Preprint typeset in JHEP style - HYPER VERSION

# Chapter 1: Abstract Group Theory

**Gregory W. Moore**

ABSTRACT: Very Abstract.May 1, 2023

## Contents


**1. Introduction** **5**

**2. Equivalence Relations** **6**

**3. Groups: Basic Definitions And Examples** **10**

**4. Homomorphism And Isomorphism** **27**
- 4.1 Kernel And Image 35

**5. Group Actions On Sets** **37**
- 5.1 Group Actions On Sets 37
- 5.2 Projective Spaces 41
- 5.3 Equivariant Maps 44
- 5.4 Some Special Relations Of $2 \times 2$ Matrix Groups And Symmetries Of Spacetime 45
  - 5.4.1 An Important Homomorphism From $SU(2)$ To $SO(3)$ 45
  - 5.4.2 $SL(2,\mathbb{R})$ And Lorentz Transformations In $2+1$ Dimensions 47
  - 5.4.3 $SL(2,\mathbb{C})$ And $O(1,3)$ 49
  - 5.4.4 $SU(2) \times SU(2)$ and Rotations in 4 Euclidean Dimensions 52
  - 5.4.5 $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ and symmetries of Anti-DeSitter space 53
- 5.5 Group Actions On Sets Induce Group Actions On Associated Function Spaces 54

**6. The Symmetric Group.** **54**
- 6.1 Cayley's Theorem 58
- 6.2 Cyclic Permutations And Cycle Decomposition 59
- 6.3 Transpositions 61
- 6.4 Diversion and Example: Card shuffling 65

**7. Cosets and conjugacy** **69**
- 7.1 Lagrange Theorem 69
- 7.2 Conjugacy 72
- 7.3 Normal Subgroups And Quotient Groups 80
  - 7.3.1 A Very Interesting Quotient Group: Elliptic Curves 93
- 7.4 Quotient Groups And Short Exact Sequences 98
- 7.5 Conjugacy Classes In $S_n$ 106
  - 7.5.1 Conjugacy Classes In $S_n$ And Harmonic Oscillators 108
  - 7.5.2 Conjugacy Classes In $S_n$ And Partitions 121

**8. More About Group Actions And Orbits** **125**
8.1 Left And Right Group Actions 125
8.1.1 More About Induced Group Actions On Function Spaces 128
8.2 Some Definitions And Terminology Associated With Group Actions 129
8.3 The Stabilizer-Orbit Theorem 134
8.4 Practice With Group Action Terminology 139
8.4.1 More About The Relation Of $SU(2)$ And $SO(3)$ 147
8.4.2 Extended Example: The Case Of $1+1$ Dimensions 152
8.4.3 Orbits Of The Lorentz Group In $d > 2$ Dimensions As Homogeneous Spaces 155
8.4.4 Lorentzian And Euclidean (anti)-deSitter Spaces As Homogeneous Spaces 159
8.4.5 Grassmannians 162
8.4.6 The Grassmannian Of Positive Definite Subspaces 165
8.5 Torsors And Principal Bundles 168

**9. Centralizer Subgroups And Counting Conjugacy Classes** **184**
9.1 $0+1$-Dimensional Gauge Theory 186
9.2 Three Mathematical Applications Of The Counting Principle 187

**10. Generators And Relations** **191**
10.1 Example Of Generators And Relations: Fundamental Groups In Topology 199
10.1.1 The Fundamental Group Of A Topological Space 199
10.1.2 Surface Groups: Compact Two-Dimensional Surfaces 207
10.1.3 Braid Groups And Anyons 212
10.1.4 Fundamental Groups Of Three-Dimensional Manifolds 215
10.1.5 Fundamental Groups Of Four-Dimensional Manifolds 217

**11. Some Representation Theory** **217**
11.1 Motivation: Wigner's Theorem In Quantum Mechanics 218
11.2 Some Basic Definitions 225
11.3 Characters 230
11.4 Unitary Representations 231
11.5 Haar Measure, a.k.a. Invariant Integration 232
11.6 The Regular Representation 241
11.7 Reducible And Irreducible Representations 243
11.7.1 Definitions 243
11.7.2 Reducible vs. Completely reducible representations 248
11.8 Schur's Lemmas 252
11.9 Pontryagin Duality 259
11.9.1 An Application Of Pontryagin Duality: Bloch's Theorem And Band Structure 262
11.10 Orthogonality Relations Of Matrix Elements And The Peter-Weyl Theorem 264

11.11 Explicit Decomposition Of A Representation 272
11.11.1 Block diagonalization of Hermitian operators 274
11.11.2 Projecting quantum wavefunctions 274
11.11.3 Finding normal modes in classical mechanics 275
11.12 Orthogonality Relations For Characters And Character Tables 278
11.12.1 Explicit Projection Onto Isotypical Subspaces 279
11.12.2 Finite Groups And The Character Table 281
11.13 Decomposition Of Tensor Products Of Representations And Fusion Coefficients 288
11.14 An Algebraic Viewpoint: The Group Algebra 291
11.14.1 Explicit Decomposition Of The Regular Representation 293
11.14.2 Class Functions And The Center Of The Group Algebra 294
11.14.3 Hopf Algebras 296
11.14.4 Frobenius Algebras 299
11.15 Representations Of The Symmetric Group 299
11.15.1 Standard Tableaux And Their Projection Operators 299
11.16 Schur-Weyl Duality And The Irreps Of $GL(d,\kappa)$ 302
11.17 Orthogonality Relations And Pontryagin Duality 310
11.17.1 The Poisson Summation Formula 311
11.17.2 The Finite Heisenberg Group And The Quantum Mechanics Of A Particle On A Discrete Approximation To A Circle 313
11.18 Induced Representations 318
11.18.1 The Geometrical Interpretation 320
11.18.2 Explicit Characters For A Finite Group $G$ 321
11.18.3 Frobenius Reciprocity 324
11.19 Representations Of $SU(2)$ 326
11.19.1 Homogeneous Polynomials 326
11.19.2 Chararacters Of The Representations $V_j$ 332
11.19.3 Unitarization 334
11.19.4 Inhomogeneous Polynomials And Mobius Transformations On $\mathbb{CP}^1$ 335
11.19.5 The Geometrical Interpretation Of $\mathcal{P}_{2j}$ 339
11.20 The Clebsch-Gordon Decomposition For $SU(2)$ 342
11.21 Lie Groups And Lie Algebras And Lie Algebra Representations 345
11.21.1 Some Useful Formulae For Working With Exponentials Of Operators 345
11.21.2 Lie Algebras 352
11.21.3 The Classical Matrix Groups Are Lie Groups 355
11.21.4 Representations Of Lie Algebras 363
11.21.5 Finite Dimensional Irreducible Representations Of $\mathfrak{sl}(2,\mathbb{C})$,$\mathfrak{sl}(2,\mathbb{R})$, and $\mathfrak{su}(2)$ 365
11.21.6 Casimirs 370
11.21.7 Lie Algebra Operators In The Induced Representations Of $SU(2)$ 372

**12. Group Theory And Elementary Number Theory** **374**
12.1 Reminder On gcd And The Euclidean Algorithm 374
12.2 Application: Expressing elements of $SL(2,\mathbb{Z})$ as words in $S$ and $T$ 379
12.3 Products Of Cyclic Groups And The Chinese Remainder Theorem 381
12.3.1 The Chinese Remainder Theorem 386

**13. The Group Of Automorphisms** **388**
13.1 The group of units in $\mathbb{Z}_N$ 394
13.2 Group theory and cryptography 400
13.2.1 How To Break RSA: Period Finding 403
13.2.2 Period Finding With Quantum Mechanics 403

**14. Semidirect Products** **408**

**15. Group Extensions and Group Cohomology** **423**
15.1 Group Extensions 423
15.2 Projective Representations 431
15.2.1 How projective representations arise in quantum mechanics 433
15.3 How To Classify Central Extensions 434
15.4 Extended Example: Charged Particle On A Circle Surrounding A Solenoid 451
15.4.1 Hamiltonian Analysis 451
15.4.2 Remarks About The Quantum Statistical Mechanics Of The Particle On The Ring 463
15.4.3 Gauging The Global $SO(2)$ Symmetry, Chern-Simons Terms, And Anomalies 469
15.5 Heisenberg Extensions 479
15.5.1 Heisenberg Groups: The Basic Motivating Example 479
15.5.2 Example: The Magnetic Translation Group For Two-Dimensional Electrons 482
15.5.3 The Commutator Function And The Definition Of A General Heisenberg Group 483
15.5.4 Classification Of $U(1)$ Central Extensions Using The Commutator Function 485
15.5.5 Pontryagin Duality And The Stone-von Neumann-Mackey Theorem 486
15.5.6 Some More Examples Of Heisenberg Extensions 491
15.5.7 Lagrangian Subgroups And Induced Representations 499
15.5.8 Automorphisms Of Heisenberg Extensions 501
15.5.9 Coherent State Representations Of Heisenberg Groups: The Bargmann Representation 514
15.5.10 Some Remarks On Chern-Simons Theory 514
15.6 Non-Central Extensions Of A General Group $G$ By An Abelian Group $A$: Twisted Cohomology 514
15.6.1 Crystallographic Groups 519

15.6.2 Time Reversal 523
15.6.3 $T^2 = (-1)^{2j}$ and the Clebsch-Gordon Decomposition 529
15.7 General Extensions 529
15.8 Group cohomology in other degrees 533
15.8.1 Definition 534
15.8.2 Interpreting the meaning of $H^{0+\omega}$ 538
15.8.3 Interpreting the meaning of $H^{1+\omega}$ 538
15.8.4 Interpreting the meaning of $H^{2+\omega}$ 539
15.8.5 Interpreting the meaning of $H^3$ 539
15.9 Some references 541

**16. Overview of general classification theorems for finite groups 541**
16.1 Brute force 542
16.2 Finite Abelian Groups 547
16.3 Finitely Generated Abelian Groups 550
16.4 The Classification Of Finite Simple Groups 552

**17. Categories: Groups and Groupoids 559**
17.1 Groupoids 570
17.2 The topology behind group cohomology 572

**18. Lattice Gauge Theory 576**
18.1 Some Simple Preliminary Computations 576
18.2 Gauge Group And Gauge Field 578
18.3 Defining A Partition Function 581
18.4 Hamiltonian Formulation 586
18.5 Topological Gauge Theory 586

**19. Example: Symmetry Protected Phases Of Matter In $1+1$ Dimensions 591**

---

## 1. Introduction

Historically, group theory began in the early 19th century. In part it grew out of the problem of finding explicit formulae for roots of polynomials. [1]. Later it was realized that groups were crucial in transformation laws of tensors and in describing and constructing geometries with symmetries. This became a major theme in mathematics near the end of the 19th century. In part this was due to Felix Klein's very influential Erlangen program that seeks to define interesting geometries using group theory. Very generally, group theory

[1] For a romantic description, see the chapter "Genius and Stupidity" in E.T. Bell's *Men of Mathematics*. For what is likely a more realistic account see chapter 6 of T. Rothman's *Science à la Mode*. For a scholarly account of the early history of group theory see *The Genesis of the Abstract Group Concept* by H. Wussig.

can be described as the mathematical expression of symmetry. For a very elegant general discussion of this, see Herman Weyl's book, *Symmetry.*

In the 20th century group theory came to play a major role in physics. Einstein's 1905 theory of special relativity is based on the symmetries of Maxwell's equations. The general theory of relativity is deeply involved with the groups of diffeomorphism symmetries of manifolds. With the advent of quantum mechanics the representation theory of groups on linear spaces and in particular Hilbert spaces came to play an important role in atomic physics. In this paradigmatic example of the expression of symmetries in quantum systems the groups $SU(2)$ and $SO(3)$ play a central role. This paradigmatic example is just one example of the general theory of symmetry in quantum mechanics discussed in section ****.

Since the 1950's group theory has played an extremely important role in particle theory. Groups help organize the zoo of subatomic particles. A particularly important example of that is the 8-fold way proposed by Gell-Mann and Ne'eman in 1961 for organizing strongly interacting particles. Moreover, the theory of forces in nature is formulated in terms of gauge theories. In order to formulate the Hamiltonian, Action, or Lagrangian of a gauge theory one must have some understanding of the theory of Lie algebras, Lie groups, and their representations.

In the late 20th and early 21st century group theory has been essential in many areas of physics including atomic, nuclear, particle, and condensed matter physics. However, the beautiful and deep relation between group theory, geometry, and physics is manifested perhaps most magnificently in the areas of mathematical physics concerned with gauge theories (especially supersymmetric gauge theories), quantum gravity, and string theory. These considerations have been central in the choices of topics covered in the following chapters.

Our discussion in these notes will be unabashedly mathematical. There has always been some resistance to increased mathematical sophistication in physics. To quote just one example, the application of the representation theory of $SU(2)$ and $SO(3)$ to atomic physics was referred to by Niels Bohr as "die Gruppenpest." The philosophical background for the choice of topics in these notes is explained more fully in the author's essay "Physical Mathematics and the Future," available on his home page.

Finally, a personal note from the author: I would like to make two requests of the reader: First, much of what follows is standard textbook material, and will be presented in similar ways in many other books, lecture notes, and articles available on the internet. But some is nonstandard. If you make use of any of the nonstandard material in something you write, please give proper acknowledgement to these notes. Second, if you find any mistakes in these notes please do not hesitate to send me an email. (But please, first do check carefully it really is a mistake.) These notes are a work in progress and are continually being updated. Thank you - Gregory Moore.

## 2. Equivalence Relations

The following ideas are very elementary, but very basic and will be used repeatedly through-out these notes. A good reference for this elementary material is I.N. Herstein, *Topics in Algebra*, sec. 1.1.

Quite generally, if $X$ and $Y$ are any two sets, we say that a *relation* $R$ is a subset of the Cartesian product $X \times Y$:

$$R \subset X \times Y \tag{2.1}$$

The idea here is that if an ordered pair $(x, y) \in X \times Y$ is in $R$, i.e. $(x, y) \in R$ then we say that "$x$ is related to $y$ by $R$." For example, if $f : X \to Y$ is a function, then the graph of the function $R = \{(x, f(x)) : x \in X\} \subset X \times Y$, is a relation. But there are other relations where, for example, there might be many points $y \in Y$ that are related to a given $x \in X$.

