Math 362

Representation Theory

Claudio Gómez-Gonzáles

Table of contents

Preface

These lectures notes are intended as the primary reference for an introductory course in Representation Theory, with a preliminary emphasis on finite groups and eventual broadening into the theory of compact groups. The material is intended for a trimester-long course at the advanced undergraduate or early graduate level, with additional topics of interest included to supplement potential forays, Any errors are due to Claudio, who welcomes feedback and corrections.

Potential additional references include:

- *Representation Theory: A First Course* (Fulton and Harris 1991). Often regarded as the definitive introductory text, though it relies on familiarity with abstract frameworks.
- *Linear Representations of Finite Groups* (Serre 1977). This is affectionately known as "Serre's little book," a concise and elegant translation that hits a variety of topics. This book was written for Josiane Heulot-Serre (the author, Jean-Pierre, was her husband) to teach students in quantum chemistry, though it is decidely a mathematical text.
- *Linear Algebra and Group Representations* (Shaw 1982). Features many examples, where much of the needed linear algebra and group theory is developed as the book progresses.
- *Character Theory of Finite Groups* (Isaacs 1994). A great reference on the relevant techniques in character theory, including an introduction to the theory of Schur covers.
- *Group Representations in Probability and Statistics* (Diaconis 1988). A succinct introduction, geared towards developing Fourier theory for finite groups.
- *Lie Groups: An Introduction through Linear Groups* (Rossmann 2002). A concrete introduction to Lie theory, emphasizing structure and representation with an eye toward applications in physics and geometry.
- *Abstract Algebra* (Dummit and Foote 2003). This textbook was used recently for Math 342 and is a generally useful reference. Its chapters 18 and 19 cover some of the material we will discuss using distinct frameworks and with different motivations.

Conventions

In these notes and during classtime we will adhere to the following conventions.

We use the standard notation $\mathbb{Z}$ and $\mathbb{Z}^+$ to stand for the sets of integers and positive integers, respectively. Perhaps controversially, we will use $\mathbb{N}$ to stand for the non-negative integers (that is, in these notes we say that $0 \in \mathbb{N}$). We use $\mathbb{Q}$ and $\mathbb{Q}^+$ for the sets of rationals and positive rationals, respectively, and similarly write $\mathbb{R}$ and $\mathbb{R}^+$ for the sets of reals and positive reals, respectively. We also denote the complex numbers by $\mathbb{C}$.

Capitalized letters will be used to stand for sets, possibly with additional structure, and for linear maps. The letters G, H, K will be used for groups, where the latter might specifically indicate the kernel of a homomorphism when made clear from context. The letter N will often be used to denote a normal subgroup, especially if it is not evidentally the kernel of some homomorphism. The letters R and S will stand for rings and algebras; F and E for fields; U, V, and W for vector spaces; X and Y will stand for sets, perhaps beset upon by a group action. The letters B and C stand in reserve for algebraic objects if we exhaust other options, (e.g., where the use of K might be confused with a kernel). In addition, letters like A, B, L, M, P, Q, and T refer to operators.

Lowercase letters like n, m, and ℓ will be reserved for integers; i, j, and k will be used for indices. Note that i, j, and k might be used for imaginary numbers or quaternions, though again such context will always be clear. Letters like g, h, and k will be used to stand for group elements. If more than one group is present, such as G and H, lower case letters will be used harmoniously with upper case letters: $g \in G$, $h \in H$, and so on. Similarly, r and s will be used for ring elements. We may use primes or indices to denote multiple elements in the same group, e.g., $g, g' \in G$ or $g_1, g_2 \in G$, when other groups are present to make something like $g, h \in G$ unnecessarily confusing.

Note that x and y can sometimes be used to stand for elements from arbitrary sets, perhaps those under the influence of a group action, or as indeterminants in a polynomial ring. In linear algebra, we use x, y, and z to stand for vectors and u, v, or w, decorated with subscripts, to stand for basis elements. The usual e_i notation is used for the standard basis elements of F^n. We will use t to denote a real parameter, with s as a potential auxiliary, and also as the indeterminant for characteristic polynomials.

The Greek letters ρ, σ, and τ will be used for group homomorphisms; ε stands for the sign representation $\mathcal{S}_n \to \mathcal{C}_2 = \{1, -1\}$. We will also use σ and τ to denote permutations if there is no danger of confusion. Note that ζ is sometimes used to stand for a root of unity, especially $\zeta_n := e^{2\pi i/n}$ as a primitive nth root. As is customary, χ is used for characters of representations. The letters φ and ψ are used especially for intertwiners; μ is used to indicate a bi-linear form. Meanwhile, α, β, γ, and λ are scalars, where the latter is used especially to indicate an eigenvalue.

We sometimes use the hooked arrow $\hookrightarrow$ to denote a one-to-one function, especially an inclusion: e.g., $X \hookrightarrow Y$. Dually, the double-headed $\twoheadrightarrow$ is used to emphasize that a map is onto. We write $\mathbb{I}_X$ for the

identity map $X \to X$, or simply $\mathbb{I}$ when context is clear. Lastly, we write

$$X \xrightarrow{\cong} Y$$

for bijections (or isomorphisms, if a category is clear) or simply $X \leftrightarrow Y$ when directionality is unimportant.

Part I

Preliminaries

Chapter 1

Operations

Here we review, with references but without great detail, the basic definitions and results that facilitate our study of representation theory.

Definition 1.1. A **binary operation** on a set X is a map

$$* : X \times X \to X.$$

We say that $*$ is:

- **unital** if there exists an element $e \in X$ such that $e * x = x = x * e$ for all $x \in X$.
- **associative** if $(x * y) * z = x * (y * z)$ for every $x, y, z \in X$.
- **commutative** if $x * y = y * x$ for every $x, y \in X$.

Example 1.1 (Arithmetic operations)**.** We are accustomed to operations on $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$:

- **Addition** $(+)$ and **multiplication** $(\cdot)$ are unital, commutative, and associative binary operations on these sets.
- **Subtraction** is a (non-unital, non-associative, non-commutative) binary operation on these sets.
- **Division** is a (non-unital, non-associative, non-commutative) binary operation on the sets $\mathbb{Q} - \{0\}, \mathbb{R} - \{0\}, \mathbb{C} - \{0\}$, but not on $\mathbb{Z} - \{0\}$.

Example 1.2 (Matrix operations)**.** If we write $\mathrm{Mat}_n(\mathbb{R})$ for the set of $n \times n$ matrices with real entries, or more generally $\mathrm{Mat}_n(F)$ with respect to some arbitrary field F, then:

- **Matrix addition** is a unital, commutative, and associative binary operation.
- **Matrix multiplication** is a unital and associative, but non-commutative, binary operation.

Example 1.3 (Function composition)**.** If X is a set, we write $\mathrm{Func}(X, X)$ to denote the set of functions $X \to X$. The operation of **composition**, i.e., assigning $f, g \in \mathrm{Func}(X, X)$ to the map $f \circ g : X \to X$ which sends $x \mapsto f(g(x))$, is associative and unital but generally non-commutative.

Example 1.4 (Cross product)**.** The standard cross-product $\times$ on $\mathbb{R}^3$ is a non-unital, non-commutative, non-associative binary operation.

Example 1.5 (Rock-paper-scissors)**.** The following operation on the set $\{r, p, s\}$ is commutative, but non-associative and non-unital:

$*$	r	p	s
r	r	p	r
p	p	p	s
s	r	s	s

In particular, $s = p * s = (r * p) * s \neq r * (p * s) = r * s = r$.

Proposition 1.1 (Uniqueness of identity)**.** *If $*$ is a unital binary operation on a set X with identity e, and there is an element $e' \in X$ satisfying $e' * x = x$ for all $x \in X$, then $e' = e$.*

Definition 1.2. Let $*$ be a unital binary operation on a set X and fix $x \in X$. If there is an element $y \in X$ such that $x * y = e = y * x$, then we say that x is **invertible** with **inverse** y.

Proposition 1.2 (Uniqueness of inverses)**.** *If $*$ is an associative unital binary operation on a set X and $x, y, z \in X$ satisfy $x * y = e = x * z$, then $y = z$.*

Example 1.6 (Negatives)**.** Every element in $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ is invertible with respect to $+$.

Example 1.7 (Fractions)**.** The element $2 \coloneqq 1 + 1$ is invertible with respect to $\cdot$ in $\mathbb{Q}$ but not $\mathbb{Z}$. Indeed, every non-zero element of $\mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$ is invertible with respect to $\cdot$.

Example 1.8 (Invertible functions)**.** A function $f \in \mathrm{Func}(X, X)$ is called a **permutation** of X if it is a bijection; we write $\mathrm{Perm}(X) \subseteq \mathrm{Func}(X, X)$ for the set of all permutations of X. These are exactly the invertible elements (with respect to composition) in Example 1.3.

Definition 1.3 (Groups)**.** (Dummit and Foote 2003, 16) A **group** is a set G together with a unital associative binary operation $*$, such that every $g \in G$ is invertible. If $*$ is commutative, then G is called an **abelian group**.[1]

Definition 1.4 (Rings)**.** (Dummit and Foote 2003, 223) A set R equipped with a pair of binary operations, usually named addition ($+$) and multiplication ($\cdot$), is called a **ring** if:

- R is an abelian group with respect to addition.
- multiplication is unital[2] and associative.
- distributive laws hold:

  $$(x + y) \cdot z = x \cdot z + y \cdot z$$

  and

  $$x \cdot (y + z) = x \cdot y + x \cdot z,$$

  for any $x, y, z \in R$.

The identity of $+$ is called 0_R while the identity of $\cdot$ is written 1_R, or simply as 0 and 1 if context is clear. If multiplication is commutative, we say that R is a **commutative ring**.

Definition 1.5 (Fields)**.** If F is a commutative ring in which every non-zero element is invertible with respect to multiplication, then we say that F is a **field**. Equivalently, F is a field if the set $F - \{0\}$ is a group with respect to multiplication.

[1] This word is in reference to the early 19th century mathematician, Niels Abel, who proved that polynomials whose Galois groups are commutative can always be *solved* in some precise sense.

[2] In some books—e.g., Dummit and Foote (2003)—rings are not required to contain a multiplicative identity. This is a matter of taste.

Remark 1.1. If F and E are fields with $F \subseteq E$, we say that F is a **subfield** of E and that E is a **field extension** of F. We often denote this relationship by E/F. While this notation may remind the reader of quotients (Definition 2.12 and Definition 3.11), we will never consider quotients of fields and so it remains unambiguous. This material is discussed thoroughly in introductory courses in Abstract Algebra and Galois Theory.

Chapter 2

Linear Algebra

We include two primary references in this section: one is the textbook for Math 232 (Bretscher 2013) and the other is a standard textbook for advanced courses in linear algebra (Petersen 2012). We include the former for its familiarity and the latter for its abstractness and formalism. For example, results in *Linear Algebra with Applications* (Bretscher 2013) are stated over the real numbers and there is no discussion of Hermitian inner products; much of *Linear Algebra* (Petersen 2012) is done over an arbitrary field and spectral theory is given a robust treatment. For those interested in a compromise, featuring a pragmatic balance of theory with a special focus on materials of interest to physicists (e.g., separable Hilbert spaces, differential and integral operators, differential equations, etc.), we submit *Introduction to Hilbert Spaces with Applications* (Debnath and Mikusinski 2005). For another (decidedly abstract) treatment, see *Abstract Algebra* (Dummit and Foote 2003).

2.1 Basic Concepts

Definition 2.1 (Vector spaces)**.** (Bretscher 2013, 167; Petersen 2012, 8) Let F be a field. An F**-vector space** is a set V with an operation $+ : V \times V \to V$ and map $F \times V \to V$, denoted by

$$(x, y) \mapsto x + y$$

and

$$(\alpha, x) \mapsto \alpha x,$$

which are called **vector addition** and **scalar multiplication**, respectively, so that

- $V, +$ is an abelian group.
- **Identity.** $1x = x$ for all $x \in V$.
- **Compatibility.** $\alpha(\beta x) = (\alpha\beta)x$ for all $\alpha, \beta \in F$, and $x \in V$.
- **Distributivity over scalars.** $(\alpha + \beta)x = \alpha x + \beta x$ for all $\alpha, \beta \in F$, and $x \in V$.
- **Distributivity over vectors.** $\alpha(x + y) = \alpha x + \alpha y$ for all $\alpha \in F$, and $x, y \in V$.

The identity vector with respect to addition is written $0 \in V$.

Definition 2.2 (Algebras)**.** (Dummit and Foote 2003, 342) If an F-vector space V is equipped with another binary operation $\cdot$ giving it a ring structure and also satisfying

$$\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y),$$

for all $\alpha \in F$ and $x, y \in V$, then V is an F-**algebra**. If $\cdot$ is commutative then V is a **commutative** F-**algebra**; otherwise, we might emphasize that V is **non-commutative**. If every nonzero element in V has a multiplicative inverse, we call it a **division algebra**.

Example 2.1. Every field F is a commutative F-algebra; the trivial ring 0 is also an F-algebra.

Example 2.2. $\mathbb{C}$ is a commutative $\mathbb{R}$-algebra with its usual field operations.

Example 2.3. The set of Hamilton quaternions $\mathbb{H} = \{a + bi + cj + dk \ : \ a, b, c, d \in \mathbb{R}\}$ is a non-commutative $\mathbb{R}$-algebra (though, not a $\mathbb{C}$-algebra). Here multiplication is defined using the relations

$$i^2 = j^2 = k^2 = ijk = -1.$$

Indeed, we will see that $\mathbb{H}$ is a division algebra.

Example 2.4. The set $\mathrm{Mat}_{m\times n}(F)$ of $m \times n$ matrices with entries in F is an F-vector space with entrywise addition. The set of $n \times n$ matrices, here written more simply as $\mathrm{Mat}_n(F)$, is a non-commutative F-algebra with the additional operation of matrix multiplication.

Example 2.5. The ring of polynomials $F[x]$ is a commutative F-algebra.

Example 2.6. The set of all real-valued functions on $\mathbb{R}$, denoted $\mathrm{Func}(\mathbb{R}, \mathbb{R})$, is an $\mathbb{R}$-algebra when equipped with pointwise operations:

$$(f + g)(x) \coloneqq f(x) + g(x)$$

and

$$(f \cdot g)(x) \coloneqq f(x)g(x).$$

Similarly, one might consider functions on a subset $\Omega \subseteq \mathbb{R}^n$ or with values in $\mathbb{C}$; we write $\mathrm{Func}(\Omega, \mathbb{R})$ and $\mathrm{Func}(\Omega, \mathbb{C})$ for these algebras (over $\mathbb{R}$ and, for the latter, also over $\mathbb{C}$).

Example 2.7. If F is a field, the set F^n is an F-vector space with coordinate-wise addition. We follow the convention of writing vectors as columns:

$$x = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in F^n.$$

Definition 2.3. (Bretscher 2013, 170; Petersen 2012, 54) Let V be an F-vector space. A subset $U \subseteq V$ is a **subspace**, denoted $U \leq V$, if it is an F-vector space with the operations from V. Equivalently, a non-empty subset U is a subspace if, for every $x, y \in U$ and $\alpha, \beta \in F$, we have $\alpha x + \beta y \in U$. If V is an F-algebra, a subspace U is a **subalgebra** if it is also a subring.

Example 2.8. Every vector space contains the **trivial subspace** $\{0\}$.

Example 2.9. For $\Omega \subseteq \mathbb{R}^n$, the following are all subalgebras of $\mathrm{Func}(\Omega, \mathbb{R})$:

$$\begin{aligned} C(\Omega) &= \text{all continuous real-valued functions on } \Omega. \\ C^k(\Omega) &= \text{all real-valued functions on } \Omega \text{ with continuous partial derivatives of order } k. \\ C^\infty(\Omega) &= \text{all infinitely-differentiable real-valued functions on } \Omega. \\ \mathcal{P}(\Omega) &= \text{functions on } \Omega \text{ that can be expressed as polynomials in } n \text{ variables.} \end{aligned}$$

We write $C(\Omega, \mathbb{C})$, $C^k(\Omega, \mathbb{C})$, $C^\infty(\Omega, \mathbb{C})$, and $\mathcal{P}(\Omega, \mathbb{C})$ for the analogous $\mathbb{C}$-valued subalgebras.

Proposition 2.1. *(Petersen 2012, 56) Let V be an F-vector space.*

- *If $V_1, V_2 \leq V$, then $V_1 \cap V_2 \leq V$.*
- *If $V_1 \leq V$ and $V_2 \leq V_1$, then $V_2 \leq V$.*

2.1.1 Bases and Dimension

From this point on, all vector spaces are understood with respect to some fixed ground field F unless stated otherwise.

Definition 2.4 (Complementary subspaces)**.** (Petersen 2012, 57) Let V be a vector space. We say that $V_1, V_2 \leq V$ are **complementary** if $V_1 \cap V_2 = \{0\}$ and every vector $x \in V$ can be expressed as $x = y + z$ for some $y \in V_1, z \in V_2$.

Proposition 2.2. *Two subspaces $V_1, V_2 \leq V$ are complementary if and only if each $x \in V$ can be written* uniquely *as $x = y + z$ for $y \in V_1$ and $z \in V_2$.*

Definition 2.5. (Bretscher 2013, 171; Petersen 2012, 56–57) Let V be a vector space. Given $S \subseteq V$, the **span** of S is the smallest subspace of V containing S, i.e., the intersection of all subspaces containing S. Equivalently, the span is the collection of all finite **linear combinations** in S:

$$\operatorname{Span}(S) \coloneqq \left\{ \sum_{i=1}^{k} \alpha_i x_i \;\middle|\; k \in \mathbb{N}, \alpha_i \in F, x_i \in S \right\}. \tag{2.1}$$

Definition 2.6 (Linear independence)**.** (Bretscher 2013, 171; Petersen 2012, 72) Let V be a vector space and $S \subseteq V$ non-empty. We say S is **linearly dependent** if there are $x_1, \dots, x_n \in S$ and $\alpha_1, \dots, \alpha_n \in F$ not all zero with

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0. \tag{2.2}$$

A set $\{x_1, \dots, x_n\}$ is instead said to be **linearly independent** if it is not linearly dependent; that is, the only solution to Equation 2.2 is given by taking all $\alpha_i = 0$.

Definition 2.7 (Bases)**.** (Bretscher 2013, 172; Petersen 2012, 14) Let V be a vector space. A collection of vectors $\mathscr{B} \subset V$ is said to be a *basis* for V if $\mathscr{B}$ is linearly independent and

$$\operatorname{Span}(\mathscr{B}) = V.$$

Often we refer to an *ordered basis*, which is a basis whose vectors are given a specific ordering, and we write $\mathscr{B}$ as a tuple rather than a set. When it will not cause confusion, we still use the notation $\mathscr{B} \subseteq V$ for ordered bases, understanding that $\mathscr{B}$ is a tuple rather than a set.

Theorem 2.1 (Dimension)**.** *(Bretscher 2013, 172; Petersen 2012, 15) Let V be a vector space. All bases of V have the same cardinality, called the* ***dimension*** *of V and written* $\dim_F V$ *(or just* $\dim V$ *if the context is clear). If this cardinality is finite, we say that V is* ***finite-dimensional*** *and write* $\dim V < \infty$*; otherwise, V is* ***infinite-dimensional****.*

Theorem 2.2. *(Bretscher 2013, 172; Petersen 2012, 15) Let V be a vector space with*

$$S = \{x_1, \dots, x_n\} \subset V$$

a linearly independent set. Then every element in $\operatorname{Span}(S)$ *can be expressed* uniquely *as a linear combination of elements in S. In particular, if* $\dim V < \infty$ *and $\mathscr{B}$ is a basis for V, then every $x \in V$ can be expressed uniquely as a linear combination of basis elements.*

Example 2.10. For $V = F^n$, the **standard basis** is given by elements $e_i \in F^n$ with a 1 in the ith entry and 0 elsewhere. The tuple $(e_1, \dots, e_n)$ is an ordered basis for F^n, which is n-dimensional.

Example 2.11. The complexes $\mathbb{C}$ are 2-dimensional over $\mathbb{R}$, with $\{1, i\}$ as a basis.

Example 2.12. The quaternions $\mathbb{H}$ are 4-dimensional over $\mathbb{R}$, with $\{1, i, j, k\}$ as a basis.

Example 2.13. The matrix algebra $\mathrm{Mat}_n(F)$ is n^2-dimensional over F.

Example 2.14. The set $\{x^n \mid n \in \mathbb{N}\}$, where we define $x^0 \coloneqq 1$, is a basis for the polynomial algebra $F[x]$ over F, which is therefore infinite-dimensional.

Example 2.15. The algebras $C(\mathbb{R})$, $C^k(\mathbb{R})$, and $C^\infty(\mathbb{R})$ have infinite dimension over $\mathbb{R}$, each containing $\mathcal{P}(\Omega)$ as a subalgebra. Note that we know $\dim_{\mathbb{R}} \mathcal{P}(\Omega)$ is infinite by a (temporarily informal, cf. Definition 2.9) comparison to $\mathbb{R}[x]$, thinking of polynomials (the formal objects) as the functions defined in terms of polynomials.

Example 2.16. The trivial vector space (Example 2.8) is 0-dimensional over F with the basis $\varnothing$.

2.1.2 Decompositions

Definition 2.8 (Direct Sum)**.** (Petersen 2012, 58) Given a pair of vector spaces V and W over the same field F, we can make the Cartesian product $V \times W$ into a vector space by defining

$$(x, y) + (x', y') \coloneqq (x + x', y + y')$$

and

$$\alpha(x, y) \coloneqq (\alpha x, \alpha y).$$

To avoid such cumbersome notation, we think of elements in $V \times W$ as formal sums $x + y$ instead of pairs (x, y), so the above rules become more aesthetically familiar: e.g., $\alpha(x + y) = \alpha x + \alpha y$. This space is the **direct sum** of V and W and is denoted $V \oplus W$.

Remark 2.1. By construction, every $x \in V \oplus W$ has a unique decomposition $x = y + z$ for $y \in V$ and $z \in W$. That is, V and W can be identified with natural complementary subspaces of $V \oplus W$.

