maximally symmetric discretization of a 3D sphere (icosahedral symmetry with reflections).

The cost of computing the GFT is not taken into account in this estimate. There exists fast GFT algorithms of complexity $\mathcal{O}\left(|\mathcal{G}|\log^{\ell}(|\mathcal{G}|)\right)$ for a number of groups, but even if we use a slow transform of complexity $\mathcal{O}\left(|\mathcal{G}^2|\right)$, the total cost of the GFT becomes just $\mathcal{O}\left(m^2|\mathcal{G}|^2\right)$, which is much less than W_{fspace}.

Example: Equilateral triangle

The smallest noncommutative group is $\mathcal{D}_3$, the symmetries of an equilateral triangle. There are six linear transformations that map the triangle onto itself, three pure rotations and three rotations combined with reflections. In Figure 3a we indicate the two generators α (rotation 120° clockwise) and β (right-left reflection). These satisfy the algebraic relations $\alpha^3 = \beta^2 = e$, $\beta\alpha\beta = \alpha^{-1}$, where e denotes the identity transform. The whole group is $\mathcal{D}_3 = \left\{e, \alpha, \alpha^2, \beta, \alpha\beta, \alpha^2\beta\right\}$.

Given an elliptic operator $\mathcal{L}$ on the triangle such that $\mathcal{L}(u\circ\alpha) = \mathcal{L}(u)\circ\alpha$ and $\mathcal{L}(u\circ\beta) = \mathcal{L}(u)\circ\beta$ for any u satisfying the appropriate boundary conditions on the triangle, let the domain be discretized with a *symmetry respecting*

discretization, see Figure 3b. In this example we consider a finite difference discretization represented by the nodes $\mathcal{I} = \{1, 2, \ldots, 10\}$, such that both α and β map nodes to nodes. In finite element discretizations one would use basis functions mapped to other basis functions by the symmetries. We define the action of $\mathcal{D}_3$ on $\mathcal{I}$ as

$$(1,2,3,4,5,6,7,8,9,10)\alpha = (5,6,1,2,3,4,9,7,8,10),$$
$$(1,2,3,4,5,6,7,8,9,10)\beta = (2,1,6,5,4,3,7,9,8,10),$$

and extend to all of $\mathcal{D}_3$ using (78). As orbit representatives, we may pick $\mathcal{S} = \{1, 7, 10\}$. The action of the symmetry group is free on the orbit $\mathcal{O}_1 = \{1, 2, 3, 4, 5, 6\}$, while the points in the orbit $\mathcal{O}_7 = \{7, 8, 9\}$ have isotropy subgroups of size 2, and finally $\mathcal{O}_{10} = \{10\}$ has isotropy of size 6.

Table 2. A complete list of irreducible representations for $\mathcal{D}_3$.

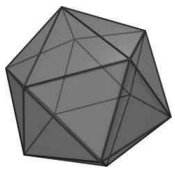
	α	β
ρ_0	1	1
ρ_1	1	-1
ρ_2	$\begin{pmatrix} -1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

The operator $\mathcal{L}$ is discretized as a matrix $\boldsymbol{A} \in \mathbb{C}^{10\times 10}$ satisfying the equivariances $\boldsymbol{A}_{ig,jg} = \boldsymbol{A}_{i,j}$ for $g \in \{\alpha, \beta\}$ and $i, j \in \mathcal{S}$. Thus we have e.g. $\boldsymbol{A}_{1,6} = \boldsymbol{A}_{3,2} = \boldsymbol{A}_{5,4} = \boldsymbol{A}_{4,5} = \boldsymbol{A}_{2,3} = \boldsymbol{A}_{6,1}$.

$\mathcal{D}_3$ has three irreducible representations given in Table 2 (extended to the whole group using (90)). To compute $\exp(\boldsymbol{A})$, we find $A = \nu(\boldsymbol{A}) \in \mathbb{C}^{3\times 3}\mathcal{G}$ from (89) and find $\hat{A} = \mathrm{GFT}(A)$ from (95). The transformed matrix $\hat{A}$ has three blocks, $\hat{A}(\rho_0), \hat{A}(\rho_1) \in \mathbb{C}^{m\times m}$ and $\hat{A}(\rho_2) \in \mathbb{C}^{m\times m}\otimes\mathbb{C}^{2\times 2} \simeq \mathbb{C}^{2m\times 2m}$, where $m = 3$ is the number of orbits. We exponentiate each of these blocks, and find the components of $\exp(\boldsymbol{A})$ using the Inverse GFT (96).

We should remark that in Lie group integrators, it is usually more important to compute $y = \exp(A)\cdot x$ for some vector x. In this case, we compute $\hat{y}(\rho_i) = \exp(\hat{A}(\rho_i))\cdot\hat{x}(\rho_i)$, and recover y by Inverse GFT. Note that $\hat{x}(\rho_2), \hat{y}(\rho_2) \in \mathbb{C}^m\otimes\mathbb{C}^{2\times 2} \simeq \mathbb{C}^{2m\times 2}$.