If $Y = X$ there is a special kind of relation known as an *equivalence relation.* By definition, an equivalence relation is a binary relation $R$ satisfying the following three conditions:

1. For all $x \in X$ we have $(x, x) \in R$.
2. *Symmetry*: [2] If $(x, y) \in R$ then $(y, x) \in R$.
3. *Transitivity*: If $(x, y) \in R$ and $(y, z) \in R$ then $(x, z) \in R$.

We often denote an equivalence relation by $\sim$. Thus $x \sim y$ means the same thing as $(x, y) \in R$. Written in this notation we can say that a binary relation $\sim$ is an equivalence relation if, $\forall a, b, c \in X$:

1. $a \sim a$
2. $a \sim b \Rightarrow b \sim a$
3. $a \sim b$ and $b \sim c \Rightarrow a \sim c$

**Examples**:

**Example 2.1** : The notion of equality satisfies these axioms of an equivalence relation. So $a \sim b$ iff $a = b$ is an equivalence relation. The main point, however, is that an equivalence relation is a more flexible notion than equality, and yet captures many of the important aspects of equality.

**Example 2.2** : $X = \mathbb{Z}$, $a \sim b$ if $a - b$ is even.

**Example 2.3** : More generally, let $X = \mathbb{Z}$, and choose a positive integer $N$. We can define an equivalence relation by saying that $a \sim b$ iff $a - b$ is divisible by $N$.

**Example 2.4** : At the other extreme from equality we could say that every element of the set $X$ is equivalent to every other element.

**Definition 2.2**: Let $\sim$ be an equivalence relation on $X$.

---

[2] This is the first of many, many, many uses of the word "symmetry" that will appear in these notes."

a.) An *equivalence class* in $X$ is a subset $\mathcal{O} \subset X$ such that for any $x \in \mathcal{O}$, if $x \sim y$ then $y \in \mathcal{O}$ and, moreover, every pair of elements in $\mathcal{O}$ are related.

b.) If $\mathcal{O} \subset X$ is an equivalence class of an equivalence relation then any element $x \in \mathcal{O}$ is called a *representative of the class* $\mathcal{O}$.

c.) The *equivalence class* associated to an element $x \in X$ is the subset

$$[x] := \{y \in X : x \sim y\} \subset X \tag{2.2}$$

**Important Remark**: In general there will be several different elements in an equivalence class. So, in general, if $[x] = [x']$ is does not follow that $x = x'$! For any equivalence class $\mathcal{O}$ we can find some representative and present it as $\mathcal{O} = [x]$, but it is a very common fallacy to conclude that just because an equivalence class is presented as $\mathcal{O} = [x]$ the equivalence class itself uniquely determines the element $x$. In other words, given an equivalence class, in general there is no way to uniquely determine a particular element $x \in X$. We will be coming back to this elementary point several times in these notes.

In the above two examples we have

**Example 2.1'** : If our equivalence relation is just equality then the equivalence class of every element has only one element: $[a] = \{a\}$. The set of equivalence classes is in bijective correspondence with the original set.

**Example 2.2'** : $[n]$ is the set of all integers with the same parity as $n$. For example,
$[1] = \{n : n \text{ is an odd integer}\}$
$[4] = \{n : n \text{ is an even integer}\}$.

**Example 2.3'** : Consider the equivalence relation $a \sim b$ iff $a - b$ is divisible by $N$. Recall that if $n$ is an integer then we can write $n = r + Nq$ in a unique way where the quotient $q$ is integral and the *remainder* or *residue modulo* $N$ is an integer $r \in \{0, 1, \dots, N-1\}$. Thus there is a bijective correspondence between the set of equivalence classes and the set $\{0, 1, \dots, N-1\}$. The equivalence class of an integer $n$ will sometimes be written as $\bar{n}$. One way to write it is

$$\bar{n} := n + N\mathbb{Z} := \{\dots, n - 2N, n - N, n, n + N, n + 2N, \dots\} \tag{2.3}$$

**Example 2.4'** : If $R = X \times X$, so that every $x$ is related to every $y$ then then there is only one equivalence class, namely the full set $X$ itself.

Here is a simple, but basic, principle:

| |
|---|
| The distinct equivalence classes of an equivalence relation on $X$ decompose $X$ into a union of mutually disjoint subsets. Conversely, given a *disjoint* decomposition $X = \amalg X_i$ we can define an equivalence relation by saying $a \sim b$ if $a, b \in X_i$. |

A disjoint decomposition of a set $X$ is sometimes called a *partition of* $X$. Thus we are claiming that there is a one-one correspondence between partitions of a set $X$ and equivalence relations on $X$.

We leave the easy proof of the above principle to the reader. As an example, the integers are the disjoint union of the even and odd integers, and the corresponding equivalence relation is the one mentioned above: $a \sim b$ iff $a - b$ is even.

---

**Exercise** *Due Diligence*

Prove that there is a one-one correspondence between partitions of a set $X$ and equivalence relations on $X$. This exercise is easy but very important. If you get stuck see Herstein's book.

---

---

**Exercise** *Examples*

In each of the examples above describe the partition of the set $X$.

---

---

**Exercise** *Equivalence Relations Which Are Graphs*

Show that if an equivalence relation $R$ is the graph of a function then that equivalence relation is just that of equality. [3]

---

---

**Exercise** *Another Characterization Of Equivalence Classes*

Suppose we view an equivalence relation as subset $R \subset X \times X$. Let $\pi_1 : X \times X \to X$ be the projection to the first factor and $\pi_2 : X \times X \to X$ be the projection on the second factor. Show that the equivalence classes in $X$ are just the sets which can be written as $\pi_i(\pi_j^{-1}(x))$ for some $x \in X$.

---

---

**Exercise** *Equivalence Relations And Fibers Of Maps*

In general, given a map $\pi : X \to Y$ we say that the *fiber above* $y \in Y$ or *preimage of* $y$ is the subset $\pi^{-1}(y) \subset X$ of elements in $X$ that map to $y$ under $\pi$.

a.) Let $p : X \to Y$ be a surjective map. Show that

$$R(p) := \{(x, x')|p(x) = p(x')\} \subset X \times X \tag{2.4}$$

is an equivalence relation.

[3] *Answer*: If $R$ is the graph of $f : X \to X$ then $(x, x) \in R$ implies that $f(x) = x$.

b.) Given an equivalence relation $R$ on $X$ denote the set of equivalence classes by $X/R$. Show that there is a map $p_R : X \to X/R$ and

$$R(p_R) = R \tag{2.5}$$

---

---

**Exercise** *Comparison Of Equivalence Relations*

We say that a partition $X = \amalg_\alpha Y_\alpha$ *refines* a partition $X = \amalg_i X_i$ if each $Y_\alpha$ is a subset of some $X_i$ and each $X_i$ is itself partitioned into a collection of sets from the collection of $Y_\alpha$.

Suppose $\sim_1$ is associated to $X = \amalg_\alpha Y_\alpha$ and $\sim_2$ is associated to $X = \amalg_i X_i$. We say that "$\sim_1$ is finer than $\sim_2$" and "$\sim_2$ is coarser than $\sim_1$."

a.) Show that if $\sim_1$ is finer than $\sim_2$ then $x \sim_1 y$ implies that $x \sim_2 y$.

b.) Show that the finest equivalence relation is equality.

c.) Show that the coarsest equivalence relation has $x \sim y$ for all $x, y \in X$.

As an example: The equivalence relation defined by equality modulo $N_1$ refines that defined by equality modulo $N_2$ if $N_1$ divides $N_2$. For example, equivalence modulo 4 refines equivalence modulo 2.

---

## 3. Groups: Basic Definitions And Examples

We begin with the abstract definition of a group.

**Definition 3.1**: A *group* is a quartet $(G, \mathbf{m}, \mathbf{I}, e)$ where

1. $G$ is a set.
2. $\mathbf{m} : G \times G \to G$ is a map, called the *group multiplication map.*
3. $\mathbf{I} : G \to G$ is a map, called the *inverse map*
4. $e \in G$ is a distinguished element of $G$ called the *identity element.*

These data $(G, \mathbf{m}, \mathbf{I}, e)$ are required to satisfy the following conditions:

1. $\mathbf{m}$ is *associative*: For all $g_1, g_2, g_3 \in G$ we have

$$\mathbf{m}(\mathbf{m}(g_1, g_2), g_3) = \mathbf{m}(g_1, \mathbf{m}(g_2, g_3)) \tag{3.1}$$

2.

$$\forall g \in G \qquad \mathbf{m}(g, e) = \mathbf{m}(e, g) = g \tag{3.2}$$

3.
$$\forall g \in G \qquad \mathbf{m}(\mathbf{I}(g), g) = \mathbf{m}(g, \mathbf{I}(g)) = e \tag{3.3}$$

The above notation is unduly heavy, and we will use it sparingly. We will use it when comparing different group multiplication laws on the same set, or when a multiplication on $X$ induces another one on a different set. Thus, we give the definition again, but more informally:

$\forall a, b \in G$ *there exists a unique element in* $G$*, called the product, and denoted* $a \cdot b \in G$

in other words, we streamline notation by writing $a \cdot b := \mathbf{m}(a, b)$. Eventually, we will drop the $\cdot$ and just write $ab$ for the group product of two elements $a, b$ in a group.

The product is required to satisfy 3 axioms:

1. Associativity: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$

2. Existence of an identity element: $\exists e \in G$ such that:
$$\forall a \in G \qquad a \cdot e = e \cdot a = a \tag{3.4}$$

3. Existence of inverses: Again, we streamline notation by writing $a^{-1} := \mathbf{I}(a)$. so that $a \cdot a^{-1} = a^{-1} \cdot a = e$

Often one speaks of a "group $G$" leaving the extra data of the multiplication, inverse, and identity implicit.

**Remarks**

1. We will often denote $e$ by 1, or, when discussing more than one group at a time, we denote the identity in a particular group $G$ by $1_G$. The identity element is also often called the *unit element*. If the set in question has more than one binary operation in play, e.g. in a ring, one needs to be careful when speaking of "the identity" or a "unit" to specify which operation is being referred to.

2. We can drop some axioms and still have objects of mathematical interest. For example, if we drop the existence of inverses the above properties define a *monoid*. That is, a monoid is a set $M$ with a multiplication map $\mathbf{m} : M \times M \to M$ which is associative such that there is an element $e \in M$ which functions as the identity for this multiplication. One can drop other combinations of the group axioms and define other mathematical objects, but in these cases the terminology is not very consistent in the literature. When proceeding from monoids to groups the further assumption of the existence of inverses turns the monoid into a group. The definition of a group seems to be in the Goldilocks region of having just enough data and conditions to allow a deep theory, but not having too many constraints to allow only a few examples. It is just right to have a deep and rich mathematical theory, together with a dazzling universe of examples.

3. We can also put further mathematical structures on the data $(G, \mathbf{m}, \mathbf{I}, e)$ (and still have a rich theory) to define important special classes of groups. For example, a *topological group* is a group $(G, \mathbf{m}, \mathbf{I}, e)$ such that $G$ is a topological space and $\mathbf{m}$ and $\mathbf{I}$ are both continuous maps of topological spaces. Similarly, a *Lie group* is a group $(G, \mathbf{m}, \mathbf{I}, e)$ such that $G$ is a manifold and $\mathbf{m}$ and $\mathbf{I}$ are real analytic in real analytic local coordinates. [4]

---

**Exercise** *Inverses And Identities In Groups Are Unique*

a.) Show that $e$ unique. [5]

b.) Given $a$ is $a^{-1}$ unique? [6]

c.) Show that axioms 2,3 above are slightly redundant: For example, just assuming $a \cdot e = a$ and $a \cdot a^{-1} = e$ show that $e \cdot a = a$ follows as a consequence.

---

Given two groups $(G_1, \mathbf{m}_1, \mathbf{I}_1, e_1)$ and $(G_2, \mathbf{m}_2, \mathbf{I}_2, e_2)$ it turns out that, in some cases, there can be many ways to use this data to define a group multiplication on the Cartesian product $G_1 \times G_2$. But there is always one canonical way this can be done:

**Definition 3.4** Let $G_1, G_2$ be two groups. The *direct product* of $G_1, G_2$ is the set $G_1 \times G_2$ with product:

$$\mathbf{m}_{G_1 \times G_2}\left((g_1, g_2), (g_1', g_2')\right) = (\mathbf{m}_{G_1}(g_1, g_1'), \mathbf{m}_{G_2}(g_2, g_2')) \tag{3.5}$$

and inverse:

$$\mathbf{I}_{G_1 \times G_2}((g_1, g_2)) = (\mathbf{I}_{G_1}(g_1), \mathbf{I}_{G_2}(g_2)) \tag{3.6}$$

---

**Exercise** *Due Diligence: Direct Product Of Groups*

a.) Check the group axioms.

[4] We will be informal about our treatment of manifolds until we study Lie groups in a separate chapter. A nice example (from mathstackexchange) of a topological group which is not a Lie group is the rational numbers with the induced topology from $\mathbb{R}$. This is a topological group: Addition and inversion of rational numbers is continuous in the induced topology from $\mathbb{R}$. But it is not a Lie group because, in the induced topology, the neighborhood of any point is not homomeomorphic to $\mathbb{R}^n$ for any $n$, so it is not a manifold. Other examples are matrix groups over $p$-adic numbers.

[5] *Answer*: Suppose that two elements $e_1, e_2 \in G$ behave as units. Consider the product $e_1 \cdot e_2$. Using $e_1$ as a unit we can say this is $e_2$. On the other hand, using $e_2$ as a unit we can say this is $e_1$. Therefore $e_1 = e_2$.