Theorem 2.3. *(Petersen 2012, 58) If $\{v_1, \dots, v_n\} \subset V$ and $\{w_1, \dots, w_m\} \subset W$ are bases, then*

$$\{v_1, \dots, v_n, w_1, \dots, w_m\}$$

is a basis for $V \oplus W$. Hence $\dim(V \oplus W) = \dim V + \dim W$.

Theorem 2.4 (Existence of complements)**.** *(Petersen 2012, 61) Let V be finite-dimensional with basis $\mathscr{B} = \{v_1, \dots, v_n\}$ and $U \leq V$. Then it is possible to choose $\{v_{i_1}, \dots, v_{i_k}\} \subseteq \mathscr{B}$ so that*

$$V = U \oplus \mathrm{Span}\{v_{i_1}, \dots, v_{i_k}\}.$$

Corollary 2.1. *If V is a vector space and $U \leq V$, then* $\dim U \leq \dim V$. *In particular, if V is finite-dimensional, then* $\dim U = \dim V$ *if and only if $U = V$.*

2.1.3 Linear Transformations

Definition 2.9. (Bretscher 2013, 178; Petersen 2012, 20) Let V and W be vector spaces over a field F. Then a **linear transformation** (a.k.a. a linear map, a linear operator, or a homomorphism of F-vector spaces) is a function $L : V \to W$ satisfying

$$L(x + y) = L(x) + L(y)$$

and

$$L(\alpha x) = \alpha L(x)$$

for all $x, y \in V$ and $\alpha \in F$. Equivalently,

$$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y) \text{ for each } x, y \in V, \alpha, \beta \in F.$$

We write $\mathrm{Hom}_F(V, W)$, or simply $\mathrm{Hom}(V, W)$ if the context is clear, for the set of all such maps. If $V = W$, we abbreviate further to $\mathrm{End}(V) \coloneqq \mathrm{Hom}(V, V)$. If L is a bijection, then its inverse is also linear; we call L a **linear isomorphism** and write $V \cong W$.

Remark 2.2. Knowing the values of L on a basis $\mathscr{B} \subset V$ determines L; we often define linear maps by deciding where to send basis elements, then **extending linearly**. In addition, we will often write Lx instead of $L(x)$ to avoid cumbersome parentheses.

Example 2.17. The identity map $V \to V$ is a linear isomorphism. More generally, scaling by $\lambda \in F^\times$ is a linear isomorphism from $V \to V$ (known as a **homothety**) as is rotation about a particular axis through the origin.

Proposition 2.3. *The composition of linear transformations is a linear transformation. Moreover, the composition of linear isomorphisms is a linear isomorphism.*

Example 2.18. Complex conjugation is an $\mathbb{R}$-linear isomorphism $\mathbb{C} \to \mathbb{C}$. Similarly, if V is a field and F is its canonical (a.k.a. prime (Dummit and Foote 2003, 511)) subfield then every element of $\mathrm{Aut}(V)$ is an F-linear isomorphism $V \to V$ (a.k.a., an F-equivariant automorphism).

Example 2.19. The map $\mathrm{D} : C^1([0,1]) \to C([0,1])$ given by $(\mathrm{D}f)(t) \coloneqq \frac{\mathrm{d}f}{\mathrm{d}t}$ is a linear map. Note that D is not injective, since every constant function maps to zero.

Example 2.20. The map $\mathrm{I} : C([0,1]) \to C^1([0,1])$ given by

$$(\mathrm{I}f)(t) = \int_0^t f(s)\,\mathrm{d}x$$

is linear. By the fundamental theorem of calculus, $\mathrm{D} \circ \mathrm{I} = \mathbb{I}$.

2.1.4 Cosets and Quotients

Definition 2.10. (Bretscher 2013, 178; Petersen 2012, 64) If $L : V \to W$ is a linear map, then

$$\ker L \coloneqq \{x \in V : Lx = 0\}$$

is called the **kernel** of L; the **image** of L is defined as

$$\mathrm{im}\, L \coloneqq \{y \in W \mid y = Lx \text{ for some } x \in V\}.$$

Recall that $\ker L \leq V$ and $\mathrm{im}\, L \leq W$ and also that some books call $\ker L$ the **nullspace**. We call $\mathrm{Null}(L) \coloneqq \dim \ker L$ the **nullity** of L and $\mathrm{rank}(L) \coloneqq \dim \mathrm{im}\, L$ the **rank** of L.

Definition 2.11. (Petersen 2012, 67) The map $V_1 \oplus V_2 \to V_1$ which forgets the V_2 term, i.e.,

$$x + y \mapsto x,$$

is linear and called a **projection** onto V_1. Equivalently, a linear map $P : V \to V$ is called a projection if $P^2 = P$, i.e., $Px = x$ for all $x \in \operatorname{im} P$. In this setup, $\mathbb{I} - P : V \to V$ is also a projection such that $V_1 \coloneqq \operatorname{im} P = \ker(\mathbb{I} - P)$, $V_2 \coloneqq \ker P = \operatorname{im}(\mathbb{I} - P)$, and $V = V_1 \oplus V_2$.

Theorem 2.5. *If $L : V \to W$ is a linear transformation and $Lx = y$, then*

$$L^{-1}(y) = x + \ker L \coloneqq \{x + z \mid z \in \ker L\}.$$

*Informally: fibers of linear maps look the same. In particular, L is one-to-one if and only if $\ker L$ is the trivial subspace. We call $x + \ker L$ the **coset** of x with respect to $\ker L$. This construction makes sense if we replace $\ker L$ with any subspace of V, since any subspace can be realized as the kernel of some linear map.*

Example 2.21. Given a differential equation of the form

$$\mathrm{D}f = g,$$

where D is some linear differential operator and g is a given function, a common technique is to first find *all* solutions to the equation $\mathrm{D}f = 0$ (i.e., $\ker D$) and then find a *particular* function f_p so that $\mathrm{D}f_p = g$. Then the space of all solutions is the coset $f_p + \ker \mathrm{D}$.

To see this in action, consider the differential equation

$$f''(t) + f(t) = -\sin t.$$

By the arduous process of educated guesses, one can find a particular solution $f_p(t) = \frac{t}{2}\cos t$. Moreover, by considering the associated characteristic equation, one can show that all real-valued solutions to $f''(t) + f(t) = 0$ are of the form $a\cos t + b\sin t$ for $a, b \in \mathbb{R}$. Hence

$$\left\{\frac{t}{2}\cos t + a\cos t + b\sin t \middle| a, b \in \mathbb{R}\right\}$$

is the set of all solutions to $f''(t) + f(t) = -\sin t$. See, e.g., (Logan 2015, 100) for further reading.

Remark 2.3. Recall the infamous $+C$ that calculus instructors are so fond of. They were generously reminding you that solutions to differential equations are really cosets!

Definition 2.12 (Quotient vector spaces)**.** (Dummit and Foote 2003, 108; Petersen 2012, 108) Let V be an F-vector space with $U \leq V$. Then the set of cosets

$$V/U \coloneqq \{x + U \mid x \in V\}$$

is a (well-defined) F-vector space with the operations

$$(x + U) + (y + U) \coloneqq (x + y) + U$$

and

$$\alpha(x + U) \coloneqq (\alpha x) + U.$$

We call V/U the **quotient** of V with respect to U and $\dim V/U$ the **codimension** of $U \leq V$. If we suppose further that V is finite-dimensional, then we have

$$\dim V/U = \dim V - \dim U,$$

i.e., the codimension of U is the same as the dimension of any complementary subspace to U.

Sketch. If $\{v_1, \dots, v_n\}$ is a basis for V and $V = U \oplus \operatorname{Span}\{v_{i_1}, \dots, v_{i_k}\}$, then show

$$\{v_{i_1} + U, \dots, v_{i_k} + U\}$$

is a basis for V/U. □

Theorem 2.6 (Nöther's isomorphism theorem)**.** *(Dummit and Foote 2003, 412; Petersen 2012, 109) Let $L : V \to W$ be a linear map. Then there is a natural isomorphism*

$$V/\ker L \cong \operatorname{im} L.$$

Moreover, there are inclusion-respecting bijections:

$$\{\textit{subspaces of } V \textit{ containing } \ker L\} \leftrightarrow \{\textit{subspaces of } V/\ker L\} \leftrightarrow \{\textit{subspaces of } \operatorname{im} L\}.$$

Corollary 2.2 (Rank-nullity theorem)**.** *If V is finite-dimensional and $L : V \to W$ is linear, then*

$$\dim L = \operatorname{Null} L + \operatorname{rank} L.$$

Corollary 2.3 (Isomorphism criteria)**.** *Suppose V and W are vector spaces of the same finite dimension and that $L : V \to W$ is a linear transformation. Then the following are equivalent:*

- *L is an isomorphism.*
- $\ker L = \{0\}$.
- $\operatorname{im} L = W$.
- *L sends a basis of V to a basis of W.*

Corollary 2.4. *Two finite dimensional vector spaces over the same field are isomorphic if and only if they have the same dimension.*

2.1.5 Matrices

Definition 2.13. (Bretscher 2013, 187; Petersen 2012, 48) Given a linear transformation $L : V \to W$ and a pair of ordered bases $\mathscr{A} = (v_1, \dots, v_n)$ and $\mathscr{B} = (w_1, \dots, w_m)$ for V and W, respectively, we associate a **matrix** (i.e., an $m \times n$ array of scalars) to L as follows. For each basis element $v_j \in V$, we know the image $Lv_j \in W = \operatorname{Span}(\mathscr{B})$ can be expressed uniquely as a linear combination:

$$Lv_j = \ell_{1,j} w_1 + \ell_{2,j} w_2 + \cdots + \ell_{m,j} w_m,$$

with each $\ell_{i,j} \in F$. For bookkeeping, we arrange these coefficients into columns:

$${}_{\mathscr{B}}[L]_{\mathscr{A}} \coloneqq \begin{pmatrix} \ell_{1,1} & \cdots & \ell_{1,n} \\ \vdots & \ddots & \vdots \\ \ell_{m,1} & \cdots & \ell_{m,n} \end{pmatrix} \in \operatorname{Mat}_{m\times n}(F).$$

When the bases are clear, we might simply write M_L for this matrix.

For an arbitrary $x \in V$ expressed in the basis $\mathscr{A}$, i.e., $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$, we can compute

$$Lx = (\ell_{1,1}\alpha_1 + \cdots + \ell_{1,n}\alpha_n) w_1 + \cdots + (\ell_{m,1}\alpha_1 + \cdots + \ell_{m,n}\alpha_n) w_m.$$

More compactly, writing $(Lx)_i$ for the w_i coefficient of Lx, we have

$$(Lx)_i = \sum_{j=1}^{n} \ell_{i,j}\alpha_j.$$

If we write a column vector $[x]_{\mathscr{A}}$ with the coefficient of v_j in the jth entry, then we can compute $L(x)$ using the standard definition of matrices acting on column vectors:

$$[L(x)]_{\mathscr{B}} = {}_{\mathscr{B}}[L]_{\mathscr{A}}\,[x]_{\mathscr{A}}.$$

Hence, given bases, much of the theory of linear maps can be reduced to that of matrices from $\mathrm{Mat}_{m\times n}(F)$ acting on column vectors from F^n (cf. Example 2.7). For example, the rank of a linear map is equal to the rank of its associated matrix, which can be computed using the concrete techniques of row reduction; see (Bretscher 2013, 26; Petersen 2012, 82).

We will most often be interested in maps $L : V \to V$ with the same basis $\mathscr{B}$ for the domain and the codomain, where we simply write $[L]_{\mathscr{B}}$ for the associated matrix.

Theorem 2.7. *Composition of linear transformations corresponds to matrix multiplication. That is, if $T : U \to V$ and $L : V \to W$ are linear maps with $\mathscr{A}$, $\mathscr{B}$, and $\mathscr{C}$ ordered bases for the vector spaces U, V, and W, respectively, then*

$${}_{\mathscr{C}}[L \circ T]_{\mathscr{A}} = {}_{\mathscr{C}}[L]_{\mathscr{B}}\,{}_{\mathscr{B}}[T]_{\mathscr{A}}.$$

Example 2.22. Consider the operator D from Example 2.46 restricted to polynomials of degree less than $n+1$. Take the ordered basis $(1, x, x^2, \dots, x^n)$. Then D is associated to the matrix

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

We can compute $\frac{\mathrm{d}}{\mathrm{d}x}(x^4 - 3x^2 + 7x + 2) = 4x^3 - 6x + 7$ by way of matrix multiplication:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 7 \\ -3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ -6 \\ 0 \\ 4 \\ 0 \end{pmatrix}.$$

Theorem 2.8. *(Bretscher 2013, 194; Petersen 2012, 48) Let V be finite-dimensional with $L : V \to V$ a linear map and consider two ordered bases $\mathscr{A}$ and $\mathscr{B} = (v_1, \dots, v_n)$. Then*

$$S\,[L]_{\mathscr{B}} = [L]_{\mathscr{A}}\,S$$

and

$$[L]_{\mathscr{B}} = S^{-1}\,[L]_{\mathscr{A}}\,S,$$

*where S is the invertible matrix whose ith column is given by v_i expressed in the basis $\mathscr{A}$. The latter is known as **conjugation** by the matrix S.*

Example 2.23. Consider complex conjugation, as in Example 2.18. The change of basis matrix from $\mathscr{A} = \{1+i, 1-i\}$ to $\mathscr{B} = \{1, i\}$ is given by

$$S = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix},$$

since $\frac{1}{2}(1+i) + \frac{1}{2}(1-i) = 1$ and $\frac{1}{2}(1+i) + \frac{-1}{2}(1-i) = i$. Noting that $S^{-1} = 2S$, we have

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = S^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} S.$$

This is an example of **diagonalizing** a matrix, i.e., changing from a less convenient basis into a more convenient basis, to be reviewed in Section 2.3.2.

Definition 2.14. If V and W are finite-dimensional and $L : V \to V$ and $T : W \to W$ are a pair of linear maps, there is an *induced* map $L \oplus T : V \oplus W \to V \oplus W$ given by

$$(L \oplus T)(x + y) \coloneqq L(x) + T(y)$$

for each $x \in V$ and $y \in W$. Given ordered bases for V and W, so that L and T are expressed by the matrices M_L and M_T, respectively, the map $L \oplus T$ is represented by the **block matrix**

$$\begin{pmatrix} M_L & 0 \\ 0 & M_T \end{pmatrix}.$$

Block matrices are especially important because they visually depict how the associated linear map respects (or fails to respect) a subspace decomposition.

Example 2.24. In the previous sense, a diagonal matrix

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

can be understood as the direct sum of many scaling transformations: $\lambda_1 \oplus \lambda_2 \oplus \cdots \oplus \lambda_n$, which each λ_i acts on the subspace spanned by e_i. In this sense, **diagonalized matrices are nice because they have been decomposed into many** 1**-dimensional linear transformations** (a.k.a. scaling) **that do not interact with one another.**

Definition 2.15. There are many equivalent and useful definitions for the **determinant** of $A \in \mathrm{Mat}_n(F)$, written $\det(A) \in F$. In particular, the determinant is the only continuous map $\phi : \mathrm{Mat}_n(\mathbb{R}) \to \mathbb{R}$ satisfying $\phi(\lambda \mathbb{1}) = \lambda^n$ and

$$\phi(AB) = \phi(A)\phi(B)$$

for all $A, B \in \mathrm{Mat}_n(\mathbb{R})$ and $\lambda \in \mathbb{R}$; there are similar characterizing properties over arbitrary fields. From this one can show, among other things, that $\det(A) = 0$ if and only if A is not invertible and $\det(A^{-1}) = \det(A)^{-1}$ for all invertible matrices.

Another, more geometric way of understanding determinants over $F = \mathbb{R}$ is in terms of the volume of the paralellepiped determined by the columns of the matrix. In other words, the determinant measures how a matrix dilates or contracts unit volume. One can use this formulation to again see that $\det(A) = 0$

if and only if $\{0\} \subset \ker A$, since a matrix with non-trivial kernel must collapse at least one axis in $\mathbb{R}^n$ to zero.

Determinants can be computed in a number of ways, in particular the recursive **Laplace expansion** in terms of a weighted sum of minors (a.k.a. submatrices obtained by deleting one row and one column) or via adjugates. Lastly, for any finite-dimensional vector space V, we can speak of the determinant of linear maps $L : V \to V$ by fixing an ordered basis $\mathscr{A}$, and computing $\det([L]_{\mathscr{A}})$. This is well-defined, since choosing a different basis $\mathscr{B}$ for V amounts to conjugation and therefore

$$\det([L]_{\mathscr{B}}) = \det(S^{-1}\,[L]_{\mathscr{A}}\,S) = \det(S)^{-1}\det([L]_{\mathscr{A}})\det(S) = \det([L]_{\mathscr{A}}).$$

Definition 2.16. (Bretscher 2013, 194; Petersen 2012, 48) The **trace** of a matrix $A \in \mathrm{Mat}_n(F)$ is given by the sum of its diagonal entries, i.e., writing $a_{i,j}$ for the entry of A in the ith row and jth column, $\mathrm{Tr}(A) \coloneqq a_{1,1} + \cdots + a_{n,n}$. The trace satisfies a cyclic invariance property:

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$$

for all $A, B \in \mathrm{Mat}_n(F)$, as well as linearity. Moreover, the trace is the unique linear map $\mathrm{Mat}_n(F) \to F$ satisfying the cyclic invariance property and $\mathrm{Tr}(\mathbb{I}) = n$. The trace of a linear map on a finite-dimensional vector space V is defined by fixing a basis and computing the trace of the associated matrix; as with determinants, this is independent of the basis choice.

Definition 2.17. The **commutator** of two linear operators $A, B : V \to V$ is defined as

$$[A, B] \coloneqq AB - BA.$$

This expression is so named because $AB = BA$ if and only if $[A, B] = 0$.

Definition 2.18. The operator $T : C^\infty([0,1]) \to C^\infty([0,1])$ given by $(Tf)(t) = t\,f(t)$ is linear. This map does not commute with $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}$, since

$$(\mathrm{D}Tf)(t) = f(t) + tf'(t) \neq tf'(t) = (T\mathrm{D}f)(t).$$

Indeed, $[\mathrm{D}, T] = \mathbb{I}$. Physicists like to make a big deal about this sort of thing.

2.2 Hermitian Inner Products

In this and subsequent sections we only consider vector spaces over $F = \mathbb{R}$ or $F = \mathbb{C}$.

Definition 2.19 (Inner product)**.** (Bretscher 2013, 249; Petersen 2012, 209) An **inner product** on any real vector space V is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ satisfying

- **Symmetry.** $\langle x, y \rangle = \langle y, x \rangle$.
- **Linearity in the second argument.** $\langle x, \alpha y + \beta z \rangle = \alpha\langle x, y \rangle + \beta\langle x, z \rangle$.
- **Positive-definiteness.** $\langle x, x \rangle > 0$ for all $x \neq 0$.

Note that combining the first two axioms also gives **linearity in the first argument**.

Remark 2.4. For a vector space V defined over the complex numbers, we cannot hope for these axioms to hold. For example, we would need to have

$$0 < \langle ix, ix \rangle = i^2\langle x, x \rangle = -\langle x, x \rangle < 0,$$

which is a contradiction. However, complex conjugation provides a suitable modification: we know $z\overline{z} = |z|^2 \geq 0$ for all $z \in \mathbb{C}$, with equality if and only if $z = 0$.

Definition 2.20 (Hermitian inner product)**.** (Bretscher 2013, 213; Debnath and Mikusinski 2005, 94) A **Hermitian inner product** on a complex vector space V is any map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ satisfying

- **Skew-symmetry, a.k.a, conjugate symmetry.** $\langle x, y \rangle = \overline{\langle y, x \rangle}$.
- **Linearity in the second argument.** $\langle x, \alpha y + \beta z \rangle = \alpha \langle x, y \rangle + \beta \langle x, z \rangle$.
- **Positive-definiteness.** $\langle x, x \rangle > 0$ for all $x \neq 0$. \end{enumerate} Combining the first two axioms, we see that this pairing is **antilinear** in the first argument:

$$\langle \alpha x + \beta y, z \rangle = \overline{\alpha} \langle x, z \rangle + \overline{\beta} \langle y, z \rangle.$$

Some books take linearity in the first argument and hence antilinearity in the second—this is a (divisive) matter of convention.

Definition 2.21. A vector space V over $\mathbb{R}$ (or $\mathbb{C}$) equipped with an inner product (respectively, a Hermitian inner product) is an **inner product space**. When emphasis is needed, we might refer to V as a *real* (respectively, *complex*) inner product space.

Proposition 2.4. *(Bretscher 2013, 218; Debnath and Mikusinski 2005, 97) Every inner product space has a natural* ***norm*** $V \to \mathbb{R}^{\geq 0}$ *given by*

$$\|x\| \coloneqq \sqrt{\langle x, x \rangle}$$

which satisfies

- ***Absolute homogeneity.*** $\|ax\| = |a| \, \|x\|$.
- ***Triangle inequality.*** $\|x + y\| \leq \|x\| + \|y\|$.
- ***Pythagorean theorem.*** $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ *if* $\langle x, y \rangle = 0$.
- ***Cauchy–Schwarz inequality.*** $\langle x, y \rangle \leq \|x\| \, \|y\|$.
- ***Parallelogram law.*** $\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$.

In other words, inner product spaces have a natural notion of distance as well as angles.

Definition 2.22. If V is an inner product space, $x, y \in V$ are called **orthogonal** if $\langle x, y \rangle = 0$. We say that $x \in V$ is a **unit vector** if $\|x\| = 1$ and we call the process of replacing a nonzero vector $x \in V$ with the unit vector $\frac{x}{\|x\|}$ **normalizing**.