Example: Icosahedral symmetry



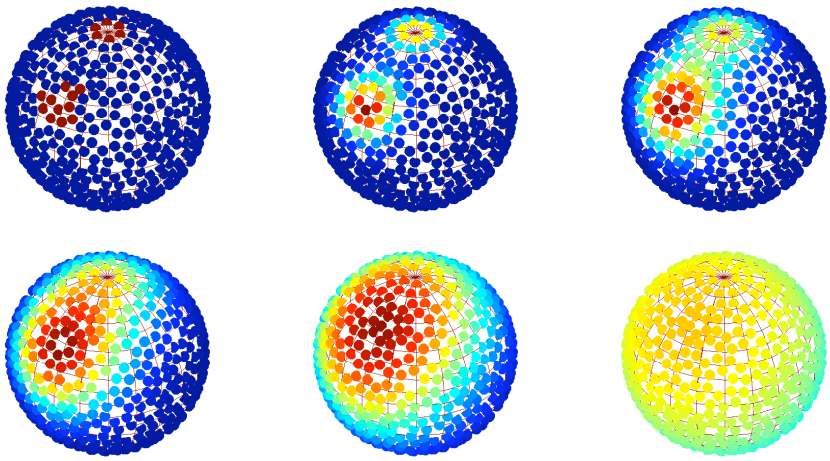
As a second example illustrating the general theory, we solve the simple heat equation

$$u_t = \nabla^2 u$$

on the surface of a unit sphere.

The sphere is divided into 20 equilateral triangles, and each triangle subdivided in a finite difference mesh respecting all the 120 symmetries of the full icosahedral symmetry group (including reflections). To understand this group, it is useful to realize that five tetrahedra can be simultaneously embedded in the icosahedron, so that the 20 triangles correspond to the in total 20 corners of these five tetrahedra. From this one sees that the icosahedral rotation group is isomorphic to A_5, the group of all 60 even permutations of the five tetrahedra. The 3-D reflection matrix $-I$ obviously commutes with any 3-D rotation, and hence we realize that the full icosahedral group is isomorphic to the direct product $C_2 \times A_5$, where $C_2 = \{1, -1\}$. The irreducible representations of A_5, listed in Lomont have dimensions $\{1, 3, 3, 4, 5\}$, and the representaitons of the full icosahedral group are found by taking tensor products of these with the two 1-dimensional representations of C_2. The fact that the full icosahedral group is a direct product is also utilized in faster computation of the GFT. This is, however, not of major importance, since the cost of the GFT in any case is much less than the cost of the matrix exponential.

The figures below show the solution of the heat equation at times 0, 2, 5, 10, 25 and 100. The initial condition consists of two located heat sources in the northern hemisphere.



5 Concluding remarks

We have in these lectures presented the basic concepts of group theory in a setting aimed at understanding computational algorithms. Some applications within computational mathematics have been discussed in detail, others in a more sketchy manner and many topics have been omitted altogether. Among the omissions, we would in particular point to the theory of multivariate Chebyshev approximations, a beautiful application of group theory which

originate from the study of kaleidoscopic reflection groups (Coxeter groups), and which has connections to many areas of mathematics, the representation theory of Lie groups in particular. The multivariate Chebyshev polynomials share the excellent approximation properties of the classical univariate case, and the multivariate polynomials are defined on domains that are related to simplexes in higher dimensions.

We have discussed Fourier analysis on abelian groups and on finite non-abelian groups. The next steps along this line is the Fourier analysis on compact Lie groups, where the fundamental Peter–Weyl theorem states that the countably infinite list of non-equivalent irreducible representations gives a complete orthogonal basis for $L^2(G)$. Certain non-compact groups (the unimodular groups) such as the Euclidean group of rigid motions in $\mathbb{R}^n$ can be derived from the compact case and is of significant interest in image and signal processing.

Finally, we would like to mention the topic of time integration of differential equations evolving on manifolds. The so-called *Lie group* integrators advance the solution by computing the action of a Lie group on the domain. This topic has been developed in substantial detail over the last two decades and has lead to many theoretical insights and practical computational algorithms [14].

Acknowledgements

I would like to express a deep gratitude towards CIME and the organisers of this summer school for inviting me to present these lectures and for their patience with me during the tortuous process of writing the lecture notes. Also, I would like to thank Ulrich von der Ohe for his careful reading and commenting upon the manuscript.

References

1. K. Åhlander and H. Munthe-Kaas. Applications of the Generalized Fourier Transform in numerical linear algebra. *BiT*, 45(4):819–850, 2005.
2. E. L. Allgower, K. Böhmer, K. Georg, and R. Miranda. Exploiting symmetry in boundary element methods. *SIAM J. Numer. Anal.*, 29:534–552, 1992.
3. E. L. Allgower, K. Georg, and R. Miranda. Exploiting permutation symmetry with fixed points in linear equations. In E. L. Allgower, K. Georg, and R. Miranda, editors, *Lectures in Applied Mathematics*, volume 29, pages 23–36. American Mathematical Society, Providence, RI, 1993.
4. E. L. Allgower, K. Georg, R. Miranda, and J. Tausch. Numerical exploitation of equivariance. *Zeitschrift für Angewandte Mathematik und Mechanik*, 78:185–201, 1998.
5. L. Auslander and R. Tolimieri. Is computing with the finite Fourier transform pure or applied mathematics. *Notices of the AMS*, 1(6), 1979.