[6] Suppose $a \cdot b_1 = e$ and $a \cdot b_2 = e$. Then $a \cdot b_1 = e$ implies $b_1 \cdot a = e$. Therefore $b_1 \cdot (a \cdot b_2) = b_1 \cdot e = b_1$. But $b_1 \cdot (a \cdot b_2) = (b_1 \cdot a) \cdot b_2 = e \cdot b_2 = b_2$. Therefore $b_1 = b_2$.

b.) Generalize this to arbitrary products: Given a map $\mathcal{G}$ from a set $I$ to the set of all groups define the product over $I$ as a group.

**Remark**: We will explore other ways of defining group structures on the Cartesian product $G_1 \times G_2$ of sets in great detail in sections **** below.

---

**Example 2.1**: As a set, $G = \mathbb{Z}, \mathbb{R}$, or $\mathbb{C}$. The group operation is ordinary addition:

$$\mathbf{m}(a,b) := a + b \tag{3.7}$$

The reader should check all the axioms.

**Example 3.2**: Now with the above example we can make new groups by considering the direct product of groups. For example, we could take $n$-tuples for a positive integer $n$: $G = \mathbb{Z}^n, \mathbb{R}^n, \mathbb{C}^n$, with the operation being vector addition, so if $\vec{x} = (x_1, \dots, x_n)$ and $\vec{y} = (y_1, \dots, y_n)$ then

$$\mathbf{m}(\vec{x}, \vec{y}) := (x_1 + y_1, \dots, x_n + y_n) \tag{3.8}$$

**Example 3.3**: $G = \mathbb{R}^* := \mathbb{R} - \{0\}$ or $G = \mathbb{C}^* := \mathbb{C} - \{0\}$ Now if $x, y \in G$ then $\mathbf{m}(x,y) := xy$ is ordinary multiplication of complex numbers. Check the axioms.

**Definition 3.2**: Suppose $(G, \mathbf{m}, \mathbf{I}, e)$ is a group and $H \subset G$ is a subset so that $\mathbf{m}$ and $\mathbf{I}$ preserve $H$, that is, the restriction of $\mathbf{m}$ takes $H \times H \to H$ and the restriction of $\mathbf{I}$ maps $H \to H$. (It then follows that $e \in H$.) In this case we say that $(H, \mathbf{m}, \mathbf{I}, e)$ is a *subgroup of* $(G, \boldsymbol{m}, \boldsymbol{I}, e)$.

---

**Exercise** *Subgroups*

a.) $\mathbb{Z} \subset \mathbb{R} \subset \mathbb{C}$ with operation $+$, define subgroups.

b.) Is $\mathbb{Z} - \{0\}$ a monoid (with $\mathbf{m}$ given by standard multiplication) ?

c.) Is $\mathbb{Z} - \{0\} \subset \mathbb{R}^*$ a subgroup?

d.) Let $\mathbb{R}^*_{>0}$ and $\mathbb{R}^*_{<0}$ denote the positive and negative real numbers, respectively. Using ordinary multiplication of real numbers, which of these are subgroups of $\mathbb{R}^*$?

e.) Consider the negative real numbers $\mathbb{R}_{<0}$ with the multiplication rule:

$$\mathbf{m}(x,y) = -xy \tag{3.9}$$

Show that this defines a group law on $\mathbb{R}_{<0}$, but that $(\mathbb{R}_{<0}, \mathbf{m}, ...)$ is <u>not</u> a subgroup of $\mathbb{R}^*$. [7]

---

---

[7] *Answer*: The point is that the multiplication of (3.9) is <u>not</u> the restriction of the multiplication on $\mathbb{R}^*$ to the negative reals. Indeed, note that the identity element is the real number $-1$.

**Exercise** *Intersections And Unions Of Subgroups*

Suppose $H_1 \subset G$ and $H_2 \subset G$ are two subgroups.

a.) Show that $H_1 \cap H_2$ is a subgroup of $G$.

b.) Is it always true that $H_1 \cup H_2$ is a subgroup of $G$ ?

c.) For an integer $\ell$ let $\ell\mathbb{Z} \subset \mathbb{Z}$ be the subset of integer multiples of $\ell$. Show that $\ell\mathbb{Z}$ is a subgroup.

d.) What is $2\mathbb{Z} \cap 3\mathbb{Z}$ ? [8]

e.) Is $2\mathbb{Z} \cup 3\mathbb{Z}$ a subgroup ? [9]

---

**Definition 3.3**: The *order* of a group $G$, denoted $|G|$, is the cardinality of $G$ as a set. Roughly speaking this is the same as the "number of elements in $G$." A group $G$ is called a *finite group* if $|G| < \infty$, and is called an *infinite group* otherwise.

Note that the direct product of two finite groups is finite. Already, with the simple concepts we have just introduced, we can ask nontrivial questions. For example:

*Does every infinite group necessarily have proper subgroups of infinite order?*

This is of course true of the examples we have just discussed. It is actually not easy to think of counterexamples, but in fact there are infinite groups all of whose proper subgroups are finite. [10]

Let us continue with an overview of examples of groups:

The groups in Examples 1,2,3 above are of infinite order. Here are examples of finite groups:

**Example 2.4: The group of $N^{th}$ roots of unity.** Choose a natural number $N$. [11] We let $\mu_N$ be the set of complex numbers $z$ such that $z^N = 1$. Thus we could write

$$\mu_N = \{1, \omega, \dots, \omega^{N-1}\} \tag{3.10}$$

where $\omega = \exp[2\pi i/N]$. This is a finite group with $N$ elements, as is easily checked.

---

[8] *Answer* $6\mathbb{Z}$.

[9] *Answer*: No. For example $2 + 3 = 5$ is not in the union, so the union is not closed under the group operation.

[10] One example are the *Prüfer groups*. These are subgroups of the group of roots of unity. They are defined by choosing a prime number $p$ and taking the subgroup of roots of unity of order $p^n$ for some natural number $n$. Even wilder examples are the "Tarski Monster groups" (not to be confused with the Monster group, which we will discuss later). These are infinite groups all of whose subgroups are isomorphic to the cyclic group of order $p$.

[11] The *natural numbers* are the same as the positive integers.

**Exercise**

a.) Show, more generally, that $\mu_N \cap \mu_{N+1} = \{1\}$.

b.) Show that if $N$ is even then the group of $N/2^{th}$ roots of unity is a subgroup of $\mu_N$.

c.) When is $\nu_N \cap \nu_M$ a nontrivial group (i.e. a group with more than one element)?

---

**Example 3.5: The residue classes modulo $N$, also called "The cyclic group of order N."** Let $N$ be a positive integer. Recall that we can put an equivalence relation on $\mathbb{Z}$ defined by $a \sim b$ iff $a - b$ is divisible by $N$, and we denoted the class of an integer $n$ by $\bar{n}$. (As mentioned above, the equivalence classes are in bijective equivalence with the elements of the set $\{0, 1, \ldots, N-1\}$.) We take $G$ to be the set of equivalence classes of integers modulo $N$. We need to define $\mathbf{m}(\bar{r}_1, \bar{r}_2)$. To do this we choose a representative $r_1$, $r_2$ from each equivalence class and take

$$\mathbf{m}(\bar{r}_1, \bar{r}_2) := \overline{(r_1 + r_2)} \tag{3.11}$$

The main thing to check here is that the equation is well-defined, since we chose representatives for each equivalence class. This group, which appears frequently in the following, will be denoted as $\mathbb{Z}_N$ or, better, $\mathbb{Z}/N\mathbb{Z}$. For example, telling railroad/military time in hours is arithmetic in $\mathbb{Z}_{24}$. The reader should note that $\mathbb{Z}_N$ "resembles" closely the group $\mu_N$. We will make that precise in the next section.

---

**Exercise**

a.) Show that, if $N$ is even then the subset of equivalence classes $\bar{r}$ with representatives $r$ which are even forms a subgroup of $\mathbb{Z}_N$.

b.) What can you say about the subgroups of $\mathbb{Z}_N$ when $N$ is odd? [12]

---

So far, all the examples we have discussed have the property that for any two elements $a, b \in G$

$$\mathbf{m}(a, b) = \mathbf{m}(b, a) \ . \tag{3.12}$$

**Definition 3.5**: When equation (3.12) holds for two elements $a, b \in G$ we say *"a and b commute."* If $a$ and $b$ commute for every pair $(a, b) \in G \times G$ then we say that $G$ is an *Abelian group*:

| If $a, b$ commute for all $a, b \in G$ we say "*G is Abelian.*" |
|---|

**Note**: Note that our abbreviated notation $a \cdot b$ for the group multiplication $\mathbf{m}(a, b)$ would actually be quite confusing when working with $\mathbb{Z}_N$. The reason is that it is also possible to define a *ring structure* (see Chapter 2) where one multiplies $r_1$ and $r_2$ as integers and then

[12] For the answer, see the section on Lagrange's theorem below.

takes the residue. This is *NOT* the same as $\mathbf{m}(r_1, r_2)$ !! For example, if we take $N = 5$ then $\mathbf{m}(2,3) = 0$ in $\mathbb{Z}_5$ because $2 + 3 = 5$ is congruent to 0 modulo 5. Of course, multiplying as integers $2 \times 3 = 6$ and 6 is congruent to 1 mod 5. Therefore, when considering Abelian groups we often prefer to use the "additive" notation

$$a + b := \mathbf{m}(a, b) \tag{3.13}$$

When we use this additive notation for Abelian groups we will write the identity element as 0 so that $a + 0 = 0 + a = a$. (Writing "$a + 1 = a$" would look extremely weird.) Note that we will not always use additive notation for Abelian groups! For example, for $\mu_N$ the multiplicative notation is quite natural. When using multiplicative notation we will not use 0 for the identity because writing "$0 \cdot a = a$" would also look extremely wierd.

Since we defined a notion of "Abelian group" we are implicitly suggesting there are examples of groups which are not Abelian. If one tries to use the group axioms to prove that $\mathbf{m}(g_1, g_2) = \mathbf{m}(g_2, g_1)$ one will fail. The only way we can know conclusively that one will fail is to provide a counterexample. The next set of examples are important classes of nonabelian groups:

**Example 3.6**: *The General Linear Group*

Let $\kappa = \mathbb{R}$ or $\kappa = \mathbb{C}$. Define $M_n(\kappa)$ to be the set of all $n \times n$ matrices whose matrix elements lie in $\kappa$. Note that this is a unital monoid under matrix multiplication: i.e. matrix multiplication is associative. But $M_n(\kappa)$ is not a group, because some matrices are not invertible. Therefore we define:

♣$\kappa$ will be our official symbol for a general field. This needs to be changed from $k$ in many places below. ♣

$$GL(n, \kappa) := \{A | A = n \times n \text{ invertible matix over } \kappa\} \subset M_n(\kappa) \tag{3.14}$$

When $\kappa = \mathbb{R}$ or $\kappa = \mathbb{C}$ $GL(n, \kappa)$ is a group of infinite order. It is Abelian if $n = 1$ and nonabelian if $n > 1$.

**Remark**: There are some important generalizations of this example: [13] We could let $\kappa$ be any field. If $\kappa$ is a finite field then $GL(n, \kappa)$ is a finite group. More generally, if $R$ is a *ring* $GL(n, R)$ is the subset of $n \times n$ matrices with entries in $R$ with an inverse in $M_n(R)$. This set forms a group. For example, $GL(n, \mathbb{Z})$ is the set of $n \times n$ matrices of integers such that the inverse matrix is also a $n \times n$ matrix of integers. This is the same set as the set of $n \times n$ matrices with integer entries whose determinant is $\pm 1$. [14] This set of matrices forms an infinite nonabelian group under matrix multiplication.

**Definition 3.5**: The center $Z(G)$ of a group $G$ is the set of elements $z \in G$ that commute with all elements of $G$:

$$Z(G) := \{z \in G | zg = gz \qquad \forall g \in G\} \tag{3.15}$$

[13] See Chapter 2 for some discussion of the mathematical notions of fields and rings used in this paragraph.

[14] In general $GL(R)$ can be characterized as the $n \times n$ matrices with matrix elements in $R$ whose determinant is a unit in $R$.

---

**Exercise** *Due Diligence: The Center* :

a.) Show that for any group $G$, $Z(G)$ is an Abelian subgroup of $G$.

b.) Show that the center of $GL(n,\kappa)$ is the subgroup of matrices proportional to the unit matrix with scalar factor in $\kappa^*$. [15]

---

**Example 3.7**: *Some Standard Matrix Groups*

A *matrix group* is a subgroup of $GL(n,\kappa)$. There are several interesting examples which we will study in great detail later. Some examples include:

The special linear group:

$$SL(n,\kappa) \equiv \{A \in GL(n,\kappa) : \det A = 1\} \tag{3.16}$$

The orthogonal and special orthogonal groups:

$$\begin{aligned} O(n,\kappa) &:= \{A \in GL(n,\kappa) : AA^{tr} = 1\} \\ SO(n,\kappa) &:= \{A \in O(n,\kappa) : \det A = 1\} \end{aligned} . \tag{3.17}$$

Another natural class are the unitary and special unitary groups:

$$U(n) := \{A \in GL(n,\mathbb{C}) : AA^{\dagger} = 1\} \tag{3.18}$$

$$SU(n) := \{A \in U(n) : \det A = 1\} . \tag{3.19}$$

**Remarks**:

1. As an exercise you should show from the definition above that the most general element of $SO(2,\mathbb{R})$ must be of the form

$$\begin{pmatrix} x & y \\ -y & x \end{pmatrix} \qquad x^2 + y^2 = 1 \tag{3.20}$$

---

[15] *Answer*: It is obvious that matrices of the form $z1_{n\times n}$ with $z \in \kappa^*$ are in the center. What is not immediately obvious is that there are no other elements of the center. Here is a careful proof that this is indeed the case: Consider the *matrix units*: $e_{ij}$. The matrix $e_{ij}$ has a 1 in the $i^{th}$ row and $j^{th}$ column and zeroes elsewhere. Note that for any matrix $A$ we have $e_{ii}Ae_{jj} = A_{ij}e_{ij}$ with no sum on $i,j$ here: on the RHS $A_{ij}$ is a matrix element, not a matrix. Now let $z$ be in the center. Check that for any pair $ij$ the matrix $1 + e_{ij}$ is invertible. Therefore, if $z$ is in the center then $z$ must commute with $1 + e_{ij}$ and hence $z$ must commute with $e_{ij}$ for all $i,j$. Now, as we observed above, $e_{ii}ze_{jj} = z_{ij}e_{ij}$ holds for any matrix, but since $z$ is also central $e_{ii}ze_{jj} = e_{ii}e_{jj}z = \delta_{ij}e_{jj}z$. So $z$ is diagonal. But for any diagonal matrix $z = \sum_k z_k e_{kk}$ we have $(zA)_{ij} = z_i A_{ij}$ and $(Az)_{ij} = A_{ij}z_j$. As long as there are matrices with $A_{ij} \neq 0$ and invertible we can conclude that $z_i = z_j$.

where the matrix elements $x, y$ are real. Thus we recognize that group elements in $SO(2, \mathbb{R})$ are in 1-1 correspondence with points on the unit circle in the plane. We can even go further and parametrize $x = \cos\phi$ and $y = \sin\phi$ and $\phi$ is a coordinate provided we identify $\phi \sim \phi + 2\pi$ so the general element of $SO(2, \mathbb{R})$ is of the form:

$$R(\phi) := \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \tag{3.21}$$

This is familiar from the implementation of rotations of the Euclidean plane in Cartesian coordinates. Note that the group multiplication law is

$$R(\phi_1)R(\phi_2) = R(\phi_1 + \phi_2) \tag{3.22}$$

so, in $\phi$ "coordinates" the group multiplication law is continuous, differentiable, even (real) analytic. Similarly, in these coordinates the inverse map is $\phi \to -\phi$, a real analytic transformation.