Example 2.25. The vector space $\mathbb{R}^n$ is an inner product space over $\mathbb{R}$ with the usual notion of dot product: $\langle x, y \rangle \coloneqq x^\top y$, where the superscript $^\top$ denotes the **transpose**. Similarly, $\mathbb{C}^n$ is an inner product space over $\mathbb{C}$ with the **Hermitian dot product**: $\langle x, y \rangle \coloneqq x^* y = \overline{x}^\top y$, i.e.,

$$\left\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \right\rangle \coloneqq \overline{\alpha_1} \beta_1 + \cdots + \overline{\alpha_n} \beta_n.$$

2.2.1 Orthogonal Complements

Proposition 2.5. *(Bretscher 2013, 236; Debnath and Mikusinski 2005, 127) If V is an inner product space and $U \leq V$, then*

$$U^\perp \coloneqq \{x \in V : \langle x, y \rangle = 0 \text{ for all } y \in U\}$$

is a subspace of V called the ***orthogonal complement*** *of U. This name is deserved because $V = U \oplus U^\perp$, i.e., every $x \in V$ can be written uniquely as $x = y + z$ for $y \in U$ and $z \in U^\perp$.*

Proposition 2.6. *If V is a finite-dimensional inner product space and $U \leq V$, then*

$$(U^\perp)^\perp = U.$$

Theorem 2.9 (Orthogonal projection)**.** *(Bretscher 2013, 213; Debnath and Mikusinski 2005, 130) Let V be a finite-dimensional inner product space with $U \leq V$. Then there is a unique projection $P : V \to V$ with $\operatorname{im} P = U$ and $\ker P = U^\perp$. We call P the **orthogonal projection** onto U.*

Remark 2.5. In general, we say that a projection P is **orthogonal** if

$$\langle Px, y\rangle = \langle x, Py\rangle,$$

for all $x, y \in V$, i.e., if P is self-adjoint (see Section 2.2.3).

Example 2.26. The matrix $P = \left(\begin{smallmatrix} 0 & 0 \\ a & 1 \end{smallmatrix}\right)$ is a projection onto the second coordinate in $\mathbb{R}^2$ and is orthogonal with respect to the usual dot product if and only if $a = 0$.

Example 2.27. If V is an inner product space and $x_0 \in V$ is a unit vector, then the linear map $P : V \to V$ given by $Px \coloneqq \langle x_0, x\rangle x_0$ is an orthogonal projection onto $U = \operatorname{Span}\{x_0\}$.

2.2.2 Orthonormal Bases

Definition 2.23. Let V be a real or complex inner product space. A collection of vectors $\{v_1, \dots, v_n\} \subset V$ is called **pairwise orthonormal** if

$$\langle v_i, v_j\rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

If $\mathscr{B} \subset V$ is a basis of pairwise orthonormal vectors, $\mathscr{B}$ is an **orthonormal basis** for V.

Example 2.28. The standard bases of $\mathbb{R}^n$ and $\mathbb{C}^n$ (see Example 2.10 and Example 2.25) are orthonormal bases with respect to the usual dot and Hermitian dot products, respectively.

Example 2.29 (L^2 inner products)**.** We can give $C([0, 2\pi])$ an inner product via

$$\langle f, g\rangle \coloneqq \frac{1}{\pi}\int_0^{2\pi} f(t)g(t)\,\mathrm{d}t. \tag{2.3}$$

Then the infinite set $\{\frac{1}{\sqrt{2}}, \sin t, \cos t, \sin(2t), \cos(2t), \sin(3t), \dots\}$ is pairwise orthonormal. If we instead consider the space $C([-\pi, \pi], \mathbb{C})$ with the inner product

$$\langle f, g\rangle \coloneqq \int_{-\pi}^{\pi} \overline{f(t)}g(t)\,\mathrm{d}t,$$

then the set $\left\{\frac{1}{\sqrt{2\pi}}e^{int}\middle| n \in \mathbb{Z}\right\}$ is pairwise orthonormal.

Remark 2.6. Orthonormal bases are convenient in many ways. In particular, to express an arbitrary vector $x \in V$ in terms of an orthonormal basis $\{v_1, \dots, v_n\} \subset V$, one can simply apply inner products. In particular, we have

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

where

$$\alpha_i \coloneqq \langle v_i, x\rangle.$$

This gives an alternative form of the Pythagorean theorem:

$$\|x\|^2 = \langle x, x\rangle = \sqrt{|\alpha_1|^2 + \cdots + |\alpha_n|^2}.$$

Moreover, if $y = \beta_1 v_1 + \cdots + \beta_n v_n \in V$ is another vector, then we have

$$\langle x, y\rangle = \overline{\alpha_1}\beta_1 + \cdots + \overline{\alpha_n}\beta_n.$$

In other words, *orthonormal bases reduce abstract inner products to dot products.*

Theorem 2.10 (Gram–Schmidt process)**.** *(Bretscher 2013, 218; Petersen 2012, 225) Given a basis $\mathscr{B} \subset V$ for V an inner product space, one can always obtain an orthonormal basis via a process of iterated projections. To be specific, given $\{v_1, \dots, v_n\}$ we produce $\{u_1, \dots, u_n\}$ via*

$$u_1 = \frac{v_1}{\|v_1\|},$$
$$u_2 = \frac{v_2 - \langle u_1, v_2\rangle u_1}{\|v_2 - \langle u_1, v_2\rangle u_1\|},$$
$$u_3 = \frac{v_3 - \langle u_1, v_3\rangle u_1 - \langle u_2, v_3\rangle u_2}{\|v_3 - \langle u_1, v_3\rangle u_1 - \langle u_2, v_3\rangle u_2\|},$$

and so on.

2.2.3 Functionals and Adjoints

Definition 2.24. If V is a vector space over F, a **functional** is a linear map $V \to F$.

Example 2.30. The maps $f(t) \mapsto \int_0^1 f(t)\,\mathrm{d}t$ and $f(t) \mapsto f(0)$ are functionals on $C([0,1])$.

Proposition 2.7. *Given functionals $f, h : V \to F$ and $\alpha, \beta \in F$, we define*

$$\begin{aligned} \alpha f + \beta h : V &\to F \\ x &\mapsto \alpha f(x) + \beta h(x). \end{aligned}$$

*This new mapping is also a functional. Moreover, these notions of addition and scaling show that the set V^* of all functionals on V is itself a vector space (called the **dual space** of V).*

Proposition 2.8. *If* $\dim V < \infty$, *then* $\dim V = \dim V^*$. *In particular, if $\{v_1, \dots, v_n\} \subset V$ is a basis, there is a canonical **dual basis** $\{f_1, \dots, f_n\} \subset V^*$, where $f_i(v_i) = 1$ and $f_i(v_j) = 0$ for all $i \neq j$.*

Lemma 2.1. *If $f : V \to F$ is a non-trivial functional on a finite-dimensional vector space V, then* $\operatorname{rank}(f) = 1$. *If V is an inner product space, then* $(\ker f)^\perp \subset V$ *is* 1*-dimensional.*

Theorem 2.11 (Riesz representation)**.** *(Debnath and Mikusinski 2005, 133) If V is a finite-dimensional inner product space and $f : V \to F$ is a functional, there is a unique $x_f \in V$ such that*

$$f(x) = \langle x_f, x\rangle$$

for all $x \in V$. In other words, an inner product gives a canonical correspondence

$$\begin{aligned} \varphi &: V \to V^* \\ (\varphi(y)(x))\, x &= \langle y, x\rangle \end{aligned}$$

that is an isomorphism if $F = \mathbb{R}$ and an anti-isomorphism *if $F = \mathbb{C}$, i.e.,*

$$\varphi(\alpha y) = \overline{\alpha}\varphi(y).$$

Sketch. If $f : V \to F$ is non-trivial, fix a unit vector $y \in (\ker f)^\perp$. The the assignment $x \mapsto \langle y, x\rangle y$ is an orthogonal projection onto $(\ker f)^\perp$, so we can write (c.f. Example 2.27):

$$x = \underbrace{x - \langle y, x\rangle y}_{\in \ker f} + \underbrace{\langle y, x\rangle y}_{\in (\ker f)^\perp}.$$

But this means that

$$f(x) - \langle y, x\rangle f(y) = 0$$

for all $x \in V$, i.e., taking $x_f = \overline{f(y)}y$ gives the desired result. □

Definition 2.25. If V and W are both F-vector spaces, a map $\mu : V \times V \to W$ is said to be a **sesquilinear** if, for all $\alpha, \beta \in F$ and $x, y, z \in V$, we have:

$$\mu(\alpha x + \beta y, z) = \overline{\alpha}\mu(x, z) + \overline{\beta}\mu(y, z)$$

and

$$\mu(x, \alpha y + \beta z) = \alpha\mu(x, y) + \beta\mu(x, z).$$

If $W = F$, μ is a **sesquilinear pairing**. In the $F = \mathbb{R}$ context, μ is simply called **bilinear**.

Example 2.31. An inner product on V is a sesquilinear pairing. If A and B are linear operators on V, then $(x, y) \mapsto \langle Ax, By\rangle$ is a sesquilinear pairing. Similarly, if $f, h : V \to F$ are linear functionals, then $(x, y) \mapsto \overline{f(x)}h(y)$ is sesquilinear.

Theorem 2.12. *If μ is a sesquilinear pairing on a a finite-dimensional inner product space V, then there is a unique linear operator $A_\mu : V \to V$ such that, for all $x, y \in V$, we have*

$$\mu(x, y) = \langle A_\mu x, y\rangle,$$

i.e., any sesquilinear pairing can be given in terms of an operator and the inner product.

Sketch. For a fixed choice of $x \in V$, the map $y \mapsto \mu(x, y)$ is a linear functional. Then, by Riesz representation (Theorem 2.11), there is a unique vector $x_\mu \in V$ such that $\mu(x, y) = \langle x_\mu, y\rangle$ for all $y \in V$. We need to show that the assignment $x \mapsto x_\mu$ defines a linear operator $A_\mu \mid V \to V$. Indeed, for any $\alpha, \beta \in F$ and $x, x', y \in V$ we have

$$\begin{aligned}
\langle A_\mu(\alpha x + \beta x'), y\rangle &= \mu(\alpha x + \beta x', y) \\
&= \overline{\alpha}\mu(x, y) + \overline{\beta}\mu(x', y) \\
&= \overline{\alpha}\langle A_\mu x, y\rangle + \overline{\beta}\langle A_\mu x', y\rangle \\
&= \langle \alpha A_\mu x, y\rangle + \langle \beta A_\mu x', y\rangle.
\end{aligned}$$

That is, $A_\mu(\alpha x + \beta x') = \alpha A_\mu x + \beta A_\mu x'$. □

Remark 2.7. The bra–ket notation so loved by physicists is related to this theorem. We have shown that *any* sesquilinear pairing on, say, $\mathbb{C}^n$ can be given in terms of a matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ and the usual dot product. A physicist might denote such a pairing by $\langle x \mid A \mid y\rangle$.

Definition 2.26 (Adjoint operators)**.** (Debnath and Mikusinski 2005, 158; Petersen 2012, 242) Let L be a linear operator on a finite-dimensional inner product space V. Then the operator $L^* : V \to V$ satisfying

$$\langle x, Ly\rangle = \langle L^* x, y\rangle, \text{ for all } x, y \in V,$$

whose existence and uniqueness is guaranteed by Theorem 2.12, is called the **adjoint** of L.

Proposition 2.9. *If V is a finite-dimensional inner product space and $A, B : V \to V$, then*

$$(AB)^* = B^* A^*.$$

Moreover, $\mathbb{I}^* = \mathbb{I}$, $(A^*)^* = A$, and

$$(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$$

for all $\alpha, \beta \in F$.

Example 2.32 (Transpose)**.** If we equip $\mathbb{R}^n$ with the standard dot product and $A \in \mathrm{Mat}_n(\mathbb{R})$, then the adjoint is given by the matrix transpose: $A^* = A^\mathsf{T}$. This follows because

$$\langle x, Ay\rangle = x^\mathsf{T} A y = (A^\mathsf{T} x)^\mathsf{T} y = \langle A^\mathsf{T} x, y\rangle.$$

Example 2.33 (Conjugate transpose)**.** If we equip $\mathbb{C}^n$ with the Hermitian dot product and $A \in \mathrm{Mat}_n(\mathbb{C})$, then the adjoint is given by the conjugate transpose: $A^* = \overline{A^\mathsf{T}}$.

2.2.4 Self-Adjoint and Unitary Operators

Definition 2.27 (Self-adjoint operators)**.** (Debnath and Mikusinski 2005, 159; Petersen 2012, 265) A linear operator A on a finite-dimensional inner product space V is called **self-adjoint** if $L^* = L$, i.e., if

$$\langle Lx, y\rangle = \langle x, Ly\rangle, \text{ for all } x, y \in V.$$

Especially if $F = \mathbb{R}$, we might also call such operators **symmetric**; if $F = \mathbb{C}$, we call them **Hermitian**. Beware that, in more general (infinite-dimensional) contexts, all three of these terms have subtly distinct meanings: see Debnath and Mikusinski (2005).

Example 2.34. The operator $(Tf)(t) = t\, f(t)$ is (formally)[1] symmetric with respect to the inner product

$$\langle f, g\rangle \coloneqq \int_0^1 f(t) g(t)\, \mathrm{d}t.$$

Example 2.35. A matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ is symmetric (with respect to the dot product) if $A = A^\mathsf{T}$. Similarly, $B \in \mathrm{Mat}_n(\mathbb{C})$ is Hermitian (with respect to the complex dot product) if $B = \overline{B^\mathsf{T}}$ (cf. Example 2.32 and Example 2.33).

Example 2.36. Take V to be the space of smooth functions $f : [0,1] \to \mathbb{C}$ where $f(0) = f(1)$ together with the inner product

$$\langle f, g\rangle \coloneqq \int_0^1 \overline{f(t)} g(t)\, \mathrm{d}t$$

[1] We are being reckless: the associated vector spaces are infinite-dimensional, so defining adjoints takes greater care. See Debnath and Mikusinski (2005) for a more cautious approach.

The "momentum" operator $A : V \to V$ given by $Af(t) = -i\frac{\mathrm{d}f}{\mathrm{d}t}$ is Hermitian. Indeed:

$$\begin{aligned}\langle f, Ag\rangle &= \int_0^1 \overline{f(t)}(-i)g'(t)\,\mathrm{d}t \\ &= \underbrace{-i\overline{f(t)}g(t)\Big|_0^1}_{=0} + \int_0^1 \overline{(-i)f'(t)}g(t)\,\mathrm{d}t \\ &= \langle Af, g\rangle,\end{aligned}$$

after integrating by parts.

Definition 2.28 (Self-adjoint operators)**.** (Debnath and Mikusinski 2005, 167; Petersen 2012, 273) An invertible linear operator L on an inner product space V is called **unitary** if $L^{-1} = L^*$. If $F = \mathbb{R}$, we use the adjective **orthogonal**. In other words, L is unitary if, for all $x, y \in V$, we have

$$\langle Lx, Ly\rangle = \langle x, y\rangle,$$

i.e., that L leaves the inner product invariant. Equivalently, L is unitary if

$$\langle Lx, y\rangle = \langle x, L^{-1}y\rangle.$$

Definition 2.29. A matrix $R \in \mathrm{GL}_n(\mathbb{R})$ is orthogonal (with respect to the dot product) if and only if $R^{-1} = R^{\mathsf{T}}$. Similarly, $Q \in \mathrm{GL}_n(\mathbb{C})$ is unitary (with respect to the complex dot product) if and only if $Q^{-1} = \overline{Q^{\mathsf{T}}}$. This is equivalent to asking that the matrix columns (equivalently, the matrix rows) are orthonormal. (cf. Example 2.32 and Example 2.33).

2.3 Spectral Decomposition

Definition 2.30 (Eigenvectors and eigenvalues)**.** (Bretscher 2013, 310; Petersen 2012, 133) If F is any field and V is an F-vector space with $L : V \to V$ a linear map, we say that a nonzero $x \in V$ is an **eigenvector** of L with **eigenvalue** $\lambda \in F$ if

$$Lx = \lambda x.$$

Example 2.37. Every non-zero $x \in V$ is an eigenvector of $\mathbb{I} : V \to V$ with eigenvalue 1.

Example 2.38. The function $f(t) = e^t$ is an eigenvector of $\mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}$ with eigenvalue 1. More generally, the function $f_\lambda(t) = e^{\lambda t}$ is an eigenvector of D with eigenvalue λ.

Example 2.39. The matrix $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ has no real eigenvectors, yet $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right) = -i\left(\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right)$.

Example 2.40. A diagonal matrix $D = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ has each standard basis element e_i as an eigenvector with eigenvalue λ_i.

Example 2.41. The operator $T : C^\infty([0,1]) \to C^\infty([0,1])$ given by $(Tf)(t) \coloneqq t\, f(t)$ (see Example 2.34) has no eigenvalues.

Lemma 2.2 (Characteristic polynomial)**.** *(Bretscher 2013, 329; Petersen 2012, 367) If L is a linear operator on a vector space V, then $\lambda \in F$ is an eigenvalue of L if and only if $L - \lambda\mathbb{I}$ has a non-trivial kernel. If V is finite-dimensional, the eigenvalues of L correspond to roots of the polynomial*

$$p_L(t) \coloneqq \det(t\mathbb{I} - L) \in F[t],$$

*called the **characteristic polynomial** of L. Note that* $\deg p_L(t) = \dim V$*, so L can have at most n distinct eigenvalues. Moreover, if F is algebraically closed (namely, $F = \mathbb{C}$), then every linear map has at least one eigenvalue.*

Theorem 2.13. *If $A \in \mathrm{Mat}_n(F)$ and $S \in \mathrm{GL}_n(F)$, A and $S^{-1}AS$ have the same eigenvalues.*

Proof. In fact, conjugate matrices *have the same characteristic polynomial*:

$$p_{S^{-1}AS}(\lambda) = \det(\lambda\mathbb{I} - S^{-1}AS) = \det(S^{-1}(\lambda\mathbb{I} - A)S) = \det(\lambda\mathbb{I} - A) = p_A(\lambda).$$

□

Theorem 2.14. *If L is an operator on a finite-dimensional inner product space V, then λ is an eigenvalue of L if and only if $\overline{\lambda}$ is an eigenvalue of L^*.*

Sketch. Note that λ is *not* an eigenvalue of $(\lambda\mathbb{I} - L)$ if and only if

$$(\lambda\mathbb{I} - L)A = \mathbb{I}$$

for some operator $A : V \to V$, i.e., if $\lambda\mathbb{I} - L$ is invertible. Taking adjoints gives

$$A^*(\overline{\lambda}\mathbb{I} - L^*) = \mathbb{I},$$

and so we see that this condition is equivalent to $\overline{\lambda}$ *not* being an eigenvalue of L^*. □

Corollary 2.5. *If $L : V \to V$ is a self-adjoint operator on a finite-dimensional inner product space V, then all eigenvalues of L are real.*

Alternative. If $Lx = \lambda x$ with x nonzero, we see directly that

$$\lambda\|x\|^2 = \langle x, \lambda x\rangle = \langle x, Lx\rangle = \langle Lx, x\rangle = \langle \lambda x, x\rangle = \overline{\lambda}\|x\|^2$$

and hence we must have $\lambda \in \mathbb{R}$. □

Corollary 2.6. *If $L : V \to V$ is a unitary operator on a finite-dimensional inner product space V, then all eigenvalues λ of L are complex numbers with $|\lambda| = 1$.*

Alternative. If $Lx = \lambda x$ with x nonzero, we have

$$|\lambda|^2\|x\|^2 = \langle \lambda x, \lambda x\rangle = \langle Lx, Lx\rangle = \langle x, x\rangle = \|x\|^2$$

and hence we must have $|\lambda| = 1$. □

2.3.1 Diagonalization

Definition 2.31. If a linear operator L on a finite-dimensional vector space V admits a basis $\mathscr{B} = \{v_1, \dots, v_n\}$ of **eigenvectors**, a.k.a. an **eigenbasis**, then $[L]_{\mathscr{B}}$ is a diagonal matrix whose entries are eigenvalues. More precisely, there is a decomposition of V into subspaces,

$$V = \text{Span}\{v_1\} \oplus \cdots \oplus \text{Span}\{v_n\},$$

such that $L = \lambda_1 \oplus \cdots \oplus \lambda_n$, where $Lv_i = \lambda_i v_i$. We say that L is **diagonalizable**.

Definition 2.32. If $L : V \to V$ is an operator and $\lambda \in F$ is an eigenvalue, we write

$$E_\lambda = \{x \in V : Lx = \lambda x\}$$

for the subspace all corresponding eigenvectors, called the **eigenspace** of λ.

Lemma 2.3. *If L is a linear transformation on a vector space V and $\lambda, \lambda' \in F$ are distinct eigenvalues of L, then $E_\lambda \cap E_{\lambda'} = \{0\}$. Moreover, a linear transformation is diagonalizable if and only if its eigenspaces span V.*

Sketch. If $x \in E_\lambda \cap E_{\lambda'}$, then $\lambda x = Lx = \lambda' x$. Thus $(\lambda - \lambda')x = 0$ yet $\lambda \neq \lambda'$, so $x = 0$. □

Corollary 2.7. *If L is a linear transformation on an n-dimensional vector space V and L has n distinct eigenvalues, then L is diagonalizable.*

Example 2.42. The matrix $\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)$ is not diagonalizable for any choice of $\lambda \in F$, since it only has a single eigenvalue, namely λ, and yet $E_\lambda = \text{Span}\{e_1\} < F^2$.

Remark 2.8. In a sense, *most* operators are diagonalizable. Briefly, given a matrix

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} \in \text{Mat}_n(F),$$

the characteristic polynomial has n roots when counted with multiplicity. Based on the theory of resultants, there is a polynomial in the n^2 entries of A which is zero if and only if A has a repeated eigenvalue (in the proper splitting field). The zero sets of non-trivial polynomials are very small compared to all of $\text{Mat}_n(F) = F^{n^2}$; over $F = \mathbb{R}$ or $F = \mathbb{C}$, the set of non-diagonalizable matrices have Lebesgue measure zero. For more, see (Dummit and Foote 2003, 619).

Example 2.43. The matrix $\left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ has distinct eigenvalues so long as $(\alpha - \delta)^2 + 4\beta\gamma \neq 0$.