6. A. Bossavit. Symmetry, groups, and boundary value poblems. a progressive introduction to noncommutative harmonic analysis of partial differential equations in domains with geometrical symmetry. *Comput. Methods Appl. Mech. and Engrg.*, 56:167–215, 1986.
7. S.H. Christiansen, H.Z. Munthe-Kaas, and B. Owren. Topics in structure-preserving discretization. *Acta Numerica*, 20(1):1–119, 2011.
8. J.H. Conway, N.J.A. Sloane, and E. Bannai. *Sphere Packings, Lattices, and Groups*, volume 290. Springer Verlag, 1999.
9. C. C. Douglas and J. Mandel. Abstract theory for the domain reduction method. *Computing*, 48:73–96, 1992.
10. A. F. Fässler and E. Stiefel. *Group theoretical methods and their applications*. Birkhäuser, Boston, 1992.
11. Claude Gasquet and Patrick Witomski. *Fourier analysis and applications: filtering, numerical computation, wavelets*, volume 30. Springer Science & Business Media, 2013.
12. K. Georg and R. Miranda. Exploiting symmetry in solving linear equations. In E. L. Allgower, K. Böhmer, and M. Golubisky, editors, *Bifurcation and Symmetry*, volume 104 of *ISNM*, pages 157–168. Birkhäuser, Basel, 1992.
13. M. E. Hoffman and W. D. Withers. Generalized Chebyshev polynomials associated with affine Weyl groups. *Transactions of the AMS*, 308(1):91–104, 1988.
14. A. Iserles, H. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. In *Acta numerica, 2000*, volume 9 of *Acta Numer.*, pages 215–365. Cambridge Univ. Press, Cambridge, 2000.
15. G. James and M. Liebeck. *Representations and Characters of Groups*. Cambridge University Press, 2 edition, 2001. ISBN 052100392X.
16. J. S. Lomont. *Applications of Finite Groups*. Academic Press, New York, 1959.
17. Saunders Mac Lane. *Categories for the working mathematician*, volume 5. Springer Science & Business Media, 2013.
18. H. Munthe-Kaas. Symmetric FFTs; a general approach. Technical report, NTNU, Trondheim, Norway, 1989. Available at: `http://hans.munthe-kaas.no`.
19. H. Munthe-Kaas. *Topics in linear algebra for vector- and parallel computers*. PhD thesis, Norwegian University of Science and Technology (NTNU), 1989.
20. H. Munthe-Kaas and T. Sørevik. Multidimensional pseudo-spectral methods on lattice grids. *Applied Numerical Mathematics*, 62(3):155–165, 2012.
21. H.Z. Munthe-Kaas. On group Fourier analysis and symmetry preserving discretizations of PDEs. *Journal of Physics A: Mathematical and General*, 39:5563, 2006.
22. HZ Munthe-Kaas, M Nome, and BN Ryland. Through the Kaleidoscope; Symmetries, Groups and Chebyshev Approximations from a Computational Point of View. *Foundations of Computational Mathematics*, pages 188–229, 2011.
23. M Scott Osborne. On the Schwartz-Bruhat space and the Paley-Wiener theorem for locally compact abelian groups. *Journal of Functional Analysis*, 19(1):40–49, 1975.
24. W. Rudin. *Fourier analysis on groups*, volume 12. Wiley-Interscience, 1990.
25. B.N. Ryland and H.Z. Munthe-Kaas. On Multivariate Chebyshev Polynomials and Spectral Approximations on Triangles. In J. S. Hesthaven and E. M. Rønquist, editors, *Spectral and High Order Methods for Partial Differential Equations*, volume 76 of *Lecture Notes Comp. Sci. and Eng.*, pages 19–41. Springer, 2011.

26. J. P. Serre. *Linear Representations of Finite Groups*. Springer, 1977. ISBN 0387901906.
27. R.J. Stanton and P.A. Tomas. Polyhedral summability of Fourier series on compact Lie groups. *American Journal of Mathematics*, 100(3):477–493, 1978.
28. Paul N Swarztrauber. The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle. *Siam Review*, 19(3):490–501, 1977.
29. S. Thangavelu. *Harmonic analysis on the Heisenberg group*, volume 159. Birkhauser, 2012.
30. G. Travaglini. Polyhedral summability of multiple Fourier series. In *Colloq. Math*, volume 65, pages 103–116, 1993.
31. Wikipedia. Smith normal form — Wikipedia, the free encyclopedia, 2015.