2. Let us consider the group $U(1)$: This is simply the group of $1 \times 1$ unitary matrices. They are not hard to diagonalize. The general matrix can be written as $z(\phi) = e^{i\phi}$ with multiplication $z(\phi_1)z(\phi_2) = z(\phi_1 + \phi_2)$ where $\phi \sim \phi + 2\pi$ yield identical group elements. Again, as with $\mu_N$ and $\mathbb{Z}_N$ the groups look like they are "the same" although strictly speaking they are different sets and therefore have different **m**'s. We will make this idea precise in the next section.

3. One of the most important groups in both mathematics and physics is $SU(2)$. Suppose we have a pair of complex numbers $(z, w) \in \mathbb{C}^2$ such that

$$|z|^2 + |w|^2 = 1 \tag{3.23}$$

Then one easily checks that

$$g = \begin{pmatrix} z & -w^* \\ w & z^* \end{pmatrix} \in SU(2) \tag{3.24}$$

We claim that, conversely, every element of $SU(2)$ can be written in this way. One can prove this by studying the 4 equations for the matrix elements in the identity $gg^\dagger = 1$. Another way to proceed makes use of some concepts from the linear algebra chapter below and goes as follows: Since $g$ is unitary it follows that the basis

$$g\begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad g\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.25}$$

Should be orthonormal. Therefore, if we write

$$g = \begin{pmatrix} z & u \\ w & v \end{pmatrix} \tag{3.26}$$

it must be that $\begin{pmatrix} z \\ w \end{pmatrix}$ is orthogonal to $\begin{pmatrix} u \\ v \end{pmatrix}$ and hence $u = -\lambda w^*$ and $v = \lambda z^*$ for some complex number $\lambda$. Moreover, since norms are preserved we know that $|z|^2 + |w|^2 = 1$ and $|u|^2 + |w|^2 = 1$ and hence $|\lambda|^2 = 1$, so $\lambda$ is a phase. So the general unitary matrix must be of the form

$$g = \begin{pmatrix} z & -\lambda w^* \\ w & \lambda z^* \end{pmatrix} \tag{3.27}$$

for some phase $\lambda$. But now if we impose $\det g = 1$ we discover that $\lambda = 1$.

**Comment: The Classical Matrix Groups Are Examples Of Lie Groups** We have noted that $SO(2)$ and $U(1)$ can be identified with the circle $S^1$. Similarly, we have shown there is a 1-1 correspondence between elements of $SU(2)$ and pairs of complex numbers $(z, w)$ with $|z|^2 + |w|^2 = 1$. If we decompose $z, w$ into real and imaginary parts:

$$\begin{aligned} z &= x_1 + \sqrt{-1}x_2 \\ w &= x_3 + \sqrt{-1}x_4 \end{aligned} \tag{3.28}$$

then the equation $|z|^2 + |w|^2 = 1$ is equivalent to the equation

$$\sum_{\mu=1}^{4} (x_\mu)^2 = 1 \tag{3.29}$$

which we can recognize as defining the unit three-dimensional sphere $S^3$ in $\mathbb{R}^4$. Later on, we will give various coordinate systems for $S^3$ making clear that the group multiplication and inverse operations are real analytic. So $SU(2)$ is a Lie group. In view of these two examples one might wonder if other Lie groups are spheres and if other spheres are Lie groups. It is not obvious, but in fact, no other spheres are groups (except for $\mathbb{Z}_2$ which could be considered the 0-dimensional sphere). This is a deep result of topology. So we should not think of Lie groups as spheres. It turns out that all the classical matrix groups we have mentioned above are examples of Lie groups. To prove they are manifolds one views the defining equations such as $A^{tr}A = 1$ as a set of equations on the matrix elements and shows that the solutions to these equations in $M_n(\kappa) \cong \kappa^{n^2}$ is a smooth manifold. In coordinates obtained from the matrix elements the group multiplication and inverse are real analytic functions.

Lie groups have vast applications in physics. For example, $G = SU(3)$ is the gauge group of a Yang-Mills theory that describes the interactions of quarks and gluons, while $G = SU(3) \times SU(2) \times U(1)$ is related to the standard model that describes all known elementary particles and their interactions. The general theory of Lie groups will be discussed in Chapter 8(?) below, although we will meet many many examples before then.

**Example 3.8**: *Some Groups Defined By Bilinear Forms* It is interesting to try to generalize the definition $A^{tr}A = 1$ of the orthogonal groups as follows: Suppose $b$ is an $n \times n$

matrix and we consider the set of matrices satisfying $A^{tr}bA = b$. Since we want this equation to imply that $A$ is invertible it is wise to restrict to $b$'s which are invertible. Taking the transpose of this equation suggests that we should also take $b^{tr} = \lambda b$, where $\lambda \in \kappa$ is a scalar. Then consistency forces $\lambda^2 = 1$ so $\lambda = \pm 1$. So we are naturally led to symmetric or antisymmetric "bilinear forms" [16] In either case, when $b$ is invertible and either symmetric or antisymmetric the "automorphism group of the bilinear form"

$$\mathrm{Aut}(b) := \{A \in M_n(\kappa) | A^{tr} b A = b\} \tag{3.30}$$

is a group. We will justify the notation on the LHS in our discussion of group actions below.

If $b$ is symmetric we can diagonalize it. (See Chapter 2). One particularly nice case is

$$b = \eta = \begin{pmatrix} -1 & 0 \\ 0 & 1_{n\times n} \end{pmatrix} \tag{3.31}$$

In this case the set of matrices $O(1, n) := \{A | A^{tr} \eta A = \eta\}$ is known as the *Lorentz group of Minkowski spacetime in* $1 + n$ *dimensions* and it plays an important role in relativistic physics. More generally, if

$$b = \begin{pmatrix} -1_{p\times p} & 0 \\ 0 & 1_{q\times q} \end{pmatrix} \tag{3.32}$$

then $Aut(b)$ is denoted $O(p, q)$. If we work with complex matrices and replace $A^{tr}$ by $A^\dagger$ then we define the groups $U(p, q)$.

If $b$ is antisymmetric it can be put into the standard form (again, see 'Linear Algebra Users Manual" ):

$$J = \begin{pmatrix} 0 & 1_{n\times n} \\ -1_{n\times n} & 0 \end{pmatrix} \in M_{2n}(\mathbb{R}) \tag{3.33}$$

which is sometimes called the standard symplectic form on $\mathbb{R}^{2n}$. Note that the matrix $J$ satisfies the properties:

$$J = J^* = -J^{tr} = -J^{-1} \tag{3.34}$$

**Definition** A *symplectic matrix* is a matrix $A$ such that

$$A^{tr} J A = J \tag{3.35}$$

We define the symplectic groups:

$$Sp(2n, \kappa) := \{A \in GL(2n, \kappa) | A^{tr} J A = J\} \tag{3.36}$$

The unitary symplectic groups are $USp(2n) := U(2n) \cap Sp(2n, \mathbb{C})$.

**Example 3.9** *Function spaces as groups.*

[16] See Chapter 2, "Linear Algebra Users Manual" for the precise definition of "bilinear form."

Suppose $G$ is a group. Suppose $X$ is any set. Consider the set of all functions from $X$ to $G$:

$$\mathcal{F} = \{f : f \text{ is a function from } X \to G\} \tag{3.37}$$

If we want to stress the role of $X$ and/or $G$ we write $\mathcal{F}[X \to G]$ for $\mathcal{F}$. We claim that $\mathcal{F}$ is also a group. The main step to show this is simply giving a definition of the group multiplication and the inversion operation. The product $\mathbf{m}_{\mathcal{F}}(f_1, f_2)$ of two functions $f_1, f_2 \in \mathcal{F}$ must be another function in $\mathcal{F}$. We define this function by giving a formula for the values of the function $\mathbf{m}_{\mathcal{F}}(f_1, f_2)$ at all values of $x \in X$:

$$\mathbf{m}_{\mathcal{F}}(f_1, f_2)(x) := \mathbf{m}_G(f_1(x), f_2(x)) \tag{3.38}$$

It is the only sensible thing we could write given the data at hand. In less cumbersome notation:

$$(f_1 \cdot f_2)(x) := f_1(x) \cdot f_2(x) \tag{3.39}$$

Similarly inverse of $f$ is the function that maps $x \to f(x)^{-1}$, where $f(x)^{-1} \in G$ is the group element in $G$ inverse to $f(x) \in G$. In formal notation

$$I_{\mathcal{F}}(f)(x) := I_G(f(x)) \qquad \forall x \in X. \tag{3.40}$$

If both $X$ and $G$ have finite cardinality then $\mathcal{F}[X \to G]$ is a finite group. If $X$ or $G$ has an infinite set of points then this is an infinite order group. If $X$ is a positive dimensional manifold and $G$ is a Lie group this is an infinite-*dimensional* space.

In the special case of the space of maps from the circle into the group:

$$LG = \mathcal{F}[S^1 \to G] \tag{3.41}$$

we have the famous "loop group" whose representation theory has many wonderful properties, closely related to the subjects of 2d conformal field theory and string theory. More recently, they have even begun to play important roles in investigations into three- and four-dimensional supersymmetric quantum field theories.

**Example 3.10** *Group Of Gauge Transformations.*

In some cases if $X$ is a manifold and $G$ is a Lie group then, taking a subgroup defined by suitable continuity and differerentiability properties, we get the *group of gauge transformations of Yang-Mills theory.* As a simple example, you are probably familiar with the gauge transformation in Maxwell theory:

$$A_\mu \to A_\mu + \partial_\mu \epsilon \tag{3.42}$$

where $A_\mu$ is the vector potential so that $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the fieldstrength tensor encoding electric and magnetic fields. Indeed, taking $\mu \in \{0, 1, 2, 3\}$ with $\mu = 0$ the time direction and a standard orientation on $\mathbb{R}^3$ we can write (in units with $c = 1$):

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B_2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix} \tag{3.43}$$

Here $\epsilon : \mathbb{M}^{1,3} \to \mathbb{R}$ is a function on 1+3 dimensional Minkowski space. (In a more careful account one would put restrictions on the allowed functions - they should be differentiable and satisfy suitable boundary conditions - etc.) The more canonical object is

$$f : x \mapsto e^{\mathrm{i}\epsilon(x)} \tag{3.44}$$

and this is a function from spacetime, $\mathbb{M}^{1,3}$ to $U(1)$, so $f \in \mathcal{F}[\mathbb{M}^{1,3} \to U(1)]$. This is a better point of view because it generalizes in interesting ways to other spacetimes: if the spacetime $X$ has nontrivial topology (e.g. is not simply connected) then for some $f : X \to U(1)$ it might not be possible to define globally a continuous function $\epsilon : X \to \mathbb{R}$ so that $f(x) = e^{i\epsilon(x)}$. So it is desirable to express the gauge transformation law in a way that only makes reference to $f$, and not to $\epsilon$. One way this can be done is to introduce the first order operators $(-\mathrm{i}\partial_\mu + A_\mu)$ which have the gauge transformation law:

$$(-\mathrm{i}\partial_\mu + A'_\mu) = f^{-1}(-\mathrm{i}\partial_\mu + A_\mu)f \tag{3.45}$$

For many reasons this is a conceptually superior way to write it. [17]

**Example 3.10**: *Permutation Groups.*

Let $X$ be any set. A *permutation* of $X$ is a one-one invertible transformation $\phi : X \to X$. The composition $\phi_1 \circ \phi_2$ of two permutations is a permutation. The identity permutation leaves every element unchanged. The inverse of a permutation is a permutation. Thus, composition defines a group operation on the permutations of any set. This group is designated $S_X$. It is an extremely important group and we will be studying it a lot. In the case where $X = M$ is a manifold we can also ask that our permutations $\phi : M \to M$ be continuous or even differentiable. If $\phi$ and $\phi^{-1}$ are differentiable then $\phi$ is a *diffeomorphism*. The composition of diffeomorphisms is a diffeomorphism by the chain rule, so the set of diffeomorphisms $\mathrm{Diff}(M)$ is a subgroup of the set of all permutations of $M$. The group $\mathrm{Diff}(M)$ is the group of gauge symmetries in General Relativity. Except in the case where $M = S^1$ is the circle, remarkably little is known about the diffeomorphism groups of manifolds. One can ask simple questions about them whose answers are unknown.