Theorem 2.15. *Suppose A and B are diagonalizable linear operators on a finite-dimensional vector space V. If $AB = BA$ then the two operators are **simultaneously diagonalizable**, i.e., there is a choice of basis $\mathscr{B} \subset V$ so that $[A]_{\mathscr{B}}$ and $[B]_{\mathscr{B}}$ are both diagonal matrices.*

Proof. Since A is diagonalizable, we have a decomposition $V = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_k}$. Next, given $x \in E_{\lambda_i}$, we have $ABx = BAx = B(\lambda_i x) = \lambda_i Bx$, i.e., Bx is still an eigenvector of A with eigenvalue λ_i. This means that B restricts to an operator B_i on each eigenspace E_{λ_i} of A. Hence, in any eigenbasis $\mathscr{A}$ for A, we have

$$[B]_{\mathscr{A}} = \begin{pmatrix} B_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & B_k \end{pmatrix},$$

a block diagonal matrix. Since B is diagonalizable, we can choose eigenbases $\mathscr{B}_i \subset E_{\lambda_i}$ to diagonalize each B_i; therefore $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_k$ is an eigenbasis for B. Remembering that each element of $\mathscr{B}$ is an eigenvector of A, we have $[A]_{\mathscr{B}}$ and $[B]_{\mathscr{B}}$ both diagonal. □

Example 2.44. The matrices $A = \left(\begin{smallmatrix} -7 & 12 & 12 \\ 2 & -5 & -4 \\ -6 & 12 & 11 \end{smallmatrix}\right)$ and $B = \left(\begin{smallmatrix} 0 & 6 & 2 \\ -2 & 6 & 4 \\ 2 & -3 & -3 \end{smallmatrix}\right)$ commute. Via $S = \left(\begin{smallmatrix} 2 & 2 & 3 \\ 0 & 1 & -1 \\ 1 & 0 & 3 \end{smallmatrix}\right)$, we have

$$S^{-1}AS = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$S^{-1}BS = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

i.e., changing bases to $\left(\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix}\right)$ is sufficient to diagonalize A but not B, yet B has distinct eigenvalues and so must be diagonalizable! Instead, take $R = \left(\begin{smallmatrix} 2 & 4 & 3 \\ 0 & 1 & -1 \\ 1 & 1 & 3 \end{smallmatrix}\right)$:

$$R^{-1}AR = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$R^{-1}BR = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Theorem 2.16. *(Bretscher 2013, 334; Petersen 2012, 368) If L is a linear map on some finite-dimensional inner product space V with $\lambda_1, \ldots, \lambda_n \in F$ the roots of $p_L(\lambda)$ counted with multiplicity, then*

$$\det(L) = \lambda_1 \cdots \lambda_n \quad \textit{and} \quad \operatorname{Tr}(L) = \lambda_1 + \cdots + \lambda_n.$$

Said more simply: the determinant is the product of eigenvalues and the trace is their sum.

Favorite sketch. If $L : V \to V$ is diagonalizable, then we can find a basis $\mathscr{B}$ such that $[L]_{\mathscr{B}}$ is diagonal; the determinant (or trace) of a diagonal matrix is just the product (or sum, respectively) of the diagonal entries. So the result is clear for diagonalizable operators.

Moreover, because "most" operators are diagonalizable in the previously alluded sense (Remark 2.8), even a non-diagonalizable $L : V \to V$ is arbitrarily "close" to a diagonalizable operator—more precisely, the set of diagonalizable operators is **dense** amongst all operators. As the operations of determinant and trace are continuous, the general formulas follow. □

Theorem 2.17. *If $P : V \to V$ is a projection on a finite-dimensional vector space V, then its only possible eigenvalues are* 0 *and* 1*. Moreover,* $\operatorname{Tr}(P)$ *is an integer and is equal to the dimension of the subspace projected onto by P.*

Definition 2.33. A linear transformation L on a finite-dimensional inner product space V is **normal** if it commutes with its adjoint, i.e., $[L, L^*] = 0$.

Example 2.45. All self-adjoint and unitary operators are normal.

Remark 2.9. Self-adjoint maps are analogous to real numbers in the algebra of operators. In fact, any operator L can be written uniquely as the sum of "real" and "imaginary" part, $L = A + Bi$, where A and B are both self-adjoint. This is accomplished by taking

$$A = \frac{1}{2}(L + L^*)$$

and

$$B = \frac{1}{2i}(L - L^*).$$

Moreover, one can check that $[L^*, L] = 2i[A, B]$, i.e., L is normal if and only if its real and imaginary parts commute. Thus, if we think of self-adjoint operators as behaving like real numbers, with normal transformations roughly analogous to complex numbers. We will later see that unitary operators behave like complex numbers of modulus 1.

2.3.2 Spectral Theory

Lemma 2.4. *Let L be a normal operator on an inner product space V. Then $x \in V$ is an eigenvector of L with eigenvalue $\lambda \in F$, i.e., $Lx = \lambda x$, if and only if $L^*x = \overline{\lambda}x$.*

Proof. First we show that L being normal means that L and L^* have the same nullspace:

$$\|Lx\|^2 = \langle Lx, Lx\rangle = \langle L^*Lx, x\rangle = \langle LL^*x, x\rangle = \langle L^*x, L^*x\rangle = \|L^*x\|^2,$$

so the left hand side is zero if and only if the right hand side is. Moreover, $L - \lambda\mathbb{I}$ is normal:

$$\begin{aligned}(L - \lambda\mathbb{I})(L - \lambda\mathbb{I})^* &= (L - \lambda\mathbb{I})(L^* - \overline{\lambda}\mathbb{I}) \\ &= LL^* - \lambda L^* - \overline{\lambda}L + |\lambda|^2\mathbb{I} \\ &= L^*L - \overline{\lambda}L - \lambda L^* + |\lambda|^2\mathbb{I} \\ &= (L^* - \overline{\lambda}\mathbb{I})(L - \lambda\mathbb{I}) = (L - \lambda\mathbb{I})^*(L - \lambda\mathbb{I}).\end{aligned}$$

Thus $(L - \lambda\mathbb{I})x = 0$ if and only if $(L^* - \overline{\lambda}\mathbb{I})x = 0$, i.e., $Lx = \lambda x$ if and only if $L^*x = \overline{\lambda}x$. □

Theorem 2.18. *If L is a normal operator on an inner product space V with $\lambda, \lambda' \in F$ distinct eigenvalues of L, then the eigenspaces E_λ and $E_{\lambda'}$ are orthogonal, i.e., we have $\langle x, y\rangle = 0$ for all $x \in E_\lambda$ and $y \in E_{\lambda'}$.*

Proof. For any $x \in E_\lambda$ and $y \in E_{\lambda'}$, we have

$$\lambda'\langle x, y\rangle = \langle x, \lambda' y\rangle = \langle x, Ly\rangle = \langle L^*x, y\rangle = \langle \overline{\lambda}x, y\rangle = \lambda\langle x, y\rangle.$$

Since $(\lambda - \lambda')\langle x, y\rangle = 0$ and $\lambda \neq \lambda'$, we must always have $\langle x, y\rangle = 0$. □

Theorem 2.19 (Spectral theorem)**.** *(Debnath and Mikusinski 2005, 196; Petersen 2012, 275) If L is a linear map on a finite-dimensional inner product space V, then L is normal if and only if there is an orthonormal basis $\{v_1, \dots, v_n\} \subset V$ consisting of eigenvectors of L, i.e., $Lv_i = \lambda_i v_i$. Thus, for any $x \in V$,*

$$Lx = \sum_{i=1}^{n} \lambda_i \langle v_i, x\rangle v_i.$$

Corollary 2.8. *If $A \in \mathrm{Mat}_n(F)$ is self-adjoint (symmetric if $F = \mathbb{R}$) or unitary (orthogonal if $F = \mathbb{R}$), then there is a unitary matrix Q whose columns are eigenvectors of A with*

$$\overline{Q^\top} A Q = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$$

Example 2.46. Consider the "momentum" operator $A = -i\frac{\mathrm{d}}{\mathrm{d}t}$ (Example 2.36) on the vector space

$$V = \mathrm{Span}\{1, \cos t, \sin t\}$$

with the inner product from Equation 2.3. In the given basis,

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}$$

is evidently self-adjoint. We compute $p_A(\lambda) = \lambda^3 - \lambda$, which has the roots $\lambda = 0, 1, -1$; these are the eigenvalues of A. To find the eigenvectors, we compute the nullspace of

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, A - \mathbb{I} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & -i \\ 0 & i & -1 \end{pmatrix}, \text{ and } \quad A + \mathbb{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -i \\ 0 & i & 1 \end{pmatrix}.$$

We can do this via the application of echelon form (Bretscher 2013, 116) or by inspection:

$$\ker A = \mathrm{Span}\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\}, \ker(A - \mathbb{I}) = \mathrm{Span}\left\{\begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}\right\}, \text{ and } \ker(A + \mathbb{I}) = \mathrm{Span}\left\{\begin{pmatrix} 0 \\ 1 \\ -i \end{pmatrix}\right\}.$$

So, the eigenvectors of A are 1, $\cos t + i \sin t = e^{it}$, and $\cos t - i \sin t = e^{-it}$, which can be normalized to the eigenbasis

$$\mathscr{B} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} e^{it}, \frac{1}{\sqrt{2}} e^{-it}\right).$$

If we wish to compute the result of A applied to $x = 3 + 2\cos t$, we can first use inner products to write x in this eigenbasis:

$$x = \underbrace{\left(\frac{1}{\pi}\int_0^{2\pi} x(s)\frac{1}{\sqrt{2}}\,\mathrm{d}s\right)}_{3\sqrt{2}}\frac{1}{\sqrt{2}} + \underbrace{\left(\frac{1}{\pi}\int_0^{2\pi} x(s)\frac{e^{-is}}{\sqrt{2}}\,\mathrm{d}s\right)}_{\sqrt{2}}\frac{e^{it}}{\sqrt{2}} + \underbrace{\left(\frac{1}{\pi}\int_0^{2\pi} x(s)\frac{e^{is}}{\sqrt{2}}\,\mathrm{d}s\right)}_{\sqrt{2}}\frac{e^{-it}}{\sqrt{2}}$$

Hence we compute

$$\begin{aligned} Ax &= 3\sqrt{2}A(\tfrac{1}{\sqrt{2}}) + \sqrt{2}A(\tfrac{1}{\sqrt{2}}e^{it}) + \sqrt{2}A(\tfrac{1}{\sqrt{2}}e^{-it}) \\ &= 3\sqrt{2}\cdot 0\cdot(\tfrac{1}{\sqrt{2}}) + \sqrt{2}\cdot 1\cdot(\tfrac{1}{\sqrt{2}}e^{it}) + \sqrt{2}\cdot(-1)\cdot(\tfrac{1}{\sqrt{2}}e^{-it}) \\ &= e^{it} - e^{-it} = 2i\sin t. \end{aligned}$$

Applying A is easy when x is written in the eigenbasis—simply scaling the individual eigenvectors by their eigenvalues—because it is *diagonalized* by this basis: $[A]_{\mathscr{B}} = \mathrm{Diag}(0, 1, -1)$. Compare this with Remark 2.6.

Chapter 3

Group Theory

The concept of a group is central to abstract algebra for its ubiquity throughout mathematics and the physical sciences: wherever one finds symmetry, one will find groups. Our primary reference for this brief review is Dummit and Foote (2003). For alternate perspectives, consider Shahriar (2017) and Artin's well-known expositions of group theory Artin (2011). As always, the best book is the third one you read!

Many of the groups we study come from an existing additive or multiplicative context, or perhaps via abstract symmetries (that is, self-bijections preserving a desired property) of some object. Indeed, this is the origin of group theory. When studying an abstract group G with some unspecified group operation $*$, aka the **group law** of G, we will often use a multiplicative notation: given $g, h \in G$, we write gh as a shorthand for $g * h$, g^{-1} for the inverse of g with respect to $*$, and g^n to mean

$$\overbrace{g * \cdots * g}^{n \text{ times}} \text{ if } n > 0,$$

or

$$\overbrace{g^{-1} * \cdots * g^{-1}}^{-n \text{ times}} \text{ if } n < 0.$$

If we have some fixed group where an additive notation is more appropriate—usually some abelian group—then we might instead write $g + h$, $-g$, and ng, respectively, for these shorthands.

3.1 Notable Families

Example 3.1 (Additive groups)**.** (Dummit and Foote 2003, 8) The sets $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathrm{Mat}_n(\mathbb{R})$ are abelian groups with respect to addition. Furthermore, the set of **integers modulo** n—the set of equivalence classes of $\mathbb{Z}$ under the relation $a \sim b$ whenever $n|(b-a)$, which is usually written $\mathbb{Z}/n\mathbb{Z}$—is a group with respect to addition. Colloquially, this is the clock with n tickmarks.

Example 3.2 (Group of units)**.** Let M be a set equipped with a unital, associative binary operation $*$. Then the subset of invertible elements,

$$M^\times \coloneqq \{x \in M \mid x \text{ is invertible}\},$$

is a group with respect to $*$ known as the **group of units**. In particular, we write

$$\begin{aligned}\mathbb{Z}^\times &= \{1, -1\},\\ \mathbb{Q}^\times &= \mathbb{Q} \setminus \{0\},\\ \mathbb{R}^\times &= \mathbb{R} \setminus \{0\},\\ \mathbb{C}^\times &= \mathbb{C} \setminus \{0\}.\end{aligned}$$

for the groups with respect to multiplication.

Example 3.3 (Group of nth roots)**.** (Bretscher 2013, 365) Let $n \in \mathbb{N}$ be fixed. The set $\mathcal{C}_n \coloneqq \{z \in \mathbb{C} : z^n = 1\}$ is the group of **nth roots of unity** with respect to multiplication. We can list these elements explicitly using Euler's formula:

$$\mathcal{C}_n = \left\{1, e^{2\pi i/n}, \dots, e^{2\pi i(n-1)/n}\right\}.$$

Example 3.4 (Trivial group)**.** The set $\{1\}$, often simply denoted 1 or $\mathcal{C}_1$, is called the **trivial group**. Notice that $\mathcal{C}_1$ trivally satisfies the group axioms when we define $1 \cdot 1 = 1$.

Example 3.5 (Torus group)**.** The unit circle, written as the set

$$\mathbb{T} \coloneqq \{a + bi \in \mathbb{C} : a^2 + b^2 = 1\}$$

of length 1 complex numbers, is the **circle** (a.k.a., torus) group with respect to multiplication.

Example 3.6 (Group of invertible matrices)**.** The **general linear group** of $n \times n$ matrices with entries in a field F, written $\mathrm{GL}_n(F)$, is the group of invertible elements in $\mathrm{Mat}_n(F)$. That is,

$$\begin{aligned}\mathrm{GL}_n(F) &\coloneqq \mathrm{Mat}_n(F)^\times\\ &= \{M \in \mathrm{Mat}_n(F) \mid M \text{ is invertible}\}\\ &= \{M \in \mathrm{Mat}_n(F) \mid \det(M) \neq 0\}\end{aligned}$$

is a group with respect to matrix multiplication. In general, if V is an F-vector space, then

$$\mathrm{GL}(V) \coloneqq \{A : V \to V \mid A \text{ is a linear isomorphism}\}$$

is the group of **linear automorphisms**. Note that $\mathrm{GL}_n(F)$ is a non-abelian for all $n > 1$; otherwise, $\mathrm{GL}_1(F) = F^\times$. Of additional interest is

$$\mathrm{U}(n) \coloneqq \left\{Q \in \mathrm{GL}_n(\mathbb{C}) \mid Q^{-1} = \overline{Q^\top}\right\},$$

the n-dimensional **unitary group**. We emphasize that $\mathrm{U}(1) = \mathbb{T}$, the torus group.

Example 3.7 (Permutation group)**.** (Dummit and Foote 2003, 29) If X is a set, then $\mathrm{Perm}(X)$ is used to denote the **group of permutations** (i.e., all self-bijections on X) under function composition. If $X = \{1, 2, \dots, n\}$ for some $n \in \mathbb{Z}^+$, we write $\mathcal{S}_n$ and call this group the **symmetric group** on n letters.

To describe permutations, we use **cycle notation**. For instance,

$$(1\ 2\ 4)(3\ 6) \in \mathcal{S}_6$$

denotes the permutation with $1 \mapsto 2$, $2 \mapsto 4$, $4 \mapsto 1$, $3 \mapsto 6$, $6 \mapsto 3$, and 5 fixed.

Example 3.8 (Quaternion group)**.** (Dummit and Foote 2003, 36) The set

$$Q_8 \coloneqq \{1, -1, i, -i, j, -j, k, -k\} \subset \mathbb{H}$$

can be made into a group by defining

$$k = ij \text{ and } -1 = i^2 = j^2 = k^2 = ijk.$$

Here we formally understand -1 as central (that is, an element commuting with all others) and an element squaring to 1. Note that $i^{-1} = -i$, and similarly $j^{-1} = -j$ and $k^{-1} = -k$. In particular, the quaternion group is non-abelian, since we must have $ij = -ji$:

$$\begin{aligned} -1 = k^2 = (ij)(ij) &\Rightarrow iji = -j = j^{-1} \\ &\Rightarrow ij = (iji)i^{-1} = j^{-1}i^{-1} = ji. \end{aligned}$$

Example 3.9. If G and H are groups, with operations $*$ and $\star$, respectively, then the set $G \times H$ can be equipped with a group operation $\odot$ as follows:

$$(g_1, h_1) \odot (g_2, h_2) \coloneqq (g_1 * g_2, h_1 \star h_2).$$

Often we suppress all the group operations and simply write

$$(g_1, h_1)(g_2, h_2) = (g_1 g_2, h_1 h_2).$$

Example 3.10. The group $\mathcal{V} \coloneqq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ is called the **Klein 4-group**.

3.2 Fundamentals

Definition 3.1. (Dummit and Foote 2003, 20) Let G be a group with $g \in G$. The **order** of g, written $|g|$, is the smallest $n \in \mathbb{Z}^+$ such that $g^n = e$; if no such n exists, we say g is of **infinite order**.

Example 3.11. The element $1 \in \mathbb{Z}$ has infinite order, but $1 \in \mathbb{Z}/n\mathbb{Z}$ has order n. Remember, the group operation here is addition!

Example 3.12. The element $i \in \mathcal{C}_4$ has order 4. The element $e^i \in \mathbb{T}$ has infinite order.

3.2.1 Subgroups

Definition 3.2. (Dummit and Foote 2003, 46) Let G be a group with $H \subseteq G$. We say that H is a **subgroup** of G, written $H \leq G$, if H is also a group with respect to the operation of G. If we want to emphasize that H is a **proper subgroup**, i.e., $H \neq G$, then we write $H < G$.

Example 3.13. The set of multiples $n\mathbb{Z}$ is a subgroup of $\mathbb{Z}$ for all $n \in \mathbb{N}$.

Example 3.14. Any group G contains the **trivial subgroup** $\{e\} \leq G$.

Example 3.15. The set $\{\mathbb{1}, (1\ 2\ 3), (1\ 3\ 2)\}$ is a subgroup of $\mathcal{S}_3$.

Example 3.16. There is a chain of subgroups $\mathbb{Z} < \mathbb{Q} < \mathbb{R} < \mathbb{C}$ with respect to addition. Similarly, there is a chain $\{\pm 1\} = \mathbb{Z}^\times < \mathbb{Q}^\times < \mathbb{R}^\times < \mathbb{C}^\times$ with respect to multiplication.

Example 3.17. For any group G and fixed $g \in G$, the set

$$\langle g \rangle \coloneqq \{g^n \mid n \in \mathbb{Z}\}$$

is a subgroup of G of order $|g|$.

Theorem 3.1. *(Dummit and Foote 2003, 62) If $H, K \leq G$, then $H \cap K \leq G$. More generally, if $\{H_\lambda\}_{\lambda \in \Lambda}$ is an arbitrary collection of subgroups in G, then $\bigcap_{\lambda \in \Lambda} H_\lambda \leq G$*

Theorem 3.2 (Lagrange)**.** *(Dummit and Foote 2003, 89) Let G be a finite group. If $H \leq G$, then $|H|$ divides $|G|$.*

Corollary 3.1. *If G is finite and $g \in G$, then $|g|$ divides $|G|$. In particular, $g^{|G|} = e$.*

Corollary 3.2 (Fermat's little theorem)**.** *Fix a prime p. Then $a^p = a$ for all $a \in \mathbb{Z}/p\mathbb{Z}$.*

3.2.2 Conjugacy

Theorem 3.3. *(Dummit and Foote 2003, 50) Let G be a group. The **center** of G, written $\mathbf{Z}(G)$, is*

$$\mathbf{Z}(G) \coloneqq \{h \in G \mid gh = hg \text{ for all } g \in G\} \leq G.$$

That is, $\mathbf{Z}(G)$ consists of the elements that commute with every element in G.

Example 3.18. For all $n > 2$, we have $\mathbf{Z}(\mathcal{S}_n) = \{\mathbb{I}\}$.

Definition 3.3. (Dummit and Foote 2003, 123) If G is a group with $g, h \in G$, we say that ghg^{-1} is the **conjugate** of h by g. Conjugacy defines an equivalence relation on G:

$$g \sim g' \text{ means } g' = hgh^{-1} \text{ for some } h \in G.$$

The equivalence classes of this relation are called **conjugacy classes** and are denoted

$$\mathrm{cl}_G(g) \coloneqq \{hgh^{-1} \mid h \in G\},$$

or simply $\mathrm{cl}(g)$ when the group is clear from context.

Example 3.19. If G is an abelian group, then each element is only conjugate to itself; more generally, for every element $g \in \mathbf{Z}(G)$ we have $\mathrm{cl}_G(g) = \{g\}$.

Example 3.20. In light of the formula

$$\sigma(1\ 2\ \cdots\ k)\sigma^{-1} = (\sigma(1)\ \sigma(2)\ \cdots\ \sigma(k)),$$

the conjugacy classes of $\mathcal{S}_n$ are in bijective correspondence with partitions of n.

Example 3.21. In the group $\mathrm{GL}_n(F)$, conjugation is simply **change of basis** (this is usually written slightly differently, cf. Theorem 2.8, where we replace a matrix A with $S^{-1}AS$ and the columns of S represent the new basis). In particular, diagonalizable matrices are those that are conjugate to a diagonal matrix.

Proposition 3.1. *For any group G, any conjugate of $g \in G$ has the same order as g itself.*

3.2.3 Generators and Relations

Definition 3.4. (Dummit and Foote 2003, 26) Let G be a group. A (possibly infinite) collection of elements $S \subset G$ is said to **generate** G, written $G = \langle S \rangle$, if every element of G can be written as a finite product of elements in S and their inverses. Alternatively but equivalently, define

$$\langle S \rangle \coloneqq \bigcap_{\substack{H \leq G \\ S \subseteq H}} H$$

as the **subgroup generated by** S. From this definition, it is evident by inspection that $\langle S \rangle$ is the *smallest subgroup of G containing S*. If G can be generated by a single element, $G = \langle g \rangle$, then we say that G is **cyclic**.

Example 3.22. The set $\{1\}$ generates the integers $\mathbb{Z}$ (for this reason, the group $\mathbb{Z}$ is often called the **infinite cyclic group}**. Generating sets are not unique; for example, the sets $\{-1\}$ and $\{2, 9\}$ also generate $\mathbb{Z}$. Similarly, $\{1\}$ generates $\mathbb{Z}/n\mathbb{Z}$. On the other hand, $e^{2\pi i/n}$ generates $\mathcal{C}_n$.

Example 3.23. The elements $\{i, j\}$ generate the quaternion group $\mathcal{Q}_8$.

Example 3.24. The Klein 4-group $\mathcal{V}$ is the smallest non-cyclic group.

Definition 3.5. (Dummit and Foote 2003, 218) Given a generating set S of a group G, equations satisfied by the generators are called **relations**. If $R_1, \dots, R_n$ is some finite list of relations from which the group law G can be completely determined, along with the standing assumptions of associativity and identity and inverses, we call the generators and relations a **presentation** for G and write

$$G = \langle S \mid R_1, \dots, R_m \rangle.$$

Example 3.25. We can present $\mathcal{C}_n$ as $\langle \zeta \mid \zeta^n = 1 \rangle$. We might imagine $\zeta = e^{2\pi i/n}$—though this identification is not necessary, and potentially misleading!

Example 3.26. A presentation for the quaternion group is

$$\mathcal{Q}_8 = \langle i, j \mid i^4 = 1, i^2 = j^2, jij = i \rangle.$$

Note that we make no mention of $k \in \mathcal{Q}_8$; it is simply understood as the product ij. More generally, the **quaternion group of order** $4n$, denoted $\mathcal{Q}_{4n}$, has the presentation

$$\mathcal{Q}_{4n} = \langle a, j \mid a^{2n} = 1, a^n = j^2, j^{-1}aj = a^{-1} \rangle.$$

In terms of the quaternion algebra $\mathbb{H}$, we can think of $a = e^{\pi i/n}$.

Example 3.27. The Klein 4-group $\mathcal{V}$ can be presented as $\langle a, b \mid a^2 = b^2 = (ab)^2 = e \rangle$.

Example 3.28. (Dummit and Foote 2003, 23) The **dihedral group** $\mathcal{D}_{2n}$, alternatively understood as the group of $2n$ rigid symmetries for a regular n-gon,[1] can be presented as

$$\langle r, s \mid r^n = \mathbb{I}, s^2 = \mathbb{I}, rs = sr^{-1} \rangle.$$

When thinking of $\mathcal{D}_{2n}$ as acting on a regular n-gon with its vertices labeled $1, \dots, n$ in a clockwise fashion, we will use the following convention:

[1] Some references, especially with geometric inclinations (e.g. Shurman (1997)), write $\mathcal{D}_n$ for this group—preferring to emphasize the number of vertices rather than the number of elements. We opt for the $2n$ notation for consistency with common sources.

- r is a rotation clockwise by $\frac{2\pi}{n}$
- s is a flip across the axis between the center of the n-gon and the vertex labeled 1.

One can show that

$$\mathbf{Z}(\mathcal{D}_{2n}) = \begin{cases} \{e\} & n \text{ odd} \\ \{e, r^{n/2}\} & n \text{ even.} \end{cases}$$

3.3 Homomorphisms

3.3.1 Basic Properties

Definition 3.6. (Dummit and Foote 2003, 36) Let G and H be groups with the operations $*$ and $\star$, respectively. A function $\rho : G \to H$ is called a (group) **homomorphism** if

$$\rho(g * g') = \rho(g) \star \rho(g') \text{ for all } g, g' \in G.$$

If ρ is also a bijection[2], then ρ is called an **isomorphism** and we write $G \cong H$ to denote the existence of such a ρ (pronounced "G is **isomorphic** to H.")

Proposition 3.2. *If $\rho : G \to H$ is a homomorphism, then $\rho(e_G) = e_H$ and $\rho(g^{-1}) = \rho(g)^{-1}$ for all $g \in G$. Moreover, if $\rho : G \to H$ is a bijection, then $\rho^{-1} : H \to G$ is also a homomorphism.*

Proposition 3.3. *If $\rho : G \to K$ and $\sigma : K \to H$ are homomorphisms, then $\sigma \circ \rho : G \to H$ is also a homomorphism. If ρ and σ are isomorphisms, then so is their composition.*

Example 3.29. For any groups G and H, there is always a homomorphism $\rho : G \to H$ given by $\rho(g) = e_H$ for all $g \in G$. This is called the **trivial homomorphism.**

Example 3.30. The groups $\mathcal{D}_6$ and $\mathcal{S}_3$ are isomorphic. One can write down an isomorphism between them by considering the action of $\mathcal{D}_6$ on the vertices of an equilateral triangle as in Example 3.28. The induced homomorphism is given by

$$\begin{aligned} e &\mapsto \mathbb{I}, \\ r &\mapsto (1\ 2\ 3), \\ r^2 &\mapsto (1\ 3\ 2), \\ s &\mapsto (2\ 3), \\ sr &\mapsto (1\ 3), \\ sr^2 &\mapsto (1\ 2). \end{aligned}$$

Theorem 3.4. *If $\rho : G \to H$ is a homomorphism and $g \in G$ has $|g| < \infty$, then $|\rho(g)|$ divides $|g|$.*

Corollary 3.3. *If $\rho : G \to H$ is an isomorphism, then $|g| = |\rho(g)|$ for all $g \in G$.*

Example 3.31. There is a homomorphism $\rho : \mathbb{T} \to \mathbb{T}$ given by $z \mapsto z^2$ which, geometrically, wraps the circle around itself twice. Note that $i \in \mathbb{T}$ has order 4 but $\rho(i) = -1$ has order 2.

Example 3.32. If V is an n-dimensional F-vector space, fixing a basis $\mathscr{B} \subset V$ gives rise to a group isomorphism $\mathrm{GL}(V) \cong \mathrm{GL}_n(F)$ by sending $L \mapsto {}_{\mathscr{B}}[L]_{\mathscr{B}}$.

[2] As we will see, it is not generally correct to define an "isomorphism" as a "bijective homomorphism." The correct definition should be: "bijective homomorphism whose inverse is also a homomorphism." We are fortunate in many algebraic categories that these notions are equivalent.

Example 3.33 (Permutation representation)**.** For any $n \in \mathbb{Z}^+$, there is a map $\mathcal{S}_n \to \mathrm{GL}_n(\mathbb{Z})$ given by sending a permutation σ to the corresponding permutation matrix. For example, when $n = 3$, this homomorphism assigns

$$(1\ 2) \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } (1\ 2\ 3) \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Example 3.34 (Determinants)**.** For any field F and $n \in \mathbb{Z}^+$, there is a homomorphism $\mathrm{GL}_n(F) \to F^\times = F - \{0\}$ given by the determinant, $A \mapsto \det(A)$. This is a rebranding of the important fact from linear algebra (cf. Definition 2.15):

$$\det(AB) = \det(A)\det(B).$$

More generally, if R is any commutative ring, we can write $\mathrm{GL}_n(R)$ for the invertible $n \times n$ with entries in R and write down a determinant homomorphism $\mathrm{GL}_n(R) \to R^\times$.

Definition 3.7. If $\rho : G \to H$ is a group homomorphism, then

$$\ker \rho \coloneqq \{g \in G \mid \rho(g) = e_H\} \leq G$$

is called the **kernel** of ρ; the **image** of ρ is defined as:

$$\operatorname{im} \rho \coloneqq \{h \in H : h = \rho(g) \text{ for some } g \in G\} \leq H.$$

Example 3.35. The kernel of the determinant map is called $\mathrm{GL}_n(F)$, the **special linear group**.

Example 3.36. (Dummit and Foote 2003, 107) There is a map $\varepsilon : \mathcal{S}_n \to \{\pm 1\}$ obtained by composing the permutation representation with the matrix determinant. For example, when $n = 3$:

$$(1\ 2) \mapsto \det \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = -1 \quad \text{and} \quad (1\ 2\ 3) \mapsto \det \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = 1.$$

The kernel of this homomorphism is called the **alternating group**, written $\mathcal{A}_n \leq \mathcal{S}_n$. The elements of $\mathcal{A}_n$ are called **even** and those in $\mathcal{S}_n \setminus \mathcal{A}_n$ are called **odd**.

3.3.2 Isomorphisms

Theorem 3.5. *A homomorphism $\rho : G \to H$ is one-to-one if and only if $\ker \rho = \{e_G\}$. Moreover, one can show that if $\rho(g) = h$, then the fiber $\rho^{-1}(h)$ can be described as the set*

$$g \ker \rho \coloneqq \{gk \mid k \in \ker \rho\}.$$

This is the (left) ***coset*** *(see Section 3.3.3) of g with respect $\ker \rho$.*

This Theorem should be reminiscent of Theorem 2.5.

Remark 3.1. If $\rho : G \to H$ is a homomorphism and $B \leq G$, then $\rho(B)$ is a subgroup of $\operatorname{im} \rho$ in H. That is, the **image** of a subgroup is a subgroup. On the other hand, if $C \leq H$, then the **inverse image** $\rho^{-1}(C)$ is a subgroup of G containing $\ker \rho$.

Example 3.37. Consider the map $\rho : \mathcal{D}_8 \to \mathcal{V}$ given by

$$\begin{aligned}\rho(e) = \rho(r^2) &= (0,0),\\ \rho(r) = \rho(r^3) &= (1,0),\\ \rho(s) = \rho(sr^2) &= (0,1),\\ \rho(sr) = \rho(sr^3) &= (1,1).\end{aligned}$$

The subgroup $\langle r\rangle$ maps to the subgroup $\langle(1,0)\rangle$; the inverse image of $\langle(0,1)\rangle$ is $\langle s, r^2\rangle$.

Remark 3.2. When defining a homomorphism $\rho : G \to H$, where $G = \langle S \mid R_1, \dots, R_m\rangle$, one need only define ρ on the generating set S. This is simply because *every* $g \in G$ *can be built with elements in* S, and so the homomorphism already "knows" how to evaluate on each g. To check that ρ is well-defined, we verify that the relations R_i are preserved by ρ.

Example 3.38. If we try to build a homomorphism $\rho : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}$ by defining $\rho(1) \coloneqq 1$, we will find that such a map is not **well-defined**, i.e., it cannot exist. Indeed, the relation

$$\underbrace{1 + \cdots + 1}_{n} = 0$$

from $\mathbb{Z}/n\mathbb{Z}$ becomes the (false!) statement $n = 0$ over the integers. From this reasoning, we can see that the only homomorphism $\mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}$ is the trivial homomorphism.

Example 3.39. We can build a homomorphism $\rho : \mathcal{D}_{12} \to \mathbb{Z}/2\mathbb{Z}$ by defining

$$\begin{aligned}\rho(r) &\coloneqq 0,\\ \rho(s) &\coloneqq 1.\end{aligned}$$

If we wish to compute, say, $\rho(sr^3)$, we can apply the homomorphism law:

$$\rho(sr^3) = \rho(s) + \rho(r^3) = \rho(s) + 3\rho(r) = 1 + 0 = 1.$$

For well-defined-ness, we check that $r^6 = e$, $s^2 = e$, and $rs = sr^{-1}$ hold after applying ρ:

$$\begin{aligned}\rho(r^6) = 6\rho(r) = 0 = \rho(e)\\ \rho(s^2) = 2\rho(s) = 2 = 0 = \rho(e)\\ \rho(rs) = \rho(r) + \rho(s) = 0 + 1 = 1 = 1 - 0 = \rho(s) - \rho(r) = \rho(sr^{-1}).\end{aligned}$$

Example 3.40 ($\mathbb{Z}$ is free)**.** Given a group G and element $g \in G$, there is a unique homomorphism $\mathbb{Z} \to G$ given by $1 \mapsto g$ and hence $n \mapsto g^n$ for all n. The image of this homomorphism is $\langle g\rangle$. We do not need to check if this map is well-defined because there are no relations on $1 \in \mathbb{Z}$ to check.

Example 3.41 (Finite cyclic groups)**.** The groups $\mathbb{Z}/n\mathbb{Z}$ and $\mathcal{C}_n$ are isomorphic; the former is sometimes called the *additive* cyclic group of order n while the latter is the *multiplicative* cyclic group of order n. We can give an explicit isomorphism by sending $1 \mapsto e^{2\pi i/n}$ and hence $k \mapsto e^{2\pi ik/n}$ for all k; note that this map is well-defined since the relation in $\mathbb{Z}/n\mathbb{Z}$,

$$\underbrace{1 + \cdots + 1}_{n \text{ times}} = 0,$$

passes to the (true) relation $\underbrace{e^{2\pi i/n} \cdot \ldots \cdot e^{2\pi i/n}}_{n \text{ times}} = 1$ in $\mathcal{C}_n$.

Remark 3.3. If $\rho : G \to H$ is a one-to-one map, then $G \cong \operatorname{im}\rho$. In such a context, we might even say that G *is* a subgroup of H by identifying G with its image, by which we understand ρ as an **inclusion**.

Example 3.42. We can identify $\mathcal{C}_n < \mathcal{D}_{2n}$ by the map $e^{2\pi i/n} \mapsto r$. Similarly, we can identify $\mathcal{D}_{2n} \leq \mathcal{S}_n$ by the action of $\mathcal{D}_{2n}$ on the vertices of a regular n-gon. We can even identify

$$\{\mathbb{I}\} = \mathcal{S}_1 < \mathcal{S}_2 < \mathcal{S}_3 < \cdots < \mathcal{S}_n < \mathcal{S}_{n+1} < \ldots$$

by thinking of a permutation $\sigma \in \mathcal{S}_n$ as acting trivially on any number $k > n$. Lastly, one can identify $\mathcal{V} \leq \mathcal{S}_4$ by using, for instance, the homomorphism $(1,0) \mapsto (1\ 2)$ and $(0,1) \mapsto (3\ 4)$.

Remark 3.4 (Non-isomorphism criteria). Given two groups G and H, one might want to prove they are *not* isomorphic. In this case, the following criteria can come in handy. G and H *cannot* be isomorphic if:

- G and H have different cardinalities.
- One of G and H is abelian but the other is not.
- For some fixed $n \in \mathbb{Z}^+$, G and H contain a different number of elements of order n.
- G and H have different subgroup lattices.

These criteria are listed by their difficulty to assess. Beware: there exist non-isomorphic groups which satisfy *all* these criteria and have to be discerned by more sophisticated techniques. We shall establish additional criteria throughout this course.

Definition 3.8. There is an equivalence relation on the collection (category) of all groups:

$$G \sim H \text{ means that there is an isomorphism } G \xrightarrow{\cong} H.$$

The equivalence classes of this relation are called the **isomorphism types**. A foundational problem in group theory is to distinguish groups, i.e., to determine whether or not a given pair of groups belong to the same isomorphism type. For small $|G| = n$, we can write out representatives of these isomorphism types as follows:

n	1	2	3	4	5	6	7	8	9	10	11	12
	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$	$\mathcal{C}_5$	$\mathcal{C}_6$	$\mathcal{C}_7$	$\mathcal{C}_8$	$\mathcal{C}_9$	$\mathcal{C}_{10}$	$\mathcal{C}_{11}$	$\mathcal{C}_{12}$
				$\mathcal{V}$		$\mathcal{S}_3$		$\mathcal{C}_4 \times \mathcal{C}_2$	$\mathcal{C}_3 \times \mathcal{C}_3$	$\mathcal{D}_{10}$		$\mathcal{C}_6 \times \mathcal{C}_2$
								$\mathcal{V} \times \mathcal{C}_2$				$\mathcal{A}_4$
								$\mathcal{D}_8$				$\mathcal{D}_{12}$
								$\mathcal{Q}_8$				$\mathcal{Q}_{12}$

Given a group G of order 6, one could use this table to see that G is isomorphic to either $\mathcal{C}_6$ or $\mathcal{S}_3$ (which are not isomorphic to one another). An easy way to tell which of these G shares isomorphism types with is by checking whether or not G is abelian, or whether G contains an element of order 6.

3.3.3 Cosets and Quotients

Definition 3.9. (Dummit and Foote 2003, 77) Let G be a group with $H \leq G$ and $g \in G$. We define

$$gH \coloneqq \{gh \mid h \in H\} \quad \text{and} \quad Hg \coloneqq \{hg \mid h \in H\}$$

to be the **left** and **right cosets**, respectively, of g with respect to H. We write G/H for the set of left cosets. The **index**[3] of H in G, denoted $[G : H]$, is the cardinality $|G/H|$. The Lagrange theorem (Dummit and

[3] Index is analogous to the notion of codimension (cf. Definition 2.12)

Foote 2003, 89) says that

$$|G| = [G : H] \ |H|$$

and its proof relies on the fact that cosets are all the same size (cf. Theorem 3.5). We also will often refer to the **conjugation** of H by g:

$$gHg^{-1} \coloneqq \{ghg^{-1} \mid h \in H\}.$$

Note that, in general, cosets are *not* subgroups (if $g \notin H$, then gH cannot contain the identify element). However, the conjugations of H are still subgroups of G.

Example 3.43. Consider $H = \langle s \rangle$ in $\mathcal{D}_8$. We can list some left and right cosets:

$$\begin{aligned} eH = \{e, s\} = He = sH = Hs, \\ rH = \{r, sr^3\}, \\ Hr = \{r, sr\}. \end{aligned}$$

We can also consider $N = \langle r^2 \rangle$, for which we have

$$\begin{aligned} eN = \{e, r^2\} = Ne = r^2H = Nr^2, \\ rN = \{r, r^3\} = Nr = r^3N = Nr^3, \\ sN = \{s, sr^2\} = Ns = sr^2N = Nsr^2, \\ srN = \{sr, sr^3\} = Nsr = sr^3N = Nsr^3. \end{aligned}$$

Proposition 3.4. *If $H \leq G$ and $g \in G$, then $gH = H$ if and only if $g \in H$.*

Definition 3.10. (Dummit and Foote 2003, 82) Let G be a group with $N \leq G$. We say that N is **normal** in G and write $N \trianglelefteq G$ if $gN = Ng$ for all $g \in G$, i.e., all the left cosets of N equal the right cosets.

Remark 3.5. (Dummit and Foote 2003, 82) There are many equivalent notions of normality. For example, a subgroup N is normal if and only if $gNg^{-1} = N$ for all $g \in G$. Often the formulation that is easiest to check is $N \trianglelefteq G$ if and only if $gxg^{-1} \in N$ for every $x \in N, g \in G$. Moreover, $N \leq G$ is normal if and only if N is the kernel of some homomorphism $G \to H$.

Example 3.44. In Example 3.43, we saw that $rH \neq Hr$ and hence the subgroup $H = \langle s \rangle$ cannot be normal in $\mathcal{D}_8$. However, $N = \langle r^2 \rangle$ *is* a normal subgroup of $\mathcal{D}_8$!

Proposition 3.5. *If $\rho : G \to H$ is a homomorphism, then* $\ker \rho \trianglelefteq G$.

Definition 3.11 (Quotient groups)**.** If $N \trianglelefteq G$, then the set of cosets G/N is a group with the operation

$$(gN)(hN) \coloneqq (gh)N.$$

We call G/N the **quotient group** of N by G. In words, G/N is what happens if we *collapse the subgroup N in G to a single element*.

Remark 3.6. Because of how the group operation in the quotient is defined, when manipulating elements we often deal in terms of **representatives** (any $g' \in gN$ is a representative for the coset gN). Thus, we should always be careful that our calculations are **well-defined**.

Remark 3.7. In the quotient G/N, every $x \in N$ represents the identity coset.

Example 3.45. Every G contains at least the normal subgroups $G \trianglelefteq G$ and $\{e_G\} \trianglelefteq G$. Notice that $G/G = \{eG\}$ is isomorphic to the trivial group $\mathcal{C}_1$, since we have collapsed all of G to a single point. On the other hand, $G/\{e_G\}$ is isomorphic to G because we did not do any collapsing at all.

Theorem 3.6 (Nöther's isomorphism theorem)**.** *(Dummit and Foote 2003, 97) If $\rho : G \to H$ is a group homomorphism and $K = \ker \rho$, there is an induced isomorphism $G/K \cong \operatorname{im} \rho$ along with inclusion-respecting bijections:*

$$\{\textit{subgroups of } G \textit{ containing } K\} \leftrightarrow \{\textit{subgroups of } G/K\} \leftrightarrow \{\textit{subgroups of } \operatorname{im} \rho\}.$$

Example 3.46. The homomorphism $\rho : \mathbb{R} \to \mathbb{T}$ given by $\rho(t) = e^{2\pi i t}$ has $\ker \rho = \mathbb{Z}$. The quotient $\mathbb{R}/\mathbb{Z}$ is called **the real line modulo** 1 and behaves algebraically like a circle: notice that

$$\left(\tfrac{1}{2} + \mathbb{Z}\right) + \left(\tfrac{1}{2} + \mathbb{Z}\right) = 1 + \mathbb{Z} = 0 + \mathbb{Z}.$$

This reflects the induced isomorphism $\mathbb{R}/\mathbb{Z} \cong \mathbb{T}$ given by $t + \mathbb{Z} \mapsto e^{2\pi i t}$.