**Example 3.11**: *Power Sets As Groups.*

Let $X$ be any set and let $\mathcal{P}(X)$ be the power set of $X$. It is, by definition, the set of all subsets of $X$. If $Y_1, Y_2 \in \mathcal{P}(X)$ are two subsets of $X$ then define the *symmetric difference*:

$$Y_1 + Y_2 := (Y_1 - Y_2) \cup (Y_2 - Y_1) \tag{3.46}$$

This defines an abelian group structure on $\mathcal{P}(X)$. The identity element 0 is the empty set $\emptyset$ and the inverse of $Y$ is $Y$ itself: That is, in this group

$$2Y := Y + Y = \emptyset = 0 \tag{3.47}$$

---

[17] On general spacetimes $X$ one can typically only define $A_\mu$ locally but the 1-form valued first order differential operator $dx^\mu(-\mathrm{i}\partial_\mu + A_\mu)$, suitably interpreted, is globally defined. For more about this look up "connection on a vector bundle."

**Exercise** *Centers Of Direct Products*

Show that the center of a direct product of groups is the direct product of the centers:

$$Z(G_1 \times G_2) = Z(G_1) \times Z(G_2) \tag{3.48}$$

And so on.

---

---

**Exercise** *Classical Matrix Groups: Due Diligence*

a.) Check that each of the above sets (3.16),(3.17),(3.18), (3.36), are indeed subgroups of the general linear group.

b.) Check that, if $b$ is invertible and either symmetric or antisymmetric then (3.30) is a matrix group.

---

---

**Exercise** *Apparent Asymmetry In The Definitions*

In (3.17) we used $AA^{tr} = 1$ but we could have used $A^{tr}A = 1$. Similarly, in (3.18) we used $AA^{\dagger} = 1$ rather than $A^{\dagger}A = 1$. Finally, in (3.36) we could, instead, have defined $Sp(2n, \kappa)$ to be matrices in $M_{2n}(\kappa)$ such that $AJA^{tr} = J$. In all three cases, writing things the other way defines the same group: Why?

(Careful: Just taking the transpose or hermitian conjugate of these equations does not help.) [18]

---

---

**Exercise** $O(2,\mathbb{R})$ *vs.* $SO(2,\mathbb{R})$

a.) Show from the definition above of $O(2,\mathbb{R})$ that the most general element of this group is the form of (3.20) above, OR, of the form

$$\begin{pmatrix} x & y \\ y & -x \end{pmatrix} \qquad x^2 + y^2 = 1 \tag{3.49}$$

btw: Note that

$$\begin{pmatrix} x & y \\ y & -x \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \tag{3.50}$$

b.) Show that no matrix in $O(2,\mathbb{R})$ is simultaneously of the form (3.20) and (3.49). Conclude that, as a manifold, $O(2,\mathbb{R})$ is a disjoint union of two circles.

[18] *Answer*: Hint: Remember that in a group the inverse matrix is in the group. Consider replacing $g \to g^{-1}$ in the definition.

**Exercise** *Symplectic groups and canonical transformations*

Let $q^i, p_i$ $i = 1, \dots n$ be coordinates and momenta for a classical mechanical system.

The **Poisson bracket** of two functions $f(q^1, \dots q^n, p_1, \dots p_n)$, $g(q^1, \dots q^n, p_1, \dots p_n)$ is defined to be

$$\{f, g\} = \sum_{i=1}^{n} \left( \frac{\partial f}{\partial q^i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} \right) \tag{3.51}$$

a.) Show that

$$\{q^i, q^j\} = \{p_i, p_j\} = 0 \qquad \{q^i, p_j\} = \delta^i{}_j \tag{3.52}$$

Suppose we define new coordinates and momenta $Q^i, P_i$ to be linear combinations of the old:

$$\begin{pmatrix} Q^1 \\ \vdots \\ Q^n \\ P_1 \\ \vdots \\ P_n \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1,2n} \\ \vdots & \ddots & \vdots \\ a_{2n,1} & \cdots & a_{2n,2n} \end{pmatrix} \cdot \begin{pmatrix} q^1 \\ \vdots \\ q^n \\ p_1 \\ \vdots \\ p_n \end{pmatrix} \tag{3.53}$$

where $A = (a_{ij})$ is a constant $2n \times 2n$ matrix.

b.) Show that

$$\{Q^i, Q^j\} = \{P_i, P_j\} = 0 \qquad \{Q^i, P_j\} = \delta^i_j \tag{3.54}$$

if and only if $A$ is a symplectic matrix.

c.) Show that $J \in Sp(2n, \mathbb{R})$. Note that it exchanges momenta and coordinates.

d.) What are the conditions on the $n \times n$ matrix $B$ so that

$$\{\begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix}\} \tag{3.55}$$

is a subgroup. [19]

e.) What are the conditions on the $n \times n$ matrix $C$ so that

$$\{\begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix}\} \tag{3.56}$$

is a subgroup. [20]

f.) Show that

$$\begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix} = J \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix} J^{-1} \tag{3.57}$$

[19] *Answer*: $B$ must be a symmetric matrix

[20] *Answer*: $C$ must be a symmetric matrix

for $C = -B$.

g.) Show that a matrix of the form

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{3.58}$$

is symplectic iff

$$\begin{aligned} (A^{tr}C)^{tr} &= A^{tr}C \\ (B^{tr}D)^{tr} &= B^{tr}D \\ A^{tr}D - C^{tr}B &= 1 \end{aligned} \tag{3.59}$$

h.) Show that if $M$ is symplectic then $M^{tr}$ is symplectic.

---

---

**Exercise** *Bogoliubov Transformations*

Suppose there is a collection of operators on a Hilbert space satisfying

$$[a_i, a_j] = [\bar{a}^i, \bar{a}^j] = 0 \qquad [a_i, \bar{a}^j] = \delta_i^{\ j} \tag{3.60}$$

with $1 \leq i, j \leq N$. That is, we have a collection of harmonic oscillators.

Show that a new collection of operators

$$\begin{aligned} b_i &= A_i^{\ j} a_j + B_{ij} \bar{a}^j \\ \bar{b}^i &= C^{ij} a_j + D^i_{\ j} \bar{a}^j \end{aligned} \tag{3.61}$$

satisfies the algebra

$$[b_i, b_j] = [\bar{b}^i, \bar{b}^j] = 0 \qquad [b_i, \bar{b}^j] = \delta_i^{\ j} \tag{3.62}$$

iff

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sp(2N, \mathbb{C}) \tag{3.63}$$

---

---

**Exercise** *The "2 Out Of 3 Property"*

A real $2n \times 2n$ matrix is said to be:

a.) Symplectic if $A^{tr} J A = J$.

b.) Complex if $A^{-1} J A = J$

c.) Orthogonal if $A^{tr} = A^{-1}$.

Show that these three conditions are not equivalent, but any two of them implies the third. This turns out to be an important fact in "Kähler geometry" a special subfield of differential geometry.

**Exercise** *The Quaternion Group And The Pauli Group*

When working with spin-1/2 particles it is very convenient to introduce the standard Pauli matrices:

$$\sigma^1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.64}$$

$$\sigma^2 := \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} \tag{3.65}$$

$$\sigma^3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.66}$$

a.) Show that they satisfy the identity, valid for all $1 \le i, j \le 3$:

$$\boxed{\sigma^i \sigma^j = \delta^{ij} + \mathrm{i}\epsilon^{ijk}\sigma^k} \tag{3.67}$$

b.) Show that the set of matrices

$$Q = \{\pm 1, \pm \mathrm{i}\sigma^1, \pm \mathrm{i}\sigma^2, \pm \mathrm{i}\sigma^3\} \tag{3.68}$$

forms a subgroup of order 8 of $SU(2) \subset GL(2,\mathbb{C})$. It is known as the *quaternion group*.

c.) Show that the set of matrices

$$P = \{\pm 1, \pm \mathrm{i}, \pm\sigma^1, \pm\sigma^2, \pm\sigma^3, \pm \mathrm{i}\sigma^1, \pm \mathrm{i}\sigma^2, \pm \mathrm{i}\sigma^3\} \tag{3.69}$$

forms a subgroup of $U(2) \subset GL(2,\mathbb{C})$ of order 16. It is known as the *Pauli group*.

**Remark**: [21] The Pauli group is often used in quantum information theory. If we think of the quantum Hilbert space of a spin 1/2 particle (isomorphic to $\mathbb{C}^2$ with standard inner product) then there is a natural basis of up and down spins: $v_1 = |\uparrow\rangle$ and $v_2 = |\downarrow\rangle$. Thinking of these as quantum analogs $|0\rangle$ and $|1\rangle$ of classical information bits $0, 1$ we see that $X = \sigma^1$ acts as a "bit flip," while $Z = \sigma^3$ acts as a "phase-flip." $Y = \mathrm{i}\sigma^2$ flips both bits and phases. These are then quantum error operators. Subgroups of the $N^{th}$ direct product of $P$ are used to construct certain quantum error-correcting codes known as stabilizer codes.

d.) Consider the direct product $P^N$ as a set of operators in $\mathrm{End}(\mathcal{H})$, where $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$ is the Hilbert space of $N$ Qbits. Show that every pair of elements in $P^N$ either commutes or anticommutes. This property is quite crucial in the theory of stabilizer codes.

e.) Let $\vec{x}, \vec{y} \in \mathbb{R}^3$. Show that

$$[\vec{x}\cdot\sigma, \vec{y}\cdot\sigma] = 2\mathrm{i}(\vec{x}\times\vec{y})\cdot\sigma \tag{3.70}$$

[21] Many terms used here will be more fully explained in Chapter 2.

$$(\vec{x}\cdot\sigma)\vec{y}\cdot\sigma\vec{x}\cdot\sigma = 2(\vec{x}\cdot\vec{y})\vec{x}\cdot\sigma - \vec{x}^2\vec{y}\cdot\sigma \tag{3.71}$$

---

---

**Exercise** *Function Groups*

Interpret the direct product $G^n$ of a group with itself $n$ times as a group of the form $\mathcal{F}[X \to G]$ for some $X$.

---

---

**Exercise** *Associative Multiplication Laws That Are Not Group Laws*

a.) Consider the candidate group law on the power set $\mathcal{P}(X)$:

$$Y_1 + Y_2 := Y_1 \cap Y_2 \tag{3.72}$$

Why is this not a group law? [22]

b.) Consider the candidate group law on the power set $\mathcal{P}(X)$:

$$Y_1 + Y_2 := Y_1 \cup Y_2 \tag{3.73}$$

Why is this not a group law? [23]

c.) Consider the candidate group law on $\mathbb{Z}$ given by

$$\mathbf{m}(n_1, n_2) = \begin{cases} 0 & n_1 + n_2 = 0 \text{ mod} 2 \\ 1 & n_1 + n_2 = 1 \text{ mod} 2 \end{cases} \tag{3.74}$$

Why is this not a group law? [24]

---

# 4. Homomorphism And Isomorphism

**Definition 4.1**: Let $(G, \mathbf{m}, \mathbf{I}, e)$ and $(G', \mathbf{m}', \mathbf{I}', e')$ be two groups,

1. A *homomorphism* from $(G, \mathbf{m}, \mathbf{I}, e)$ to $(G', \mathbf{m}', \mathbf{I}', e')$ is a mapping that preserves the group law. That is, it is a map of sets $\varphi : G \to G'$ such that, for all $g_1, g_2 \in G$ we have:
$$\varphi(\mathbf{m}(g_1, g_2)) = \mathbf{m}'(\varphi(g_1), \varphi(g_2)) \tag{4.1}$$

2. If $\varphi(g) = 1_{G'}$ implies that $g = 1_G$ then $\varphi$ is said to be *injective* or *into*.

[22] *Answer*: The identity would have to be $e = X$. But then there is no inverse.
[23] *Answer*: The identity would now be $\emptyset$ but there is no inverse.
[24] *Answer*: What is the identity element?

3. If, for any $g' \in G'$ there exists some $g \in G$ such that $\varphi(g) = g'$ then $\varphi$ is said to be *surjective* or *onto*.

4. If $\varphi$ is both into and onto, i.e. if it is both injective and surjective, then $\varphi$ is called an *isomorphism*.

5. One often uses the term *automorphism* of $G$ when $\varphi$ is an isomorphism and $G = G'$, that is $G$ and $G'$ are literally the same set with the same multiplication law.

**Remarks**

1. We will henceforward be more informal and simply say that $\varphi : G \to G'$ is a homomorphism of groups if, for all $g_1, g_2 \in G$:

$$\varphi(\underbrace{g_1 g_2}_{\text{product in G}}) = \overbrace{\varphi(g_1)\varphi(g_2)}^{\text{product in G}'} \tag{4.2}$$

2. A common slogan is: "isomorphic groups are the same."

**Example 0**: *The stupid homomorphism.* Given <u>any</u> two groups $G, G'$ there is always at least one homomorphism $\varphi : G \to G'$, namely, $\varphi(g) = 1_{G'}$ for all $g \in G$. It is an easy exercise to check this is a homomorphism.

**Example 1: $\mu_N$ is isomorphic to $\mathbb{Z}_N$**: Let $N$ be a positive integer. Then we can define a homomorphism

$$\varphi : \mathbb{Z}_N \to \mu_N \tag{4.3}$$

as follows. We want to define $\varphi(\bar{r})$. Recall that $\bar{r} = r + N\mathbb{Z}$ is an equivalence class. We choose <u>any</u> representative $r' \in r + N\mathbb{Z}$. Then we set:

$$\varphi(\bar{r}) := \exp\left(2\pi \mathrm{i} \frac{r'}{N}\right) \tag{4.4}$$

There is a crucial thing to check here: We need to check that the map is actually well-defined. We know that any two representatives $r'_1$ and $r'_2$ for $\bar{r}$ must have the property that $r'_1 - r'_2 = 0 \ \mathrm{mod} N$, that is $r'_1 - r'_2 = \ell N$ for some integer $N$ and now by standard properties of complex numbers we see that indeed $\exp\left(2\pi \mathrm{i} \frac{r'_1}{N}\right) = \exp\left(2\pi \mathrm{i} \frac{r'_2}{N}\right)$.