Example 3.47. In Example 3.44, we saw that $N = \langle r^2 \rangle \trianglelefteq \mathcal{D}_8$. The quotient

$$\mathcal{D}_8/N = \{eN, rN, sN, srN\}$$

is a group since N is normal. Moreover, $\mathcal{D}_8/N$ is isomorphic to the Klein 4-group via $\rho : \mathcal{D}_8/N \to \mathcal{V}$ defined by

$$\begin{aligned} \rho(eN) &= (0,0), \\ \rho(rN) &= (1,0), \\ \rho(sN) &= (0,1), \\ \rho(srN) &= (1,1). \end{aligned}$$

Example 3.48. (Dummit and Foote 2003, 82) Given $N \trianglelefteq G$, one may consider the **canonical homomorphism** $G \to G/N$ given by $g \mapsto gN$. This homomorphism is surjective and has kernel N. In particular, a subgroup of G is normal if and only if it is the kernel of some homomorphism.

Remark 3.8 (Sub/quotient-object duality)*.* Throughout algebra, one-to-one maps $X \to Y$ correspond to subobjects of Y; surjective maps $X \to Y$ correspond to quotients of X.

3.4 Group Actions

Remark 3.9. Since their official inception in early 19th century work on solving equations, groups were made to *act*—the original definition of group was in terms of the permutations of roots of polynomials that preserve algebraic relations over $\mathbb{Q}$ (i.e., subgroups of S_n). This is reconciled with our modern definition via the Cayley theorem (Dummit and Foote 2003, 120).

3.4.1 Basic Properties

Definition 3.12. (Dummit and Foote 2003, 41) Let G be a group and X a set. A **group action**, sometimes denoted by the shorthand $G \curvearrowright X$, is a homomorphism $\rho : G \to \operatorname{Perm}(X)$. We often suppress ρ and simply write $g \cdot x$ to mean

$$g \cdot x \coloneqq \rho(g)(x)$$

Many books refer to a set X equipped with such an action a G-**set**.

Example 3.49. For any group G and set X, there is always the **trivial action** given by the trivial homomorphism $G \to \operatorname{Perm}(X)$, i.e., $g \cdot x = x$ for all $g \in G$ and $x \in X$.

Example 3.50. (Dummit and Foote 2003, 92) The group $\mathcal{A}_4$ acts on a regular tetrahedron by rotation: given a labeling of the vertices, there is exactly one even permutation corresponding to the action of each rigid symmetry on those vertices. Moreover, this correspondence sends compositions of rotations to composition of permutations.

Example 3.51 (Left multiplication)**.** A group G always acts on itself by left multiplication:

$$\begin{aligned} G &\to \mathrm{Perm}(G) \\ g &\mapsto (x \mapsto gx). \end{aligned}$$

Example 3.52 (Conjugation)**.** A group G always acts on itself by conjugation:

$$\begin{aligned} G &\to \mathrm{Perm}(G) \\ g &\mapsto (x \mapsto gxg^{-1}). \end{aligned}$$

Example 3.53 (Left cosets)**.** Given any $H \leq G$, the group G acts on G/H by left multiplication:

$$\begin{aligned} G &\to \mathrm{Perm}(G/H) \\ g &\mapsto (xH \mapsto (gx)H). \end{aligned}$$

3.4.2 Orbit–Stabilizer Theorem

Definition 3.13. If $G \curvearrowright X$ and $x \in X$, then the **orbit** Gx of x is the set of all elements $y \in Y$ that can be reached from x via the G-action. More precisely,

$$Gx \coloneqq \{y \in X \mid y = g \cdot x \text{ for some } g \in G\}.$$

Orbits are equivalence classes for the relation

$$x \sim y \text{ means } y = g \cdot x \text{ for some } g \in G.$$

In particular, the orbits partition X.

Example 3.54. Consider the circle group $\mathbb{T}$ acting on $\mathbb{C}$ by multiplication. The orbits of this action can be visualized as circles centered at the origin of any positive radius r, along with the origin itself as a singleton orbit.

Definition 3.14. If $G \curvearrowright X$ via $\rho : G \to \mathrm{Perm}(X)$, then the **kernel** of the action is the subgroup of elements in G that act trivially on X. In symbols:

$$\ker(G \curvearrowright X) \coloneqq \ker(\rho) = \{g \in G \mid g \cdot x = x \text{ for all } x \in X\} \trianglelefteq G.$$

For a fixed $y \in X$, we can identify the elements that leave specifically y unchanged:

$$\mathrm{Stab}_G(y) \coloneqq \{g \in G \mid g \cdot y = y\}.$$

This subgroup is the **stabilizer** of y.[4] Note that $\ker(G \curvearrowright X) \leq \mathrm{Stab}_G(y) \leq G$ and

$$\ker(G \curvearrowright X) = \bigcap_{x \in X} \mathrm{Stab}_G(x).$$

In addition, the **fixed set** of $G \curvearrowright X$ is given by

$$X^G \coloneqq \{x \in X \mid g \cdot x = x \text{ for all } g \in G\}.$$

[4] Some books, like Dummit and Foote (2003), write G_x for $\mathrm{Stab}_G(x)$, but this notation is nearly indistinguishable from the orbit Gx.

Remark 3.10. Often when considering a group action $G \curvearrowright X$, one might **restrict** to a subgroup $H \leq G$ by considering only how the elements of H act on X. That is, we have an associated action $H \curvearrowright X$. In this case, for an element $x \in X$, we have $\mathrm{Stab}_H(x) = \mathrm{Stab}_G(x) \cap H$.

Remark 3.11. Another common notation for the fixed set of $G \curvearrowright X$ is $\mathrm{Fix}(G) \coloneqq X^G$; more generally, if $H \leq G$, we write

$$\mathrm{Fix}(H) = X^H \coloneqq \{x \in X \mid h \cdot x = x \text{ for all } h \in H\},$$

and, for a specific element $g \in G$, we write simply $\mathrm{Fix}(g) \coloneqq \mathrm{Fix}(\langle g \rangle)$.

Example 3.55. When $G \curvearrowright X$ trivially, the kernel of the action is all of G.

Example 3.56. Consider the action of $\mathcal{D}_8$ on the two *diagonals* of a square, so that r and sr both act by interchanging the diagonals (recall our convention in Example 3.28). We can see that the kernel of this action is $K = \langle r^2, s \rangle$, since rotation by 180° and flipping across a diagonal axis do not interchange the diagonals.

Example 3.57. Consider the natural action of $\mathcal{S}_n$ on the set $X = \{1, \dots, n\}$, which has trivial kernel. For any $x \in X$, we can see that

$$\mathrm{Stab}_{\mathcal{S}_n}(x) = \{\sigma \in \mathcal{S}_n \mid \sigma(x) = x\} \cong \mathcal{S}_{n-1}.$$

Example 3.58. When $G \curvearrowright G$ by left multiplication, the stabilizer of every $x \in G$ is trivial in light of the **cancellation law**: $gx = gy$ always means $x = y$. Moreover, the orbit of any $x \in G$ is all of G, since $y = (yx^{-1}) \cdot x$ for any $y \in G$. Actions with only one orbit are called **transitive**.

Example 3.59. When $G \curvearrowright G$ by conjugation, the stabilizer of $x \in G$ is called the **centralizer**:

$$\mathrm{C}_G(x) \coloneqq \{g \in G \mid gxg^{-1} = x\}.$$

Remember that the kernel is equal to the intersection of all stabilizers, i.e.,

$$\mathbf{Z}(G) = \bigcap_{g \in G} \mathrm{C}_G(g).$$

Lastly, the orbit of x is the conjugacy class $\mathrm{cl}_G(x)$.

Theorem 3.7. *Suppose $G \curvearrowright X$ and fix $g \in G$ and $x \in X$. Then*

$$\mathrm{Stab}_G(g \cdot x) = g \,\mathrm{Stab}_G(x) g^{-1}.$$

Example 3.60. When $G \curvearrowright G/H$ by left multiplication, the stabilizer of the identity coset eH is simply H. This action is transitive and, in light of Theorem 3.7,

$$\mathrm{Stab}_G(gH) = gHg^{-1}.$$

Hence the kernel of this action is given by

$$\ker(G \curvearrowright G/H) = \bigcap_{g \in G} gHg^{-1}.$$

This is called the **normal core** of H; it is the largest normal subgroup of G contained in H.

Theorem 3.8 (Orbit-stabilizer). *If $G \circlearrowright X$ and $x \in G$, then*

$$[G : \mathrm{Stab}_G(x)] = |Gx|.$$

Example 3.61. For $G \circlearrowright G$ by left multiplication (Example 3.51), each stabilizer is trivial and so Theorem 3.8 gives the unimpressive tautology $|G| = |G|$.

Example 3.62. For $G \circlearrowright G$ by conjugation (Example 3.52), applying Theorem 3.8 yields:

$$[G : \mathrm{C}_G(x)] = |\,\mathrm{cl}_G(x)\,|.$$

Example 3.63. For $G \circlearrowright G/H$ by left multiplication (Example 3.53), Theorem 3.8 to $eH \in G/H$ recovers Theorem 3.2.

3.4.3 Group Algebra

Definition 3.15. Fix a field F and a finite group G. Of great importance to this course is the **group algebra**, denoted $F[G]$, which is a $|G|$-dimensional F-algebra. The group algebra comes with a canonical basis, one element for every $g \in G$ which we denote $e_g \in F[G]$. We define multiplication in this ring by

$$e_g \cdot e_h \coloneqq e_{gh}.$$

In other words, $F[G]$ is built from formal F-linear sums of the elements in $G = \{g_1, \ldots, g_n\}$,

$$\alpha_{g_1} e_{g_1} + \cdots + \alpha_{g_n} e_{g_n} \in F[G],$$

and inherits a notion of multiplication from the group law. In particular, note that $F[G]$ is a commutative F-algebra if and only if G is abelian. The group algebra comes with a natural G action, namely by extending $g \cdot e_h \coloneqq e_{gh}$ linearly.

Example 3.64. The group algebra $\mathbb{R}[\mathbb{Z}/3\mathbb{Z}]$ consists of vectors of the form $\alpha_0 e_0 + \alpha_1 e_1 + \alpha_2 e_2$. We can write down the action $\mathbb{Z}/3\mathbb{Z} \to \mathrm{Aut}(\mathbb{R}[\mathbb{Z}/3\mathbb{Z}])$ explicitly:

$$0 \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$1 \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$
$$2 \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

These matrices all commute, as they are in the image of a homomorphism from an abelian group, so we hope that they might diagonalize simultaneously. Indeed, writing $\zeta = e^{2\pi i/3}$, we have:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta^2 & 0 \\ 0 & 0 & \zeta \end{pmatrix}$$

after changing to the basis $\{e_0 + e_1 + e_2, e_0 + \zeta^2 e_1 + \zeta e_2, e_0 + \zeta e_1 + \zeta^2 e_2\}$. Over $\mathbb{R}$, however, the best we can do is change to a basis like $\{e_0 + e_1 + e_2, e_0 - e_1, e_0 + e_1 - 2e_2\}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & -3/2 \\ 0 & 1/2 & -1/2 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 3/2 \\ 0 & -1/2 & -1/2 \end{pmatrix}$$

While not diagonal, these transformations have all been placed into a block diagonal form. Hence, we have obtained a vector space decomposition $\mathbb{R}[\mathbb{Z}/3\mathbb{Z}] = U \oplus V$, where

$$U = \text{Span}\{e_0 + e_1 + e_2\} \text{ and } V = \text{Span}\{e_0 - e_1, e_0 + e_1 - 2e_2\},$$

that *respects* the action of $\mathbb{Z}/3\mathbb{Z}$; furthermore, we can see that the action of $\mathbb{Z}/3\mathbb{Z}$ on U is trivial and the action on V resembles rotations of the plane by increments of 120°. Breaking actions down into more comprehensible pieces is the *raison d'être* of representation theory.

Part II

Representation Theory

Chapter 4

Introduction

From a certain point of view, the basic premise of representation theory is that groups and their actions are complicated, yet there are certain families of groups (like matrix groups) that act on objects (like vector spaces) which we understand quite well. One can ask: how can we bring to bear the concrete insights of linear algebra to our favorite abstract groups?

4.1 Representations

Definition 4.1. Let G be a group. A **representation** of a group G in a vector space V is a homomorphism

$$\rho : G \to \mathrm{GL}(V).$$

We say that $\dim V$ is the **degree** of the representation; in these notes we will only consider finite-degree representations unless explicitly noted.

Remark 4.1. There is a natural inclusion $\mathrm{GL}(V) \hookrightarrow \mathrm{Perm}(V)$, where the latter is understood as the set of all[1] (that is, paying no heed to linear structure) bijections $V \to V$. In this sense, every representation of G is a group action (recall Definition 3.12) on V.

We refer to representations as *linear group actions* $G \circlearrowright V$: we associate to every $g \in G$ a linear map (or maybe even a matrix!) $\rho(g)$ that knows how to act on any $x \in V$. When the context is clear, we will often use the notation $g \cdot x$ to stand for $\rho(g)(x)$; similarly, we often refer to V as a representation of G without explicitly invoking $\rho : G \to \mathrm{GL}(V)$. Further, if V is finite dimensional over a field F with a chosen ordered basis $(v_1, \dots, v_n)$, then such a map is equivalent to a homomorphism $G \to \mathrm{GL}_n(F)$.

Example 4.1. Every G admits the **trivial representation** via the trivial map $G \to F^\times$.

Example 4.2. The group $\mathbb{Z}/n\mathbb{Z}$ has a family of representations $\mathbb{Z}/n\mathbb{Z} \to \mathbb{C}^\times$ given by

$$1 \mapsto e^{(2\pi j/n)i}.$$

These maps are distinct for $0 \le j < n$; we will denote $\mathbb{C}$ equipped with this action by $U_{j,n}$.

[1] Note that, simply from the perspective of cardinalities, $\mathrm{Perm}(V)$ is *much* larger than $\mathrm{GL}(V)$.

Example 4.3. The group $\mathbb{Z}$ has uncountably many representations $\mathbb{Z} \to \mathbb{C}^\times$. Given any $\theta \in \mathbb{R}$,

$$1 \mapsto e^{\theta i}$$

extended to a homomorphism (recall Example 3.40). Note that each choice of $\theta \in [0, 2\pi)$ gives rise to a distinct homomorphism.

Example 4.4. The circle group has countably many representations $\omega_k : \mathbb{T} \to \mathbb{C}^\times$ defined by

$$\omega_k(z) = z^k.$$

These maps are distinct for every choice of $k \in \mathbb{Z}$; geometrically, they correspond to winding a circle around itself k times.

Example 4.5. The coordinate-wise action (Example 3.33) of $\mathcal{S}_n$ on F^n, i.e., by sending permutations to their corresponding permutation matrices, is called the **permutation representation**.

Example 4.6. The map $\varepsilon : \mathcal{S}_n \to \mathcal{C}_2$ which sends 2-cycles to -1 is the **sign representation** (Example 3.36). The dihedral group admits a similar representation via pre-composing with $\mathcal{D}_{2n} \hookrightarrow \mathcal{S}_n$ (Example 3.42), i.e., recording the permutations induced on vertices, that detects changes in orientation.

Example 4.7. The dihedral group $\mathcal{D}_{2n}$ acts naturally on $\mathbb{R}^2$ via

$$r \mapsto \begin{pmatrix} \cos(\frac{2\pi}{n}) & -\sin(\frac{2\pi}{n}) \\ \sin(\frac{2\pi}{n}) & \cos(\frac{2\pi}{n}) \end{pmatrix},$$
$$s \mapsto \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Taking determinants $\mathrm{Mat}_{2\times 2}(\mathbb{R}) \to \mathbb{R}$ recovers the orientation-detecting representation.

Example 4.8. The group algebra $\mathbb{C}[G]$ with the action described in Definition 3.15 is called the **group algebra** of G.

Example 4.9. For the case $G = \mathcal{D}_6$, using the ordered basis $(e_e, e_r, e_{r^2}, e_s, e_{sr}, e_{sr^2})$, we have:

$$e \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad s \mapsto \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

$$r \mapsto \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \qquad sr \mapsto \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$r^2 \mapsto \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \qquad sr^2 \mapsto \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Compare the block structure of these matrices with that of Example 3.64.

Example 4.10. Suppose that group G acts on a finite set $X = \{x_1, \dots, x_n\}$. Then there is an induced representation of G on $F[X]$, the free F-vector space on X, which is a vector space where elements of X are thought of as formal vectors. Moreover, $F[x]$ comes equipped with a canonical basis, up to ordering: $\{\vec{x}_1, \dots, \vec{x}_n\}$. In particular, if we fix an ordering on the elements of G, then the action on the group algebra $\mathbb{C}[G]$ is is encoded by a homomorphism $\rho : G \to \mathrm{GL}_{|G|}(\mathbb{C})$ and $\operatorname{im}\rho$ consists of permutation matrices.

Example 4.11. Consider the action of $\mathcal{D}_8$ on the diagonals $X = \{d_1, d_2\}$ of a square (Example 3.56); r and s both interchange these diagonals, but r^2 leaves them fixed. This action induces a linear action on formal sums $\alpha\vec{d}_1 + \beta\vec{d}_2$, where $\alpha, \beta \in F$, which constitute the vector space $F[X]$. In the basis $(\vec{d}_1, \vec{d}_2)$, we can encode this action via

$$r \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$s \mapsto \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

4.2 Subrepresentations

When considering a new type of algebraic object, there are a number of important questions we should keep in mind. Among them are: "What sorts of internal structure do these things have?" Just as vector spaces admit subspaces, groups have subgroups, rings have subrings, etc.—what should we mean when we talk about **subrepresentations**?

We have many options! For example, we could take a subgroup $H \leq G$ and then only consider the subset of automorphisms $\rho(h)$ for $h \in H$. This is not the notion we are really after, however, and is referred to as the **restriction** of ρ to H. We will instead orient our analysis around a fixed group G and its different

actions across many vector spaces: G is constant and V, together with its automorphisms compatible with G via some homomorphism, becomes the object of study. One might even call this a category

Remember, a representation is a special type of action on vector spaces, and these have a natural notions of subobject. A subrepresentation, then, should consist of the same automorphisms $\rho(g)$, *but only acting on a particular subspace of* V. Not any subspace will do, however, since for many subspaces W and $g \in G$ we might have $w \in W$ mapping to $g \cdot w \notin W$; that is, $\rho(g)$ is an automorphism of V but not necessarily of an arbitrary subspace W. However, if W is **closed under the G-action**,[2] meaning that $g \cdot w \in W$ for every $w \in W$ and $g \in G$, then it makes sense to think of W as a G-representation in its own right.

Definition 4.2. Let G be a group and $\rho : G \to \mathrm{GL}(V)$ a representation. If W is a subspace of V closed under the G-action, then we say that W is a **subrepresentation** of V. We denote this relationship by writing $W \leq_G V$. If V has admits a proper subrepresentation, we say V is **reducible**; otherwise, V is **irreducible**.[3]

Example 4.12. Every 1-dimensional representation is irreducible, simply because such a space admits no proper non-trivial subspaces.

Example 4.13. Recall the direct sum of linear maps (Definition 2.14). Given a pair of representations $\rho : G \to \mathrm{GL}(V)$ and $\sigma : G \to \mathrm{GL}(W)$, there is an induced representation on $V \oplus W$:

$$\begin{aligned} \rho \oplus \sigma : G &\to \mathrm{GL}(V \oplus W) \\ g &\mapsto \rho(g) \oplus \sigma(g). \end{aligned}$$

In particular, given bases for V and W, the matrices for $\rho \oplus \sigma$ are of the form

$$(\rho \oplus \sigma)(g) = \begin{pmatrix} \rho(g) & 0 \\ 0 & \sigma(g) \end{pmatrix},$$

i.e., $V \oplus W$ contains V and W as (complementary) subrepresentations. If $a \in \mathbb{N}$, we write

$$V^{\oplus a} := \underbrace{V \oplus \cdots \oplus V}_{a \text{ times}}$$

for the a-fold direct sum of V with itself.

Example 4.14. Recall the induced representation from the action of $\mathcal{D}_8$ on the diagonals of a square (Example 4.11). The subspaces $U = \mathrm{Span}\{\vec{d}_1 + \vec{d}_2\}$ and $U' = \mathrm{Span}\{\vec{d}_1 - \vec{d}_2\}$ are subrepresentations. Moreover, if we take $x \in U$ and $y \in U'$, then we see that:

$$r \cdot x = x = s \cdot x \quad \text{and} \quad r \cdot y = -y = s \cdot y.$$

If we listen carefully, the math is trying to tell us that we are studying the representation using the wrong basis! Switching from the basis $\{\vec{d}_1, \vec{d}_2\}$ to $\{\vec{d}_1 + \vec{d}_2, \vec{d}_1 - \vec{d}_2\}$, the matrices representing r and s (and thus

[2] Sometimes W is said to be a G-**invariant subspace** of V, but we will avoid this language here for confusion with related concepts.

[3] We have ignored until this point (and will continue to do so after this point) the idea of a 0-**dimensional** representation. Indeed, for any fixed ground field F, there is a unique 0-dimensional vector space $\{0\}$ (with basis the empty set), which in turn only has the zero map $0 : \{0\} \to \{0\}$ (an isomorphism!) as its only self-map. We can certainly define the 0-representation $G \to \mathrm{GL}(\{0\})$ by sending every $g \mapsto 0$, but this is non-interesting for our purposes. We will only consider positive-dimensional representations in this course, and we do not consider the 0-representation to be reducible or irreducible.

each matrix in the representation) are diagonal:

$$r \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$
$$s \mapsto \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We say that $\mathcal{D}_8 \circlearrowright \mathbb{R}[X]$ has been **decomposed** as $U \oplus U'$ (cf. Example 2.24).

Definition 4.3. A representation $\rho : G \to \mathrm{GL}(V)$ is called **decomposable** if there are proper subrepresentations $W, W' \leq_G V$ such that $V = W \oplus W'$. In particular, any decomposable representation is always reducible.

Remark 4.2. Decomposition is equivalent to finding a basis of V such that *each* $L \in \mathrm{im}\,\rho(g)$ *has a block diagonal form* (according to some uniform block structure). Indeed, let $x \in V$. Recall that if $(w_1, \dots, w_k)$ and $(w'_1, \dots, w'_{n-k})$ are bases for W and W', respectively, then in the basis for V given by $\mathscr{B} = (w_1, \dots, w_k, w'_1, \dots, w'_{n-k})$ we have

$$[L]_{\mathscr{B}} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ and } [x]_{\mathscr{B}} = \begin{pmatrix} y \\ z \end{pmatrix}$$

where A is a $k \times k$ matrix, B is $k \times (n-k)$, C is $(n-k) \times k$, D is $(n-k) \times (n-k)$, y is a k-dimensional vector, and z is $(n-k)$-dimensional. Moreover, matrix multiplication works out nicely, e.g.

$$[Lx]_{\mathscr{B}} = \begin{pmatrix} Ay + Bz \\ Cy + Dz \end{pmatrix}$$

For W and W' to be subrepresentations, L must sends vectors in W to vectors in W and, similarly, those in W' should stay in W'. This amounts to

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} \overset{\text{req}}{=} \begin{pmatrix} * \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 \\ z \end{pmatrix} \overset{\text{req}}{=} \begin{pmatrix} 0 \\ * \end{pmatrix},$$

i.e., we need $B = 0$ and $C = 0$. That's block diagonalization!

Example 4.15. If we take U to be the span of $(1 \ \cdots \ 1)^\mathsf{T} \in \mathbb{C}^n$, then U is a subrepresentation of the permutation representation $\mathcal{S}_n \circlearrowright \mathbb{C}^n$. Consider the case when $n = 3$, writing A for the permutation matrix induced by (1 2) and B for that induced by (1 2 3), where we switch to the basis $\{e_1 + e_2 + e_3, e_2, e_3\}$. We can use, among other tools, the SageMath computer algebra system (2025) to compute this basis change:

$$\texttt{A in new basis:} \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

$$\texttt{B in new basis:} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

We can see from the form of these matrices that this complementary subspace $\mathrm{Span}\{e_2, e_3\}$ of U is not itself an $\mathcal{S}_3$ representation, as it is not closed under the $\mathcal{S}_3$-action.

Listing 4.1 Changing basis.

```
A = matrix([[0,1,0],[1,0,0],[0,0,1]])
B = matrix([[0,0,1],[1,0,0],[0,1,0]])
v1 = vector([1,1,1])
e2 = vector([0,1,0])
e3 = vector([0,0,1])
S = column_matrix([v1,e2,e3])
show("A in new basis: ", S.inverse()*A*S)
show("B in new basis: ", S.inverse()*B*S)
```

The orthogonal complement of U under the usual dot product *is* a subrepresentation,

$$V \coloneqq U^{\perp} = \{\alpha_1 e_1 + \cdots + \alpha_n e_n : \alpha_1 + \cdots + \alpha_n = 0\},$$

since the sum of entries is invariant under permutations of those entries. This subspace V is called the **standard representation** of S_n. For $n = 3$, we could take the vectors $v_2 = e_1 - e_2$ and $v_3 = e_2 - e_3$ as a basis for V, then switch to the basis $\{e_1 + e_2 + e_3, v_2, v_3\}$:

Listing 4.2 Block diagonalization.

```
v2 = vector([1,-1,0])
v3 = vector([0,1,-1])
S = column_matrix([v1,v2,v3])
show("A in new basis: ", S.inverse()*A*S)
show("B in new basis: ", S.inverse()*B*S)
```

$$\texttt{A in new basis:} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\texttt{B in new basis:} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}$$

While these matrices are not diagonal, they are in a *block diagonal* form since $\mathbb{C}^3 = U \oplus V$ is a decomposition of representations. This is the best we can hope for: since A and B do not commute, they cannot simultaneously diagonalize.

The last improvement we mention comes from choosing a better basis for V, in particular an eigenbasis for the submatrix of B. If we set $\zeta_3 \in \mathbb{C}$ as a non-trivial third root of unity, i.e. a root of the polynomial $x^3 - 1 = (x-1)(x^2 + x + 1)$ that is not equal to 1, then $\zeta_3^2 + \zeta_3 + 1 = 0$. It then follows that

$$w_2 = \zeta_3^2 e_1 + \zeta_3 e_2 + e_3 \text{ and } w_3 = \zeta_3 e_1 + \zeta_3^2 e_2 + e_3$$

lie in (and are a basis for) V. Notice that the collection $\{e_1 + e_2 + e_3, w_2, w_3\}$ is nearly an orthonormal set, needing only to be normalized. We have:

$$\texttt{A in new basis:} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Listing 4.3 Unitary block diagonalization. Notice that Sage prints $\zeta_3^2 = -\zeta_3 - 1$ by preferring the expression with lower-order powers of ζ_3. We use the field F defined here for these calculations because computer algebra over the complex numbers can be problematic.

```
F.<zeta> = CyclotomicField(3)
# This initializes a field extension of the rationals
# containing a third root of unity to do our calculations in.

w2 = vector([zeta^2,zeta,1])
w3 = vector([zeta,zeta^2,1])
S = column_matrix([v1,w2,w3])
show("A in new basis: ", S.inverse()*A*S)
show("B in new basis: ", S.inverse()*B*S)
```

$$\texttt{B in new basis:} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta_3 & 0 \\ 0 & 0 & -\zeta_3 - 1 \end{pmatrix}$$

We will see that the standard representation V is an irreducible representation of S_n.

Example 4.16. The representation $\rho : \mathbb{Z} \to \mathrm{GL}_2(\mathbb{C})$ given by

$$\rho(k) := \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$$

is reducible but not decomposable. In particular, e_1 is an eigenvector for every matrix in the representation, and so $\mathrm{Span}\{e_1\}$ is a subrepresentation of $\mathbb{C}^2$. However, $\rho(k)$ is not diagonalizable for every $k \neq 0$, since the characteristic polynomial of $\rho(k)$ is $(t-1)^2$ but

$$\mathrm{Null}(\rho(k) - \mathbb{I}) = 1.$$

For those familiar with the language, the geometric multiplicity of the eigenvalue 1 is less than the algebraic multiplicity whenever $k \neq 0$. Problematic representations like this are why we temporarily restrict to finite groups in our initial treatment of representation theory.

Notice that the work of diagonalizing our matrices, i.e., their spectral theory, seems to be related to how a representation might decompose into subrepresentations!

Example 4.17. If the image of a representation $V = \mathbb{C}^n$ consists of permutation matrices (e.g., $S_n \circlearrowright \mathbb{C}^n$ or $G \circlearrowright \mathbb{C}[G]$) and $W \leq_G V$, then the orthogonal complement

$$W^\perp = \{y \in V : x^* y = 0 \text{ for all } x \in W\}$$

under the standard Hermitian dot product is a subrepresentation of V. Moreover, since $V = W \oplus W^\perp$ (recall Proposition 2.5), we have a decomposition of S_n-representations.

Indeed, if $g \in G$ and $y = \beta_1 e_1 + \cdots \beta_n e_n$, then

$$g \cdot y = \beta_1 e_{\sigma_g(1)} + \cdots + \beta_n e_{\sigma_g(n)}.$$

for some $\sigma_g \in S_n$. Furthermore, if $x = \alpha_1 e_1 + \cdots \alpha_n e_n \in W$, then

$$g^{-1} \cdot x = \alpha_{\sigma_g(1)} e_1 + \cdots + \alpha_{\sigma_g(n)} e_n \in W$$

because W is a subrepresentation (and thus is closed under the G-action). Therefore

$$\begin{aligned} x^*(g \cdot y) &= (\alpha_1 e_1 + \cdots \alpha_n e_n)^* (\beta_1 e_{\sigma_g(1)} + \cdots + \beta_n e_{\sigma_g(n)}) \\ &= \sum_{i=1}^{n} \overline{\alpha_i} \beta_{\sigma_g^{-1}(i)} \\ &= \sum_{j=1}^{n} \overline{\alpha_{\sigma_g(j)}} \beta_j \\ &= (\alpha_{\sigma_g(1)} e_1 + \cdots + \alpha_{\sigma_g(n)} e_n)^* (\beta_1 e_1 + \cdots \beta_n e_n) \\ &= (g^{-1} \cdot x)^* y = 0, \end{aligned}$$

so $W^\perp$ is closed under the G-action.

4.3 Weyl's Trick

Henceforth, all representations will be understood as over the complex numbers unless explicitly stated otherwise.

In general, there are many ways to take orthogonal complements, based on how one chooses to measure angles. Given a subrepresentation W of V, most inner products $\langle \cdot, \cdot \rangle$ will *not* give rise to a complement $W^\perp$ that is also a subrepresentation of V. However, if one is sufficiently crafty, it turns out there *is* an inner product on V that guarantees the orthogonal complement is G-invariant for a wide class of examples. Hermann Weyl published such a result in 1925, based on a technique by Adolf Hurwitz from 1897, and therefore it is sometimes referred to as Weyl's unitary trick. However we decide to name it, the trick comes down to the power of *averaging*.

Lemma 4.1 (Unitary trick)**.** *Let G be a finite group with a representation $\rho : G \to \mathrm{GL}(V)$ over $\mathbb{C}$. Then there exists a Hermitian inner product $\langle \cdot, \cdot \rangle_G$ on V such that*

$$\langle g \cdot x, g \cdot y \rangle_G = \langle x, y \rangle_G$$

for all $g \in G$ and $x, y \in V$. That is, the $\rho(g)$ are unitary *with respect to $\langle \cdot, \cdot \rangle_G$; we say that $\langle \cdot, \cdot \rangle_G$ is a* ***G-invariant inner product****.*

Remark 4.3. In light of the unitary trick, given a representation $\rho : G \to \mathrm{GL}(V)$ with G finite and V a finite-dimensional complex vector space, we can choose an orthonormal basis $\{v_1, \ldots, v_n\} \subset V$ via the G-invariant inner product $\langle \cdot, \cdot \rangle_G$ and the Gram–Schmidt process. Writing down matrices in this basis gives an isomorphism $\mathrm{GL}(V) \cong \mathrm{GL}_n(\mathbb{C})$ such that each linear transformation $\rho(g)$ passes to a unitary matrix.

If we had chosen a basis from the get-go, so that the $\rho(g)$ are thought of as matrices rather than abstract transformations, then this lemma shows that a change of basis will bring the image of ρ into a subset of the unitary group $\mathrm{U}(n)$, where $n = \dim(V)$. In either case, Weyl's trick guarantees that our representations factor through a unitary group

$$\begin{array}{ccc} G & \xrightarrow{\rho} & \mathrm{GL}(V) \\ \vdots & & \updownarrow \cong \\ \mathrm{U}(n) & \hookrightarrow & \mathrm{GL}_n(\mathbb{C}). \end{array}$$

of Lemma 4.1. Let $\langle \cdot, \cdot \rangle$ be an arbitrary Hermitian inner product on V; we will use it to construct a new inner product satisfying the desired property. For $x, y \in V$, we define:

$$\langle x, y \rangle_G \coloneqq \frac{1}{|G|} \sum_{h \in G} \langle h \cdot x, h \cdot y \rangle. \tag{4.1}$$

Note that if $\langle \cdot, \cdot \rangle$ already satisfies the desired property, then $\langle \cdot, \cdot \rangle_G = \langle \cdot, \cdot \rangle$.

First we check to make sure that this new form $\langle \cdot, \cdot \rangle_G$ is still a Hermitian inner product. Fortunately, linearity in the second argument and conjugate-symmetric follow immediately from the respective properties of $\langle \cdot, \cdot \rangle$. The positive-definiteness condition is also clear, since

$$\langle x, x \rangle_G = \frac{1}{|G|} \sum_{h \in G} \langle h \cdot x, h \cdot x \rangle = \frac{1}{|G|} \sum_{h \in G} \|h \cdot x\|^2.$$

That is, $\langle x, x \rangle_G = 0$ if and only if $\|h \cdot x\| = 0$ for each $h \in G$, which can only happen if $x = 0$.

Lastly, we will show the unitary condition. This calculation relies on the fact that the mapping $G \to G$ given by right multiplication ($h \mapsto hg$) for some fixed g, while not a homomorphism, *is a bijection*. We compute:

$$\begin{aligned} \langle g \cdot x, g \cdot y \rangle_G &= \frac{1}{|G|} \sum_{h \in G} \langle h \cdot (g \cdot x), h \cdot (g \cdot y) \rangle \\ &= \frac{1}{|G|} \sum_{h \in G} \langle (hg) \cdot x, (hg) \cdot y \rangle \\ &= \frac{1}{|G|} \sum_{h' \in G} \langle h' \cdot x, h' \cdot y \rangle = \langle x, y \rangle_G. \end{aligned}$$

□

Indeed, though we will not prove it here, this *G-invariant inner product* is uniquely determined (up to scaling) on each irreducible representation V.

Example 4.18. A representation $G \to \mathrm{GL}_n(\mathbb{C})$ whose image consists of permutation matrices (cf. Example 4.17) or, more generally, unitary matrices, has $\langle x, y \rangle_G = x^* y$.

Example 4.19. We have seen the action of S_3 on the standard representation $V \leq_{S_3} \mathbb{C}^3$ with the basis $\{e_1 - e_2, e_2 - e_3\}$ written down as

$$(1\,2) \mapsto A \coloneqq \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix},$$
$$(1\,2\,3) \mapsto B \coloneqq \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}.$$

We can compute $\langle \cdot, \cdot \rangle_{S_3}$ as in Equation 4.1 via SageMath (Listing 4.4):

Listing 4.4 Computing an invariant inner product using symbolic algebra.

```
# Define a polynomial ring over the rationals (QQ)
R.<alpha_1, alpha_2, beta_1, beta_2> = QQ[]

A = matrix([[-1, 1], [0, 1]])
B = matrix([[0, -1], [1, -1]])
x = vector([conjugate(alpha_1), conjugate(alpha_2)])
y = vector([beta_1, beta_2])
inner_prod = 0

# Loop over the image of the representation:
# + identity_matrix(2): the 2x2 identity matrix
# + B, B^2 (the 3-cycles)
# + A, A*B, A*B^2 (the 2-cycles)
for g in [identity_matrix(2), B, B*B, A, A*B, A*B*B]:

    inner_prod += (conjugate(g)*x) * (g*y) / 6
    # Note that the conjugate isn't necessary here, since
    # everything is real... but just for completeness!

show(inner_prod.expand())
```

$$\frac{4}{3}\beta_1\overline{\alpha_1} - \frac{2}{3}\beta_2\overline{\alpha_1} - \frac{2}{3}\beta_1\overline{\alpha_2} + \frac{4}{3}\beta_2\overline{\alpha_2}$$

If we think carefully about what is happening in this calculation, we realize that we don't need the symbolic variables at all—but they are helpful to have as we get our footing! Listing 4.5 computes the invariant inner product more directly:

Listing 4.5 Computing an invariant inner product by linear algebra.

```
A = matrix([[-1,1],[0,1]])
B = matrix([[0,-1],[1,-1]])
G = [identity_matrix(2), B, B*B, A, A*B, A*B*B]
M = sum(conjugate(g).transpose() * g for g in G) / 6
show(M)
```

$$\begin{pmatrix} \frac{4}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{4}{3} \end{pmatrix}$$

In any case, the matrices A and B (and all products therein) are unitary with respect to the inner product (cf. Theorem 2.12):

$$\left\langle \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \right\rangle_{S_3} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}^* \begin{pmatrix} 4/3 & -2/3 \\ -2/3 & 4/3 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}.$$

Corollary 4.1. *If $\rho : G \to \mathrm{GL}(V)$ is a finite-dimensional complex representation of a finite group, the eigenvalues of each $\rho(g)$ are all valued in the unit circle $\mathbb{T} = \mathrm{U}(1) \subset \mathbb{C}$.*

4.4 Maschke's Theorem

Unitary matrices preserve angles—in light of Weyl's trick, it makes perfect sense to turn to orthogonal complements in order to find G-invariant subspaces.

Theorem 4.1 (Maschke's Theorem)**.** *Let G be a finite group with V a representation of G over $\mathbb{C}$ and $W \leq_G V$. Then there exists a complementary subrepresentation $W' \leq_G V$ such that $V = W \oplus W'$, i.e., V is reducible if and only if it is decomposable.*

Proof. Take $W' = W^\perp$ under the inner product $\langle \cdot, \cdot \rangle_G$ of Weyl's unitary trick, Equation 4.1, so that $W \oplus W^\perp = V$ (as in Proposition 2.5). To show that W' is also a subrepresentation of V, take arbitrary elements $x \in W', y \in W$, and $g \in G$. Then:

$$\langle g \cdot x, y \rangle_G = \langle x, \underbrace{g^{-1} \cdot y}_{\in W} \rangle_G = 0,$$

where we know $g^{-1} \cdot y \in W$ since $W \leq_G V$, and so we conclude $g \cdot x \in W'$. □

Corollary 4.2 (Complete Reducibility)**.** *If G is a finite group with V a finite-dimensional representation over $\mathbb{C}$, then V decomposes as a sum of irreducible subrepresentations:*

$$V = V_1 \oplus \cdots \oplus V_r.$$

Proof. Given a representation V of G, either V is irreducible or reducible. In the former case, we are done; otherwise, decompose V via Maschke's theorem into subrepresentations $V = W \oplus W'$. Next consider W: if W is irreducible, proceed to analyzing W'; otherwise, apply the theorem again to decompose W into subrepresentations. Since V is finite-dimensional, this process of breaking representations into irreducible representations must stop. □

The equivalence of reducibility and decomposability, together with Corollary 4.2, orients us to several fundamental priorities for the representation theory of finite groups.

Remark 4.4. For a given finite group G:

- How can we tell if a representation $G \to \mathrm{GL}(V)$ is irreducible?
- How "many" irreducible representations are there of G, up to some notion of equivalence?
  - Can we classify them in some way?
  - How should we compare representations?
- Can we decompose a given representation into irreducible representations?
- Are irreducible decompositions of representations unique in some sense?

Chapter 5

Intertwiners

Coming soon

Chapter 6

Constructions

Coming soon

Part III

Character Theory

Chapter 7

Introduction

Coming soon

Chapter 8

Character Tables

Coming soon

$\mathcal{D}_8$	1 e	1 r^2	2 r	2 s	2 sr
trivial	1	1	1	1	1
$\langle r \rangle$-kernel	1	1	1	-1	-1
$\langle s, r^2 \rangle$-kernel	1	1	-1	1	-1
$\langle sr, r^2 \rangle$-kernel	1	1	-1	-1	1
action on plane	2	-2	0	0	0

Chapter 9

Fourier Theory

Coming soon

Part IV

Compact Groups

Chapter 10

Introduction

Coming soon

Chapter 11

The Group $U(1)$

Coming soon

Chapter 12

The Group $SU(1)$

Coming soon

References

Artin, M. 2011. *Algebra*. Pearson Education.

Bretscher, O. 2013. *Linear Algebra with Applications*. Pearson Education.

Debnath, L., and P. Mikusinski. 2005. *Introduction to Hilbert Spaces with Applications*. Elsevier Science.

Diaconis, P. 1988. *Group Representations in Probability and Statistics*. IMS Lecture Notes. Institute of Mathematical Statistics.

Dummit, D. S., and R. M. Foote. 2003. *Abstract Algebra*. Wiley.

Fulton, W., and J. Harris. 1991. *Representation Theory: A First Course*. Graduate Texts in Mathematics. Springer New York.

Isaacs, I. M. 1994. *Character Theory of Finite Groups*. Dover Books on Advanced Mathematics. Dover.

Logan, J. D. 2015. *A First Course in Differential Equations*. Undergraduate Texts in Mathematics. Springer International Publishing.

Petersen, P. 2012. *Linear Algebra*. Undergraduate Texts in Mathematics. Springer New York.

Rossmann, Wulf. 2002. *Lie Groups: An Introduction Through Linear Groups*. Oxford University Press.

Serre, Jean-Pierre. 1977. *Linear Representations of Finite Groups*. Vol. 42. Graduate Texts in Mathematics. Springer-Verlag.

Shahriar, S. 2017. *Algebra in Action: A Course in Groups, Rings, and Fields*. Pure and Applied Undergraduate Texts. American Mathematical Society.

Shaw, R. 1982. *Linear Algebra and Group Representations*. Linear Algebra and Group Representations. Academic Press.

Shurman, Jerry. 1997. *Geometry of the Quintic*. John Wiley & Sons.

The Sage Developers. 2025. *SageMath, the Sage Mathematics Software System*. Version 10.8. https://www.sagemath.org.

Appendices

Syllabus

Math 362: Representation Theory

Course Basics

Instructor	Claudio Gómez-Gonzáles	**Class Time**	2-3c
Email	cgonzales@carleton.edu	**Class Location**	CMC 209
Office	CMC 321	**Drop-In Location**	CMC 306 and CMC 328

Websites

Moodle, Course Notes, Google Drive

Textbook

Representation Theory Notes, by Claudio Gómez-Gonzáles. Additional resources include:

- *Representation Theory: A First Course*, by Fulton and Harris.
- *Character Theory of Finite Groups*, by Isaacs.
- *Group Representations in Probability and Statistics*, by Diaconis.

Prerequisites

Math 342 or instructor permission

Course description

Representation theory is the study of abstract structures via the tools of linear algebra. The first objects to be studied in this way were finite groups, motivated by the powerful framework of characters in number theory, but the field has generalized incredibly due to the prevalence of symmetry throughout mathematics, physics, and beyond. Topics include group modules, semisimple algebras, and Maschke's Theorem; characters, orthogonality relations, and character tables; Fourier transformations and random walks. Mathematical communication will be an important aspect of the course.