Next we check that

$$\varphi(\bar{r}_1 + \bar{r}_2) = \varphi(\bar{r}_1)\varphi(\bar{r}_2) \tag{4.5}$$

If you unwind the definitions you should find this follows from a standard property of the exponential map.

Equation (4.5) implies that (4.3) is a homomorphism. In fact one easily checks:

a.) If $\varphi(\bar{r}) = 1$ then $\bar{r} = \bar{0}$. Thus $\varphi$ is injective.

b.) Every element of $\mu_N$ is of the form $\varphi(\bar{r})$ for some $\bar{r}$. Thus, $\varphi$ is surjective. Note that this is equivalent to saying that every element in $\mu_N$ is of the form $\omega^j$ where $\omega = e^{2\pi \mathrm{i}/N}$.

Thus, $\varphi$ is in fact an isomorphism. As we mentioned above, the two groups appeared to be "the same." We have now given precise meaning to that idea.

**Example 2: A family of homomorphisms $\mu_N \to \mu_N$**: For each integer $k$ we can define the $k^{th}$ power map

$$p_k : \mu_N \to \mu_N \tag{4.6}$$

by

$$p_k(z) = z^k \tag{4.7}$$

where $z$ is any $N^{th}$ root of unity. Note that $z^k$ is also an $N^{th}$ root of unity. Moreover $(z_1 z_2)^k = z_1^k z_2^k$ by elementary properties of complex numbers, so $p_k$ is a homomorphism. Note that it is not always injective or surjective. For example, if $k$ is a multiple of $N$ it is the stupid homomorphism. In fact $p_{k+N} = p_k$.

**Example 3: A family of homomorphisms $\mathbb{Z}_N \to \mathbb{Z}_N$**:

For any integer $k$ we can define the "$k^{th}$ multiplication map"

$$m_k : \mathbb{Z}_N \to \mathbb{Z}_N \tag{4.8}$$

by the equation:

$$m_k(\bar{r}) := \overline{kr} \tag{4.9}$$

where on the right hand side $\overline{kr}$ is defined by choosing a representative $r$ for the class $\bar{r}$ and then using ordinary multiplication of integers $k \times r$ (e.g. $2 \times 3 = 6$) and then reducing modulo $N$. Again, one needs to check the equation is well-defined. Note that $m_{k+N} = m_k$.

**Example 4: Relating the homomorphisms in the previous three examples**: Since $\mathbb{Z}_N$ and $\mu_N$ are isomorphic, one should expect that homomorphisms $\mathbb{Z}_N \to \mathbb{Z}_N$ and $\mu_N \to \mu_N$ should be related. Moreover, one should have the intuition that $p_k$ and $m_k$ somehow have the "same effect." Indeed, note that

$$p_k(\omega^j) = (\omega^j)^k = \omega^{jk}. \tag{4.10}$$

is the essential identity. More formally, one easily checks that

$$\varphi \circ m_k = p_k \circ \varphi \tag{4.11}$$

Or, since $\varphi$ is invertible,

$$p_k = \varphi \circ m_k \circ \varphi^{-1} \ . \tag{4.12}$$

In mathematics one often uses *commutative diagrams* to express identities such as (4.11). In this case the diagram looks like

$$\begin{array}{ccc} \mathbb{Z}_N & \xrightarrow{m_k} & \mathbb{Z}_N \\ \downarrow{\scriptstyle\varphi} & & \downarrow{\scriptstyle\varphi} \\ \mu_N & \xrightarrow{p_k} & \mu_N \end{array} \tag{4.13}$$

We say *a diagram commutes* if the following condition holds: The diagram describes a graph with sets associated to vertices and maps associated with oriented edges. Consider following the arrows around any two paths on the graph with the same beginning and final points. We compose the maps associated with those arrows to get two maps from the initial set to the final set. The diagram commutes iff any pair of maps obtained this way are equal.

**Remark** We will discuss in detail later on that when $k$ is an integer relatively prime to $N$ the map $p_k$ is an automorphism of $\mu_N$ and $m_k$ is an automorphism of $\mathbb{Z}_N$. For example in $\mathbb{Z}/3\mathbb{Z} = \{\bar{0}, \bar{1}, \bar{2}\}$ if we take $k = 2$, or any even integer not divisible by 3, then $m_k$ exchanges $\bar{1}$ and $\bar{2}$. (Check that such an exchange is indeed a homomorphism!) We will discuss this kind of example in greater detail in Section §13 below.

**Example 5:** $GL(V)$ **and** $GL(n, \kappa)$: [25] Let $V$ be a finite dimensional vector space over a field $\kappa$. (For example, take $\kappa = \mathbb{R}$ or $\kappa = \mathbb{C}$.) We can define a group $GL(V)$ to be the group of invertible linear transformations from $V$ to itself. The group law is composition. Let $b = \{v_1, \dots, v_n\}$ be an ordered basis for $V$. Given $b$ we define a homomorphism

$$\varphi_b : GL(V) \to GL(n, \kappa) \tag{4.14}$$

as follows. Given $T \in GL(V)$ we have

$$T(v_i) = \sum_j (A_b(T))_{ji} v_j \tag{4.15}$$

The $n \times n$ matrix with matrix $A_b(T)$, with matrix elements $(A_b(T))_{ij}$ defines a matrix associated with $T$ in such a way that

$$A_b(T_1 \circ T_2) = A_b(T_1) A_b(T_2) \tag{4.16}$$

where on the RHS we have ordinary matrix multiplication. The subscript $b$ stresses the dependence of the matrix on the ordered basis $b$. It follows that $A_b(T)$ is invertible if $T$ is and that

$$\varphi_b : GL(V) \to GL(n, \kappa) \tag{4.17}$$

defined by $\varphi_b(T) := A_b(T)$ is a homomorphism of groups. An easy exercise below asks you to show that it is an isomorphism of groups.

[25] See Chapter 2 for definitions of the linear algebra terms

One kind of homomomorphism is especially important:

**Definition 4.2**:

a.) Let $G$ be a group, then a (finite dimensional) representation of $G$ is a finite-dimensional vector space $V$ together with a group homomorphism $\varphi : G \to GL(V)$. Sometimes $V$ is referred to as the *carrier space*.

b.) A *matrix representation* of a group $G$ is a homomorphism

$$\varphi : G \to GL(n, \kappa) \tag{4.18}$$

for some positive integer $n$ and field $\kappa$.

**Remarks**

1. One can also have matrix representations in $GL(n, R)$ where $R$ is a ring.
2. It follows from our definitions that if $\varphi : G \to GL(V)$ is a representation of $G$ and we have an ordered basis $b$ of $V$ then we can produce a corresponding matrix representation $\varphi_b \circ \varphi$ of $G$.
3. In later discussions of representation we will usually denote the homomorphism from $G$ to $GL(V)$ by $T : G \to GL(V)$. We did not use the notation here because it would clash with the previous example. We often refer to "the representation $(T, V)$ of $G$" or sometimes "the representation $(V, T)$ of $G$" etc. We won't be ultra-fastidious about the notation.
4. **Important Remark**: Looking ahead, the basic idea of the expression of symmetry in quantum mechanics is that a group $G$, the "symmetry group," is linearly (and unitarily) represented on the "Hilbert space $\mathcal{H}$ of physical states" of a physical system. This statement, while quite common, is both inaccurate and incomplete. But it is a good start.

---

**Exercise** *Preservation Of Structure*

Show that, for <u>any</u> group homomorphism $\varphi$ we always have:

$$\varphi(1_G) = 1_{G'} \tag{4.19}$$

$$\varphi(g^{-1}) = \mu(g)^{-1} \tag{4.20}$$

---

---

**Exercise** *The Stupid Homomorphism*

Consider the map $\mu : G \to G'$ defined by $\mu(g) = 1_{G'}$. Show that this is a homomorphism.

---

---

**Exercise** *Some Simple Isomorphisms*

a.) Show that the exponential map $x \to e^x$ defines an isomorphism between the additive group $(\mathbb{R}, +)$ and the multiplicative group $(\mathbb{R}_{>0}, \times)$.

b. ) Show that $SO(2)$ and $U(1)$ are isomorphic groups.

c.) Show that $z \to z^{-1}$ is an automorphism of $U(1) \to U(1)$. What is the corresponding automorphism of $SO(2)$?

---

---

**Exercise** *A group "with one free generator"*

Consider a group with a nontrivial element $g_0$ such that every element in the group if a power of $g_0$ or $g_0^{-1}$ and $g_0^n = g_0^m$ iff $n = m$ in the integers.

Show that this group is isomorphic to $\mathbb{Z}$.

Remark: This is an example of what we will call below a group freely generated by one element.

---

---

**Exercise** *Sometimes diagrams don't commute*

Show that the diagram

$$\begin{array}{ccc} \mathbb{Z}_N & \xrightarrow{m_{k_1}} & \mathbb{Z}_N \\ \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \varphi} \\ \mu_N & \xrightarrow{p_{k_2}} & \mu_N \end{array} \tag{4.21}$$

commutes iff $k_1 = k_2 \ \mathrm{mod} N$.

---

---

**Exercise** *The Quaternion Group*

Construct a homomorphism

$$\mu : Q \to \mathbb{Z}_2 \times \mathbb{Z}_2 \tag{4.22}$$

where $Q$ is the Quaternion group (3.68).

---

---

**Exercise**

Let $S_2$ be any set with two elements

a.) Show that there are exactly two possible group structures on $S_2$, and in each case construct an isomorphism of $S_2$ with $\mu_2 \cong \mathbb{Z}_2$. [26]

b.) Consider the matrix group of two elements:

$$\hat{S}_2 = \{\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\} \tag{4.23}$$

with multiplication being matrix multiplication. Construct an isomorphism with $S_2$. [27]

---

---

**Exercise** *Due Diligence*

Show that if $V$ is a finite dimensional vector space of dimension $n$ and $b$ is an ordered basis for $V$ then

$$\varphi_b : GL(V) \to GL(n, \kappa) \tag{4.25}$$

is an isomorphism of groups. [28]

---

**Exercise** *Some Simple Representations Of* $\mu_N$

Let $\omega = e^{2\pi i/N}$.

a.) Show that for any integer $k$ the $k^{th}$ power map

$$p_k(\omega^j) = \omega^{jk} \tag{4.26}$$

defines a representation of $\mu_N$ by $1 \times 1$ matrices.

b.) Show that

$$\mu : \omega^j \mapsto R(\frac{2\pi j}{N}) := \begin{pmatrix} \cos(\frac{2\pi j}{N}) & \sin(\frac{2\pi j}{N}) \\ -\sin(\frac{2\pi j}{N}) & \cos(\frac{2\pi j}{N}) \end{pmatrix} \tag{4.27}$$

defines a two-dimensional matrix representation of $\mathbb{Z}_N$.

[26] *Answer*: We need to choose which element of the set is the identity. Call it $e$. Call the other element $\sigma$. Then we must have $\sigma^2 = e$.

[27] *Answer*: Write $S_2 = \{e, \sigma\}$ with $e$ the identity and $\sigma^2 = e$. Define $\mu : S_2 \to \hat{S}_2$

$$\begin{aligned} \mu(e) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ \mu(\sigma) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned} \tag{4.24}$$

[28] *Answer*: Suppose $\varphi_b(T) = 1_{n\times n}$. Then $T(v_i) = v_i$ and since $b$ is a basis $T$ is the identity. Moreover, given a basis, any invertible matrix defines an invertible linear transformation.

c.) Let $P$ be the $N \times N$ "shift matrix" all of whose matrix elements are zero except for 1's just below the diagonal and $P_{1,N} = 1$. See equation (7.133) below. Show that

$$\mu(\omega^j) = P^j \tag{4.28}$$

is an $N \times N$ dimensional representation of $\mu_N$.

---

---

**Exercise** *Two Characterizations Of Abelian Groups*

Let $G$ be a group.

a.) Consider the map: $\mu : G \to G$ given by squaring: $\mu(g) = g^2$. Show that $\mu$ is a group homomorphism iff $G$ is Abelian.

b.) Consider the map:

$$G \times G \to G \tag{4.29}$$

defined by group multiplication: $\mu(g_1, g_2) = \mathbf{m}(g_1, g_2) = g_1 g_2$. Show that $\mu$ is a group homomorphism iff $G$ is Abelian.

c.) As a generalization of (b) suppose $H \subset G$ is a subgroup. Show that the map $H \times G \to G$ defined by group multiplication is a homomorphism iff $H$ is a subgroup of the center of $G$. [29]

---

---

**Exercise** *Commutative Diagrams*

Suppose $X, Y, Z, W$ are arbitrary sets and we have a map $f : X \to Y$ and $1-1$ maps $g : X \to Z$ and $h : Y \to W$. Show that there is a unique map $\tilde{f} : Z \to W$ such that the diagram:

$$\begin{array}{ccc} X & \xrightarrow{F} & Y \\ \Big\downarrow{\scriptstyle g} & & \Big\downarrow{\scriptstyle h} \\ Z & \xrightarrow{\tilde{f}} & W \end{array} \tag{4.30}$$

commutes.