How to succeed

- Pre-read actively and come to class with questions,
- Begin homework immediately after it is assigned,
- Work with each other and work towards building a supportive community,
- Visit my office hours frequently and ask lots of questions!

Talk to me! I have specifically worked to design a course that challenges the ways we think and allows us to grow together. If you need help, or even if you don't and just have suggestions or thoughts, please come to my office! I promise to treat your problems with respect and to keep any sensitive conversations as confidential as I can, but note that (under Title IX) I am a mandatory reporter of sexual misconduct.

Course Outline

Schedule is subject to change! While the material we learn is building up a coherent and interrelated set of ideas, the primary content evaluated in a particular Midterm corresponds to the Module of the same number (the final Exam corresponds largely to the final Module).

Module 1: Fundamentals & Structure

Week 1: Fundamentals

- **Tuesday:** *Algebra review.* §2.1 and §3.
  - Syllabus discussion,
  - Abstract linear algebra and interplay with group theory.
- **Thursday:** *Complex linear algebra.* §2.2–2.3.
  - Hermitian inner products,
  - Orthogonal complements and decompositions,
  - Functionals, Riesz representation, and adjoints.
- **Homework 0** by Friday at 4 PM.
- **Journal 1** by Saturday at midnight.

Week 2: Representations

- **Tuesday:** *Representations.* §4.1–4.2.
  - Unitary diagonalization,
  - Basic properties of representations,
  - Subrepresentations and examples.
- **Thursday:** *Maschke's theorem.* §4.3–4.4.
  - Weyl's unitary trick and complete reducibility,
  - Introduction to `SageMath`.
- **Homework 1** due by Wednesday at 4 PM.
- **Journal 2** due by Saturday at midnight.

Week 3: Structure and Classification

- **Tuesday:** *Intertwiners.* §5.1.
  - Morphisms of actions as G-linearity,
  - Properties and examples.
- **Thursday:** *Fundamental theorem.* §5.2 and §6.1.
  - Schur's Lemma,
  - Uniqueness of decompositions.
- **Homework 2** due by Tuesday at 4 PM.
- **Journal 3** due by Saturday at midnight.

Module 2: Character Theory

Week 4: Aside on Linear Algebra

- **Tuesday:** *Duals and symmetric powers.* §6.2–6.3.
  - Decomposition revisited,
  - Dual representations and polynomials.
- **Thursday:** *Hom spaces and tensors.* §6.4–6.5.
  - Maps and multilinearity,
  - G-linearity as G-invariance.
  - **Begin Midterm 1.**
- **Homework 3** due by Tuesday at 4 PM.
- **Journal 4** due by Saturday at midnight.

Week 5: Character Theory

- **Tuesday:** *Characters.* §7.1.
  - Basic properties,
  - Invariant subrepresentations,
  - A projection formula.
- **Thursday:** *Orthonormality.* §7.2.
  - Spaces of intertwiners,
  - Schur's Lemma revisited.
- **Homework 4** due by Tuesday at 4 PM.
- **Midterm 1** due in class on Thursday.
- **Journal 5** due by Saturday at midnight.

Week 6: Character Tables

- **Tuesday:** *Fundamental Theorem.* §8.1.
  - Orthonormality of characters.
- **Thursday:** *Applications.* §8.2.
  - **Begin Midterm Project.**
- **Homework 5** due by Tuesday at 4 PM.
- **Journal 6** due by Saturday at midnight.

Week 7: Fourier Theory I

- **Tuesday:** *Random walks on finite groups.* §9.1.
  - Probability distributions and convolutions,
  - Fourier transforms.
- **Thursday:** *Plancherel formula.* §9.2.
  - Fourier inversion,
  - Shuffling cards.
- **Homework 6** due by Tuesday at 4 PM.
- **Journal 7** due by Saturday at midnight.

Module 3: Compact Groups

Week 8: Topology of Infinite Groups

- **Tuesday:** *Review.* §9.3.
  - Wrapping up Fourier for finite groups,
  - Accounting of our progress.
- **Thursday:** *Groups and integrals.* §10.1–10.3.
  - How can we generalize?
  - Topological necessities,
  - Averaging revisited.
  - **Begin Midterm 2.**
- **Homework 7** due by Tuesday at 4 PM.
- **Midterm Project** due by Thursday at 4 PM.
- **Journal 8** due by Saturday at midnight.

Week 9: Fourier Theory II

- **Tuesday:** *The group* U(1). §10.4 and §11.1.
  - Pontryagin duality,
  - Exponentials and path lifting.
- **Thursday:** *The representation theory of* U(1). §11.2–11.3.
  - Fourier theory,
  - Irreducible representations.
- **Homework 8** due by Tuesday at 4 PM.
- **Midterm 2** due in class on Thursday.
- **Journal 9** due by Saturday at midnight.

Week 10: Further Topics

- **Tuesday:** *The representation theory of* SU(2). §12.1, §12.3–12.4.
  - Haar measure,
  - Connections to quaternions and the group SO(3).
- **Thursday:** *Peter–Weyl and beyond.* §12.5.
  - Wrapping up, review, and where to go next.
  - **Begin Final Exam.**
- **Homework 9** due by Tuesday at 4 PM.

Final Exam Period

Week 11:

- **Review Sessions**
- **Final Exam** due 2026-03-16 by 4 PM.

Grade Details

Grading

This class will be graded on an A–F scale, as detailed below.

A– [90%, 93%)	**A** [93%, 100%]	
B– [80%, 83%)	**B** [83%, 87%)	**B+** [87%, 90%)
C– [70%, 73%)	**C** [73%, 77%)	**C+** [77%, 80%)
D [60%, 67%)		**D+** [67%, 70%)
F [0, 59%)		

I reserve the right to change this distribution, but will only do so in a way that would make your grade better (never worse). In general, a B indicates that you have learned the key concepts of this course and could reliably apply them in the future. An A indicates that you have demonstrated a deeper understanding not only of how to apply ideas but also in communicating and exploring concepts beyond the scope of the course.

Grade breakdown

Final marks for the course will be computed using the following weights.

- **Community**, worth 2% of your grade. There are chances to do this every day—by contributing to discussions, working with others on homework, representing your group after breakouts, and contributing to the Notes repository. Your presence benefits you, your classmates, and me!
- **Journaling**, done weekly, in total worth 3% of your grade. These are written assignments due at the end of each week that provide space for self-evaluation, chronicling what you've learned, synthesizing concepts from previous classes, and providing ongoing course feedback.
- **Homework**, ten assignments, in total worth 20% of your grade. These problem sets allow you to exercise techniques that we have discussed in class and also build experience in mathematical communication. You are encouraged to work in groups! Much of the work of this class will happen in these assignments, where you connect with peers and cement your own knowledge. Turn in homework during class or to the course mailbox. Late Homework is accepted with a penalty, detailed in the Late work policy; Homework that is more than a week late will not be accepted. Your lowest assignment will be dropped in computing your final grade.
- **Midterm Project** (aka, the Great Character Hunt), worth 20% of your grade. This collaborative project will involve computing the character table invariant for various finite groups, and submitting results in the style of a professional mathematical paper. You will practice summarizing methods of calculation, describing computed results, and proof writing.
- **Exams**, two midterms and a final, in total worth 55% of your grade. The exams are spaces for you to demonstrate the material that you have mastered on your own. If you require additional accommodations for these forms of assessment, let me know well in advance.

Support and Other Policies

Drop-in hours

These are times that I set aside during my week to be available for you. Just show up! You do not need to make an appointment and you are not annoying me.

Late work

For every day that a Homework assignment is turned in late, the associated grade will be dropped $\frac{1}{2}$ of a letter grade (e.g., an assignment turned in Wednesday that was due Monday and would have received an A would receive a B). You will have 4 Late Credits to apply to late Homework assignments. Each Late Credit is worth 1 day to turn in an assignment late without penalty: e.g., turning in an assignment due Wednesday on the following Friday would use 2 Late Credits. Late assignments will not be accepted beyond a week after the original deadline, except in explicit cases of extension. As per College policy, Late Credits cannot extend deadlines past the last day of class.

Collaboration and academic integrity

Math is a collaborative activity! Even when we publish a paper alone, mathematicians are part of a sociopolitical fabric animated by our roles in institutions of education, research, and industry. You should work with your classmates to learn this material; you will do this in class! However, you may not copy anyone else's work. Rather, you must write up your own solutions and give credit to classmates and other collaborators for important insights. Cases of academic dishonesty are taken seriously by the College and I am required to report them.

Large language models and generative AI

I understand the increasing ubiquity of these technologies, but I discourage their use in this course. While I can imagine some applications—helping you experiment with SageMath, for example—I encourage you to always critically reflect on whose labor is being replaced or which relationships lose out when you invoke these tools. If you are interested in thinking about sociological dimensions of automation and mathematical labor, here's a good place to start. Remember, I ask you to grapple with complex ideas in this course because we grow from productive struggle.

> In general, I view "generational use" of LLMs as inappropriate for this class, while "assistive use" can be appropriate. You may not ask an LLM to solve a problem for you or ask questions of such a system that you would not ask of a classmate or me. I do not authorize the sharing of my course materials with AI platforms. Moreover, I ask you to only turn to LLMs towards understanding a homework problem (*never* on exams or course projects) after working with classmates and me. Transcription software is not allowed in class except with explicit permission via accommodation request.

Sources of support

You should expect to be challenged in this course! If you are stuck, know that you are following in the footsteps of all who came before you. Here is a list of available resources, some of which will be expanded upon below:

- Your classmates and me!
- The Academic Support Center, Student Health and Counseling, and the Dean of Students Office.

Student health

Your well-being should be your first priority. It is important to recognize stress you may be facing, which can be personal, emotional, physical, financial, or academic. Sleep, exercise, and building a supportive community are important! Please do not come to class if you are sick—instead, stay in communication with me and other students. If you cannot attend class for an extended period of time, reach out to me on how we can make accommodations for missed material and late work.

Accommodations for students with disabilities

The Office of Accessibility Resources is the campus office that collaborates with students who have disabilities to provide and/or arrange reasonable accommodations. If you have, or think you may have, a disability (e.g., mental health, attentional, learning, autism spectrum disorders, chronic health, traumatic brain injury and concussions, vision, hearing, mobility, or speech impairments), please contact `oar@carleton.edu` to arrange a confidential discussion regarding equitable access and reasonable accommodations. The College also makes available assistive technologies including audio recording Smartpens, text-to-speech and speech-to-text software, and more.

Personal electronics

There are no restrictions on phones, tablets, laptops, or other electronic equipment in the classroom, provided that said equipment is being used respectively and non-disruptively. Please silence your devices and be mindful of others, especially in discussions or other collaborative contexts.

Title IX

Be aware that all Carleton faculty and staff members, with the exception of Chaplains and SHAC staff, are "responsible employees." Responsible employees are required to share any information they have regarding incidents of sexual misconduct with the Title IX Coordinator. Carleton's goal is to ensure community members are aware of all the options available and have access to the resources they need. If you have questions, please contact Carleton's Title IX Coordinator or visit the Sexual Misconduct Prevention and Response website.

Land Acknowledgement: Carleton derives wealth and prestige through its ownership of ancestral homelands of the Wahpekute and Mdewakanton bands of the Dakota Nation and more broadly as an academic institution in the United States. I urge you to support the organizing work of Indigenous peoples seeking liberation through direct action, advocacy, and education.

Part V

Homework

Homework 0

Exercise 0.1. Let V and W be F-vector spaces and let $L : V \to W$ be a linear map. Take $x \in V$ and set $y = L(x) \in W$. Prove that

$$L^{-1}(y) = x + \ker L.$$

Exercise 0.2. Let F be a field and suppose that $\phi : \mathrm{Mat}_n(F) \to F$ satisfies $\phi(AB) = \phi(A)\phi(B)$ for all $A, B \in \mathrm{Mat}_n(F)$ and that $\phi(\mathbb{I}) = 1$.

a) If $A \in \mathrm{GL}_n(F)$, show that $\phi(A^{-1}) = \phi(A)^{-1}$.

b) Suppose henceforth that ϕ is non-trivial, i.e., that ϕ is not the constant map sending every matrix to 1. Show $\phi(0) = 0$.

c) Show that that $\phi(J) = 0$, where J is the matrix with 1s along the superdiagonal and 0s elsewhere.

Hint: What is J^2? What about other powers of J?

d) Note that if the columns of $M \in \mathrm{Mat}_n(F)$ are $c_1, c_2, \dots, c_n$, then the columns of MJ are $0, c_1, \dots, c_{n-1}$. Now, suppose $\mathrm{rank}(A) = n - 1$. Show that there is always a basis change $S \in \mathrm{GL}_n(F)$ such that the first column of $S^{-1}AS$ is zero. Conclude that $A = TJS^{-1}$ for some $S, T \in \mathrm{GL}_n(F)$ and hence $\phi(A) = 0$.

Since any non-invertible matrix A may be written as a product of rank $n - 1$ matrices, we can conclude that $\phi(A) = 0$ for all non-invertible $A \in \mathrm{Mat}_n(F)$ and hence ϕ is determined by its values on $\mathrm{GL}_n(F)$.

Exercise 0.3. Show $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$ for all $A, B \in \mathrm{Mat}_n(F)$; conclude that the trace is independent of the choice of basis.

Exercise 0.4. Show that any linear map $f : \mathrm{Mat}_n(F) \to F$ satisfying $f(AB) = f(BA)$ for all $A, B \in \mathrm{Mat}_n(F)$ and $f(\mathbb{I}) = n$ is the trace.

Hint: Consider the standard basis of matrices, i.e., those of the form $E_{i,j} \in \mathrm{Mat}_n(F)$ for all $1 \le i, j \le n$. What happens when you multiply these matrices?

Exercise 0.5. Let $\lambda \in \mathbb{C} \setminus \{0\}$. Show that the Heisenberg relation $AB - BA = \lambda\mathbb{I}$ cannot hold for matrices, i.e., quantum mechanics requires infinite dimensions.

Hint: Take the trace of both sides.

Exercise 0.6. A matrix $S \in \mathrm{Mat}_n(F)$ is called a **permutation matrix** if it has exactly one entry of 1 in each row and each column and 0 elsewhere—such a matrix induces a permutation $\sigma \in \mathcal{S}_n$ through its action on column vectors. For example

$$(1\,2\,4) \in \mathcal{S}_4 \mapsto \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Recall that i is a **fixed point** of $\sigma \in \mathcal{S}_n$ if $\sigma(i) = i$. Show that

$$|\mathrm{Fix}(\sigma)| = \mathrm{Tr}(S).$$

Homework 1

Exercise 1.1. Let $P : V \to V$ be a projection.

a) Show that $\mathbb{I} - P : V \to V$ is also a projection, that $\ker P = \operatorname{im}(\mathbb{I} - P)$, and that $V = \ker P \oplus \operatorname{im} P$.

b) Show that an eigenvalue of a projection P can only be 0 or 1. Conclude that $\operatorname{Tr}(P) = \dim(\operatorname{im} P)$, i.e., the trace computes the dimension of the subspace projected onto by P.

Exercise 1.2. If V is an inner product space and $\|x_0\| = 1$, show $P(x) \coloneqq \langle x_0, x\rangle x_0$ is an orthogonal projection (cf. Remark 2.5) onto $\operatorname{Span}\{x_0\}$.

Exercise 1.3. Show that a pairwise orthonormal set $\mathscr{A}$ of an inner product space V is necessarily a linearly independent set.

Exercise 1.4. Verify that the set $S = \left\{ \frac{1}{\sqrt{2\pi}} e^{int} \;\middle|\; n \in \mathbb{Z} \right\}$ is pairwise orthonormal in the inner product space $C([-\pi, \pi], \mathbb{C})$ with respect to

$$\langle f, g\rangle \coloneqq \int_{-\pi}^{\pi} \overline{f(t)} g(t)\, \mathrm{d}t.$$

Exercise 1.5. Let $Q \in \operatorname{Mat}_n(\mathbb{C})$. Prove that the following are equivalent:

i) $Q \in \mathrm{U}(n)$ (see Definition 2.30).
ii) The columns of Q are orthonormal with respect to the Hermitian dot product.
iii) The rows of Q are orthonormal with respect to the Hermitian dot product.

Exercise 1.6. Let $A = \begin{pmatrix} -10 & 18 & 18 \\ 3 & -7 & -6 \\ -9 & 18 & 17 \end{pmatrix}$. Compute A^{10}.

Remark: You can easily do this by brute force with a computer algebra system like `SageMath`, but that isn't the point—we can do this manually! The hint here is: if $S^{-1}AS$ is diagonal, computing its powers is easy. Does that help with powers of A?

Exercise 1.7. Let V be an inner product space with $L : V \to V$ a linear transformation.

a) Show that L can be *uniquely* decomposed as $L = A + Bi$, where A and B are self-adjoint operators.

b) Show that $[L^*, L] = 2i[A, B]$, i.e., L is normal if and only if its real and imaginary part commute.

c) (*Bonus*) Suppose we have proven the spectral theorem (Theorem 2.19) for *Hermitian* matrices. Prove the spectral theorem for normal matrices. That is, if self-adjoint matrices can be *unitarily* diagonalized, the same is true for normal matrices.

Hint: Use (b) and modify the proof of Theorem 2.15.

d) Show that the sum and product of *commuting* normal matrices is normal.

Remark: You may use the conclusion of (c) even if you did not prove it.

Homework 2

Exercise 2.1. Let G be a group and take $h, k \in G$. The (group-wise) *commutator* of h and k is given by

$$[h,k] \coloneqq h^{-1}k^{-1}hk.$$

Notice that $[h,k] = e$ if and only if $hk = kh$, so in particular the commutator in an abelian group is always trivial. One might say that commutators, taken all together, measure how badly a group fails to be abelian.

a) Show that $g[h,k]g^{-1}$ can be written as a commutator.[1] Conclude that the subgroup *generated* by all commutators in G, known as the *commutator subgroup* and written $[G,G]$, is normal in G.

b) The *abelianization* of G is defined as the quotient $G^{\mathrm{ab}} \coloneqq G/[G,G]$. Show that G^{ab} is abelian.[2]

c) Let $\pi : G \twoheadrightarrow G^{\mathrm{ab}}$ be the canonical homomorphism. If $\varphi : G \to A$ is a group homomorphism and A is abelian, show that there is a homomorphism $\tilde{\varphi} : G^{\mathrm{ab}} \to A$ such that $\varphi = \pi \circ \tilde{\varphi}$. This is the *universal property of abelianizations*.

Exercise 2.2. Let G be finite, and let $N \trianglelefteq G$. For any $\rho : G \to \mathrm{GL}(V)$, define the *$N$-invariants* of V as

$$V^N \coloneqq \{x \in X \mid \rho(n)x = x \text{ for all } n \in N\}.$$

a) Prove that V^N is a G-subrepresentation of V.

b) Does V^N need to be trivial as a G-representation, i.e., must $g \cdot x = x$ for all $g \in G$ and $x \in V^N$? Prove or disprove.

Exercise 2.3. Let G be a finite group with $\rho : G \to \mathrm{GL}(V)$.

a) If ρ has degree 2 or 3, show that ρ is irreducible if and only if there is no common eigenvector for every $A \in \operatorname{im}\rho$.

Hint: $3 = 2 + 1$ and $2 = 1 + 1$.

b) Given an example to show the claim is false if the degree of ρ is 4.

Exercise 2.4.

a) Prove that there is a representation $\rho : Q_8 \to \mathrm{GL}_3(\mathbb{C})$ with

$$\rho(i) = \frac{1}{2}\begin{pmatrix} 1-i & 0 & -1-i \\ 0 & 2i & 0 \\ -1-i & 0 & 1-i \end{pmatrix} \quad \text{and} \quad \rho(j) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{2} & 1 \\ -\sqrt{2} & 0 & -\sqrt{2} \\ 1 & \sqrt{2} & -1 \end{pmatrix}$$

b) Decompose ρ into irreducible representations.

[1] A possible approach is using that, for any fixed $g \in G$, the map $\psi_g : G \to G$ given by $\psi_g(h) = ghg^{-1}$ is a homomorphism (indeed, an isomorphism).

[2] In this sense, G^{ab} is the *largest abelian quotient* of G, since we obtained it by collapsing the smallest subgroup containing all commutators.

Exercise 2.5. Here we work towards classifying the finite subgroups of SO(3), i.e., real representations of degree 3. Towards that end, let $G \leq \mathrm{SO}(3)$ be finite and nontrivial, and write $S^2 \subset \mathbb{R}^3$ for the unit sphere. Note that $\mathrm{SO}(3) \circlearrowright S^2$ by left multiplication, since

$$\|Rx\|^2 = \langle Rx, Rx\rangle = \langle x, x\rangle = \|x\|^2 = 1$$

for every $R \in \mathrm{SO}(3)$ and $x \in S^2$. Thus, via inclusion, there is a natural action $G \circlearrowright S^2$.

a) Let $R \in \mathrm{SO}(3)$. Show that R has 1 as an eigenvalue.

Hint: Consider $\det(R - \mathbb{I})$ and recall that $(-1)^3 = -1$.

b) If $R \neq \mathbb{I}$, show that $\mathrm{Fix}(R) \coloneqq \{x \in S^2 \mid Rx = x\}$ consists of *two antipodal points*: $\mathrm{Fix}(R) = \{y, -y\}$.

c) Let $Y \subset S^2$ be the set of elements fixed by at least one non-trivial $g \in G$:

$$Y \coloneqq \bigcup_{g \in G \setminus \{\mathbb{I}\}} \mathrm{Fix}(g)$$

Show that $G \circlearrowright Y$.

d) Decomposing $G \circlearrowright Y$ into distinct orbits $O_1, \ldots, O_r$, use the Orbit-Stabilizer Theorem to show

$$2 - \frac{2}{|G|} = \sum_{i=1}^{r} \left(1 - \frac{1}{k_i}\right),$$

where $k_i = \dfrac{|G|}{|O_i|}$.

e) (*Bonus*) We must have $r \leq 3$ (why?) and we know $2 \leq k_i$ divides $|G|$ (Lagrange's Theorem). What solutions can you find?