---

---

**Exercise** *Isomorphisms And Preservation Of Structure*

a.) Suppose $\varphi : G_1 \to G_2$ is an isomorphism. Show that $\varphi^{-1}$ is an isomorphism.

b.) Suppose that $\varphi : G_1 \to G_2$ is an isomorphism, and $\varphi' : G_1' \to G_2'$ is an isomorphism. Suppose also that $\nu_1 : G_1 \to G_1'$ is a homomorphism (not necessarily an

[29] *Answers*: (a): $\mu(g_1 g_2) = \mu(g_1)\mu(g_2)$ implies $g_1^2 g_2^2 = (g_1 g_2)^2 = g_1 g_2 g_1 g_2$. Now cancel the $g_1$ on the left and the $g_2$ on the right. (b): ETC.

isomorphism). Show that there is a unique homomorphism $\nu_2 : G_2 \to G_2'$ so that we have the commmutative diagram:

$$\begin{array}{ccc} G_1 & \xrightarrow{\nu_1} & G_1' \\ \downarrow{\scriptstyle \varphi} & & \downarrow{\scriptstyle \varphi'} \\ G_2 & \xrightarrow{\nu_2} & G_2' \end{array} \tag{4.31}$$

---

---

**Exercise** *Inverse Image Of Subgroups Under Homomorphism Are Subgroups*
Suppose $\varphi : G_1 \to G_2$ is a homomorphism and $H_2 \subset G_2$ is a subgroup.
Show that $\varphi^{-1}(H_2) \subset G_1$ is a subgroup. [30]

---

---

**Exercise** *Fiber Products*

Given groups $G_1$ and $G_2$, and homomorphisms $\psi_1 : G_1 \to H$ and $\psi_2 : G_2 \to H$ one can define a subset of $G_1 \times G_2$ known as a *fiber product*:

$$G_1 \times_{\psi_1,\psi_2} G_2 := \{(g_1, g_2)|\psi_1(g_1) = \psi_2(g_2)\} \ . \tag{4.32}$$

Show that the fiber product is in fact a subgroup of $G_1 \times G_2$, where $G_1 \times G_2$ has the direct product group structure.

---

---

**Exercise** *Structure Of The Power Set Group*

In equation (3.46) above we noted that for any set $X$, the power set $\mathcal{P}(X)$ has the structure of an Abelian group. Show that this group is isomorphic to a direct product of $\mathbb{Z}_2$ factors, each factor being generated by the singleton elements $\{x\} \subset X$ for each $x \in X$.

---

### 4.1 Kernel And Image

Given an arbitrary homomorphism

$$\varphi : G \to G' \tag{4.33}$$

there is automatically a "God-given" subgroup of both $G$ and $G'$:

**Definition 7.4.1**:

[30] *Answer*: Suppose $h, h' \in \varphi^{-1}(H_2)$ then $\varphi(h), \varphi(h') \in H_2$, but since $H_2 \subset G_2$ is a subgroup we know that $\varphi(h)\varphi(h') \in H_2$ but this means $\varphi(hh') \in H_2$. Therefore $hh' \in \varphi^{-1}(H_2)$.

a.) The *kernel* of $\varphi$ is

$$K = \ker\varphi := \{g \in G | \varphi(g) = 1_{G'}\} \tag{4.34}$$

b.) The *image* of $\varphi$ is

$$\mathrm{im}\varphi := \varphi(G) \subset G' \tag{4.35}$$

---

**Exercise** *Due Diligence*
b.) Check that $\varphi(G) \subseteq G'$ is indeed a subgroup.
a.) Check that $\ker(\varphi) \subset G$ is indeed a subgroup. [31]

---

**Example 1**: Consider $m_2 : \mathbb{Z}_4 \to \mathbb{Z}_4$. Then

$$\ker(m_2) = \mathrm{im}(m_2) = \{\bar{0}, \bar{2}\} \cong \mathbb{Z}_2 \tag{4.36}$$

**Example 2**: In the previous example $\ker(m_2) = \mathrm{im}(m_2)$ but this is a very exceptional case. Consider $m_{12} : \mathbb{Z}_{15} \to \mathbb{Z}_{15}$. The reader should check that

$$\begin{aligned} \ker(m_{12}) &= \{\bar{0}, \bar{5}, \bar{10}\} \cong \mathbb{Z}_3 \\ \mathrm{im}(m_{12}) &= \{\bar{0}, \bar{3}, \bar{6}, \bar{9}, \bar{12}\} \cong \mathbb{Z}_5 \end{aligned} \tag{4.37}$$

**Example 3**: More generally, consider $m_k : \mathbb{Z}_N \to \mathbb{Z}_N$ for positive integers $N, k$. We can write them as $N = gs$ and $k = gt$ where $s, t, g$ are relatively prime and $g$ is the greatest common divisor of $N$ and $k$, denoted $g = (N, k)$. Then (see the section on elementary number theory below) it is not hard to see that

$$\ker(m_k) = \{\bar{0}, \bar{s}, 2\bar{s}, \dots, (g-1)\bar{s}\} \cong \mathbb{Z}_g \tag{4.38}$$

while (using Bezout's theorem):

$$\mathrm{im}(m_k) = \{\bar{0}, \bar{g}, 2\bar{g}, \dots, (s-1)\overline{g}\} \cong \mathbb{Z}_s \ . \tag{4.39}$$

**Example 4**: Consider the homomorphism

$$\varphi : U(1) \to SU(2) \tag{4.40}$$

defined by

$$\varphi(z) = \begin{pmatrix} z^N & 0 \\ 0 & z^{-N} \end{pmatrix} \tag{4.41}$$

[31] *Answer*: If $k_1, k_2 \in K$ then $\varphi(k_1 k_2) = \varphi(k_1)\varphi(k_2) = 1_{G'}$. So $K$ is closed under multiplication. The group properties of $K$ now follow.

Then

$$\ker(\varphi) \cong \mu_N \tag{4.42}$$

while the image is the subgroup of diagonal matrices in $SU(2)$.

---

**Exercise** *Kernel Of Projection Map In A Fiber Product*

Let $\pi_1 : G_1 \times G_2 \to G_1$ be the projection map $\pi_1 : (g_1, g_2) \mapsto g_1$.

a.) Show that $\pi_1$ restricted to any subgroup of $G_1 \times G_2$ is a homomorphism.

b.) Show that the kernel of $\pi_1$ restricted to the fiber product of (4.32) is isomorphic to $\ker(\psi_2)$.

---

# 5. Group Actions On Sets

## 5.1 Group Actions On Sets

Recall that we said that if $X$ is any set then a *permutation* of $X$ is a 1-1 and onto mapping $X \to X$. The set $S_X$ of all permutations forms a group under composition.

We now define the notion of an action of a group on a set. This is a very important notion, and we will return to it extensively when discussing examples. If the following discussion seems too abstract the reader should consult section 8 for a number of concrete examples beyond the ones we are about to give. There is further material related to group actions in Chapter 3.

There are three ways to think about a group action on a set:

**First Way**: A *transformation group* on $X$ is a subgroup of $S_X$.

**Second Way**: We define a *left $G$-action on a set $X$* to be a map $\phi : G \times X \to X$ compatible with the group multiplication law as follows:

$$\phi(g_1, \phi(g_2, x)) = \phi(g_1 g_2, x) \tag{5.1}$$

We would also like $x \mapsto \phi(1_G, x)$ to be the identity map. Now, equation (5.1) implies that

$$\phi(1_G, \phi(1_G, x)) = \phi(1_G, x) \tag{5.2}$$

which is compatible with, but does not quite imply that $\phi(1_G, x) = x$. Thus in defining a group action we must also impose the condition:

$$\phi(1_G, x) = x \qquad \forall x \in X. \tag{5.3}$$

---

**Exercise**

Give an example of a map $\phi : G \times X \to X$ that satisfies (5.1) but not (5.3). [32]

---

The above two ideas are connected to each other by the

**Third Way**: Yet another way to say this is the following. First, for a fixed $g \in G$ define a map $\phi(g, \cdot) : X \to X$ that takes $x \mapsto \phi(g, x)$. Then using both axioms for a group action one proves that $\phi(g, \cdot) \in S_X$. Call this map $\Phi(g) \in S_X$. Now define a map $\Phi : G \to S_X$ that takes $g \mapsto \Phi(g)$. Again using the axioms for a group action one easily checks that $\Phi(g_1) \circ \Phi(g_2) = \Phi(g_1 g_2)$ Therefore, to say we have a group action of $G$ on $X$ is to say that $\Phi$ is a homomorphism of $G$ into the permutation group $S_X$.

**Definition**: If $X$ has a group action by a group $G$ we say that $X$ is a $G$-set.

**Notation**: Again, our notation is overly cumbersome because we want to stress the concept. Usually one writes a left $G$-action as

$$g \cdot x := \phi(g, x) \tag{5.4}$$

The key axioms become

$$\begin{aligned} g_1 \cdot (g_2 \cdot x) &= (g_1 g_2) \cdot x \\ 1_G \cdot x &= x \end{aligned} \tag{5.5}$$

We can think of $\Phi(g)$ as the map that sends $x \to g \cdot x$.

**Definition/Discussion: Orbits**: If $G$ acts on a set $X$ then we can define an equivalence relation on $X$ by saying that two elements $x_1, x_2 \in X$ are equivalent, $x_1 \sim x_2$ if there is some $g \in G$ with $\phi(g, x_1) = x_2$. The reader should check that this is indeed an equivalence relation. The equivalence class $[x]$ with this equivalence relation is known as the *orbit of G through a point x*. So, concretely it is the set of points $y \in X$ which can be reached by the action of $G$:

$$O_G(x) = \{y : \exists g \quad \text{such that} \quad y = g \cdot x\} \tag{5.6}$$

The notion of orbits is very important in geometry, gauge theory and many other subjects. The *set of orbits* is denoted $X/G$. We will discuss many examples below of this extremely important concept.

**Remark**: If $X, G$ are topological spaces, with the $G$ action on $X$ continuous then $X/G$ carries a natural topology. It has the largest number of open sets so that the map $p : X \to X/G$ taking $x \in X$ to the equivalent class $[x] \in X/G$ is continuous. So a subset $U \subset X/G$ is open iff $p^{-1}(U) \subset X$ is open.

**Examples**

[32] *Answer*: As the simplest example, choose any element $x_0 \in X$ and define $\phi(g, x) = x_0$ for all $g, x$. For a slightly less trivial example consider $G = S_2$ and let $\phi(e, x) = \phi(\sigma, x) = f(x)$. Then if $f \circ f(x) = f(x)$ the condition (5.1) will be satisfied, but there certainly exist functions with $f \circ f = f$ which are not the identity map.

1. Consider rotations around the origin of $\mathbb{R}^2$. They act on the points of $\mathbb{R}^2$ as a $G$-action. The distinct orbits are circles centered on the origin. The origin is also an orbit by itself.

2. Consider rotations around the origin by multiples of $2\pi/3$. Check that this group is isomorphic to $\mathbb{Z}_3$. Consider an equilateral triangle centered on the origin. Then the rotations act on the triangle preserving it. Thus, $\mathbb{Z}_3$ acts as a group of symmetries of the equilateral triangle. Intuitively, group theory is the theory of symmetry. We have just illustrated how that idea can be formalized through the notion of group action on a set. ♣FIGURE HERE! ♣

3. A group representation of $G$ on a vector space $V$ is the same thing as a $G$-action on $V$ where $\Phi(g)$ are linear transformations.

4. *Dynamical Systems As Group Actions: Action Of $\mathbb{Z}$ on any set $X$.* Recall that we noted below that if there is an element $g_0 \in G$ so that every element in $G$ is uniquely of the form $g_0^n$ for $n \in \mathbb{Z}$ (where negative integers are powers of $g_0^{-1}$ and the group law is just $g_0^n g_0^m = g_0^{n+m}$ then such a group is isomorphic to $\mathbb{Z}$. [33] The action of such a group on any set $X$ is of the following form: There is an invertible map $f : X \to X$ and can define a group action on $X$ by

$$g_0^n \cdot x = \begin{cases} \underbrace{f \circ \cdots \circ f}_{n \quad \text{times}}(x) & n > 0 \\ x & n = 0 \\ \underbrace{f^{-1} \circ \cdots \circ f^{-1}}_{|n| \quad \text{times}}(x) & n < 0 \end{cases} \tag{5.7}$$

Any $\mathbb{Z}$-action must be of this form since we can define $f(x) := g_0 \cdot x$. Thus the orbit of a point $x \in X$ is the set of all images of successive actions of $f$ and $f^{-1}$. This is known as a discrete dynamical system. The map $f$ defines evolution by discrete time steps.

5. *A simple example of the induced topology on the set of orbits $X/G$.* Consider the action of $\mathbb{Z}$ on $\mathbb{R}$ where $n : x \mapsto x + n$. The orbit of a real number $r$ is $r + \mathbb{Z}$. Note that the value of the function $p(x) := e^{2\pi i x}$ uniquely determines an orbit. So we can identify the space of orbits $X/G = \mathbb{R}/\mathbb{Z}$, with this action of $\mathbb{Z}$, with the points on the circle. Moreover, with the quotient topology defined above it is the circle with its usual topology.

6. Let $G = GL(n, \kappa)$ and $X = \kappa^n$, the $n$-dimensional vector space over $\kappa$. Then the usual linear action on vectors defines a group action of $G$ on $X$. One can check that there are only two orbits: The zero vector gives one orbit. This will be an important observation in our discussion of projective space in section 5.2.

[33] As we will soon see, $G$ is the free group on one generator.

7. If $G = \mathbb{Z}_2$ acts linearly on $\mathbb{R}^{n+1}$ (i.e. $V = \mathbb{R}^{n+1}$ is a representation of $\mathbb{Z}_2$) then we can choose coordinates so that the nontrivial element $\sigma \in G$ acts by

$$\sigma \cdot (x^1, \dots, x^{n+1}) = (x^1, \dots, x^p, -x^{p+1}, \cdots, -x^{p+q}) \tag{5.8}$$

where $p + q = n + 1$. Note that this action preserves the equation of the sphere $\sum_i (x^i)^2 - 1 = 0$ and hence descends to a $\mathbb{Z}_2$-action on the sphere $S^n$. The case $p = 0, q = n + 1$ is the antipodal map. This also descends to a group action on $\mathbb{RP}^N$ for $N = p + q - 1$.

8. Now consider a set of integers $(q_1, \dots, q_n) \in \mathbb{Z}^n$. Then for each such set of integers there is a $\mathbb{C}^*$-action on $\mathbb{CP}^{n-1}$ defined by

$$\mu \cdot [X^1 : \cdots : X^n] := [\mu^{q_1} X^1 : \cdots : \mu^{q_n} X^n] \tag{5.9}$$

for $\mu \in \mathbb{C}^*$. (Check it is well-defined!)

9. The group $G = SL(2, \mathbb{R})$ acts on the complex upper half plane:

$$\mathcal{H} = \{\tau | \mathrm{Im}\tau > 0\} \tag{5.10}$$

via

$$g \cdot \tau := \frac{a\tau + b}{c\tau + d} \tag{5.11}$$

where

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{5.12}$$

The reader should check that indeed:

$$g_1 \cdot (g_2 \cdot \tau) = (g_1 g_2) \cdot \tau \tag{5.13}$$

and moreover, for $g \in SL(2, \mathbb{R})$ we have

$$\mathrm{Im}(g \cdot \tau) = \frac{\mathrm{Im}(\tau)}{|c\tau + d|^2} \tag{5.14}$$

so that the $G$ action preserves the upper half plane. Note this would <u>not</u> be true if $g \in SL(2, \mathbb{C})$.

10. *The Space Of Gauge Inequivalent Fields In Maxwell Theory.* Let $\mathbb{M}^{1,d}$ be $(d+1)$-dimensional Minkowski space and consider the first order differential operators $(-i\partial_\mu + A_\mu)$, which parametrize the data of the electromagnetic gauge potential. Then the group $\mathcal{G} = Map[\mathbb{M}^{1,d} \to U(1)]$ is a group which acts on the set $\mathcal{A}$ of all gauge potentials by $A_\mu \to A_\mu - \mathrm{i} f^{-1} \partial_\mu f$. The space of orbits $\mathcal{A}/\mathcal{G}$ parametrizes gauge-inequivalent field configurations.

---

**Exercise** *G Actions And Equivalence Relations*

a.) Suppose $G$ acts on a set $X$. Then $G$ acts on the Cartesian product $X \times X$ via $g \cdot (x_1, x_2) := (g \cdot x_1, g \cdot x_2)$. Suppose $R \subset X \times X$ is an equivalence relation which is preserved by the $G$-action. That is, $(x_1, x_2) \in R$ implies $(g \cdot x_1, g \cdot x_2) \in R$. Show that $G$ acts on the set of equivalence classes. [34]

b.) Suppose that $G_1$ and $G_2$ act on $X$ so that their group actions commute. Show that $G_1$ acts on the set of orbits $X/G_2$ and vice versa.

---

---

**Exercise** *G-actions And Homomorphisms*

Suppose $G_2$ acts on a set $X$ via $\phi_2 : G_2 \times X \to X$.

Suppose $\varphi : G_1 \to G_2$ is a homomorphism.

Show that $\phi_1 : G_1 \times X \to X$ defined by

$$\phi_1(g_1, x) := \phi_2(\varphi(g_1), x) \tag{5.15}$$

is a $G_1$-action on $X$.

---

### 5.2 Projective Spaces

Let $\kappa = \mathbb{R}$ or $\kappa = \mathbb{C}$. The multiplicative group $\kappa^*$ acts on $X = \kappa^n$ by scaling all the coordinates. Note that scaling a nonzero vector by a nonzero scalar gives a nonzero vector so $G = \kappa^*$ also restricts to act on $\tilde{X} = \kappa^n - \{0\}$. The set of orbits in the two cases differs only by the addition of one point: The orbit of the zero vector. However, the induced topology on $X/G$ is quite different in the two cases. With the quotient topology $\kappa^n/\kappa^*$ is not a Hausdorff. To see this, consider what would constitute an open set containing the orbit of the zero vector. It would have to be a set $\mathcal{U} \subset \kappa^n/\kappa^*$ so that $p^{-1}(\mathcal{U})$ is an open set of the origin in $\kappa^n$. But, being the inverse image of $p$, that set would have to be invariant under scaling by all elements in $\kappa^*$. But this means that $p^{-1}(\mathcal{U}) = \kappa^n$, and hence $\mathcal{U} = X/G$. In other words, the only open set in $\kappa^n/\kappa^*$ containing $[\vec{0}]$ is all of $X/G$. But this means we cannot find open sets separating $[\vec{0}]$ from $[v]$ for any nonzero vector $v$ and hence $\kappa^n/\kappa^*$ is not Hausdorff.

The situation is very different if we discard the problematic zero vector. Then it is not hard to show that $(\kappa^n - \{0\})/\kappa^*$ is a nice Hausdorff space, and in fact a smooth manifold. This important manifold is often denoted $\mathbb{RP}^{n-1}$ for $\kappa = \mathbb{R}$ and $\mathbb{CP}^{n-1}$ for $\kappa = \mathbb{C}$. When $\kappa = \mathbb{C}$ it is even a complex manifold, meaning that the transition functions on patches are holomorphic functions. See Chapter 3 for more discussion. We will often denote elements of $\kappa\mathbb{P}^{n-1}$ by $[X^1 : \cdots : X^n]$ where the $X^i$ cannot all be zero and they are complex, or real,

[34] *Answer.* Define $g \cdot [x] := [g \cdot x]$ and check this definition is well-defined.

depending on whether $\kappa = \mathbb{C}$ or $\kappa = \mathbb{R}$. This stands for the equivalence class of a vector $(X^1, \dots, X^n) \in \kappa^n - \{0\}$.

**Examples**

1. $\mathbb{RP}^1$ is the moduli space of lines through the origin in $\mathbb{R}^2$. EXPLAIN IN DETAIL.

2. More generally, $\mathbb{RP}^n$ or $\mathbb{CP}^n$ is the moduli space of one-dimensional subspaces of $\mathbb{R}^{n+1}$ or $\mathbb{C}^{n+1}$, respectively.

3. It is natural to identify $\mathbb{CP}^1$ with the extended complex plane $\widehat{\mathbb{C}}$ and hence also with the unit sphere $S^2 \subset \mathbb{R}^3$. We can map

$$\pi : [z_1 : z_2] \mapsto \begin{cases} z_1/z_2 & z_2 \neq 0 \\ \infty & z_2 = 0 \end{cases} \tag{5.16}$$

and this gives a $1-1$ correspondence between points in $\mathbb{CP}^1$ and points on the extended complex plane. Another way to say the same thing is as follows: We cover $\mathbb{CP}^1$ with two open sets

$$\mathbb{CP}^1 = \mathcal{U}_N \cup \mathcal{U}_S \tag{5.17}$$

where $\mathcal{U}_N$ are the points $[z_1 : z_2]$ with $z_2 \neq 0$ and $\mathcal{U}_S$ are the points with $z_1 \neq 0$. Now consider the map:

$$\phi_N : \mathcal{U}_N \to \mathbb{C} \qquad \phi_N([z_1 : z_2]) := z_N := z_1/z_2 \tag{5.18}$$

This defines a coordinate system in $\mathcal{U}_N$. Note that if we tried to extend the domain of $\phi_N$ to include the point $[1 : 0]$ the image would be the "point at infinity." Similarly, on the set $\mathcal{U}_S$ where $z_1 \neq 0$ we can define a map

$$\phi_S : \mathcal{U}_S \to \mathbb{C} \qquad \phi_S([z_1 : z_2]) := z_S := z_2/z_1 \tag{5.19}$$

Note that the point $[0 : 1]$ corresponds to a "point at infinity" in this mapping. Also note that $\mathcal{U}_N \cap \mathcal{U}_S$ is all of $\mathbb{CP}^1$ except for two points, $[1 : 0]$ and $[0 : 1]$. On this intersection it makes sense to compare coordinates and we always have:

$$z_S z_N = 1 \tag{5.20}$$

We have thus described $\mathbb{CP}^1$ as a manifold.

Now, it is also true that $\widehat{\mathbb{C}} \cong S^2$ via stereographic projection. The stereographic projection from the <u>north pole</u> is the mapping:

$$z_N = \frac{\hat{x}^1 + \mathrm{i}\hat{x}^2}{1 - \hat{x}^3} \tag{5.21}$$

The stereographic projection from the <u>south pole</u> is the mapping:

$$z_S = \frac{\hat{x}^1 - \mathrm{i}\hat{x}^2}{1 + \hat{x}^3} \tag{5.22}$$

Once again

$$z_S z_N = 1 \tag{5.23}$$

so we are describing the same manifold. So one can identify the three manifolds $\mathbb{CP}^1$, $\widehat{\mathbb{C}}$, and $S^2$.

**Remarks**

1. *Quantum Mechanics Is Projective.* Projective spaces play an important role in quantum mechanics. The of pure quantum states in a quantum system with an $n$-dimensional Hilbert space is a copy of $\mathbb{CP}^{n-1}$. The pure state is a one-dimensional projection operator onto a line, and $\mathbb{CP}^{n-1}$ is the moduli space of lines. Often people speak of "the state $\psi \in \mathcal{H}$" where $\mathcal{H}$ is a Hilbert space. But the only meaningful thing is the equivalence class of $\psi$ under multiplication by $\mathbb{C}^*$. If we normalize $\psi$ then this ambiguity becomes multiplication by a phase and people speak - misleadingly - or a "ray in Hilbert space." The only physical information is in the density matrix, which in this case is a rank one projection operator onto the line through the origin and $\psi$:
$$P_\psi = \frac{\psi\psi^\dagger}{(\psi,\psi)} = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle} \tag{5.24}$$
where we have given both math and physics notation, respectively. Note that $\psi' = \alpha\psi$ where $\alpha \in \mathbb{C}^*$ if and only if $P_\psi = P_{\psi'}$. The rank one projection operators are in 1-1 correspondence with the one-dimensional subspaces of Hilbert space and these are in 1-1 correspondence with the <u>points</u> of $\mathbb{CP}^n$ if the Hilbert space is $(n+1)$ dimensional. These remarks generalize to infinite-dimensional Hilbert space.

♣Much more explanation needed here. ♣

2. *Coordinate Systems For Projective Spaces.* Consider $\mathbb{RP}^n$. For each $1 \le i \le n+1$ define a subset $\mathcal{U}_i \subset \mathbb{RP}^n$ to be the set of points $[x^1 : \cdots : x^{n+1}]$ with $x^i \neq 0$. If $x^i \neq 0$ then there is a <u>unique</u> representative vector in $\mathbb{R}^{n+1}$ of the form $(y^1, \dots, y^{n+1})$ with $y^i = 1$. Then the remaining coordinates $y^j$ define a point in $\mathbb{R}^n$ and are coordinates defined on the patch $\mathcal{U}_i$. For example, on $\mathcal{U}_1$, each point $[x^1 : \cdots : x^{n+1}]$ has a unique representative of the form $(1, y^2, \dots, y^{n+1})$. Then $(y^2, \dots, y^{n+1})$ is a set of coordinates on this patch. Now consider points on the overlap $\mathcal{U}_i \cap \mathcal{U}_j$ for $i \neq j$. Then a point $[x^1 : \cdots : x^{n+1}]$ in the intersection has both $x^i$ and $x^j$ nonzero. The same point has a unique representative $(y^1, \dots, y^{n+1})$ such that $y^i = 1$ and a second unique representative $(w^1, \dots, w^{n+1})$ such that $w^j = 1$. Therefore we have
$$(w^1, \dots, w^{n+1}) = \frac{1}{y^j}(y^1, \dots, y^{n+1}) \tag{5.25}$$
and also
$$(y^1, \dots, y^{n+1}) = \frac{1}{w^i}(w^1, \dots, w^{n+1}) \tag{5.26}$$
so that $w^k = y^k/y^j$ for all $1 \le k \le n+1$ and similarly $y^k = w^k/w^i$. The reader should check that these two equations are compatible. This gives the change of coordinates on patch overlaps $\mathcal{U}_i \cap \mathcal{U}_j$.

### 5.3 Equivariant Maps

**Definition** Let $X, X'$ be two $G$-spaces. We say that $f : X \to X'$ is an *equivariant map*, (a.k.a. a *morphism of $G$-spaces*) if for all $x \in X$, and $g \in G$ we have

$$f(g \cdot x) = g \cdot f(x) \tag{5.27}$$

or more formally:

$$f(\phi(g, x)) = \phi'(g, f(x)) \tag{5.28}$$

or, equivalently, the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & X' \\ {\scriptstyle \Phi(g)}\downarrow & & \downarrow{\scriptstyle \Phi'(g)} \\ X & \xrightarrow{f} & X' \end{array} \tag{5.29}$$

commutes. [35]

**Examples**

1. Consider the $\mathbb{Z}_2$ action on $\mathbb{R}^{n+1}$ in equation (5.8) above. Any linear equivariant map from $\mathbb{R}^{n+1}$ to itself for this action is of the form

$$\begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \tag{5.30}$$

   where $A \in M_p(\mathbb{R})$ and $D \in M_q(\mathbb{R})$.

2. Let $\mathcal{H}$ be the upper half complex plane and $\overline{\mathcal{H}}$ the lower complex half-plane. Then $z \to \bar{z}$ is an equivariant map for the $SL(2, \mathbb{R})$ action discussed above.

3. The projection map from $\mathbb{R}^{N+1} - \{0\} \to \mathbb{RP}^N$ is equivariant for the group action defined in (5.8).

4. The group $G = SL(2, \mathbb{C})$ acts on $\mathbb{CP}^1$ via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot [z_1 : z_2] := [az_1 + bz_2 : cz_2 + dz_2] \tag{5.31}$$

   and the group $G = SL(2, \mathbb{C})$ also acts on $\widehat{\mathbb{C}}$ via

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z := \frac{az + b}{cz + d} \tag{5.32}$$

   Note that, unless $c = 0$, one must use $\widehat{\mathbb{C}}$ and not $\mathbb{C}$ because the Mobius transformation takes $z = -d/c$ to $\infty$. One can check that the map $\pi : \mathbb{CP}^1 \to \widehat{\mathbb{C}}$ is equivariant for these actions of $SL(2, \mathbb{C})$.

---

[35] We could actually generalize this to a pair of maps $F : X \to X'$ and a homomorphism $\varphi : G \to G'$ such that $F(g \cdot x) = \varphi(g) \cdot F(x)$. This generalization is quite natural in the context of category theory. See section 17 below.

---

**Exercise** *Equivariant Maps For The Action Of* $\mathbb{Z}$ *On* $\mathbb{R}$ *By Translation*

Consider the action of $\mathbb{Z}$ on $\mathbb{R}$ by translation: $\phi_n : x \mapsto x + n$, for $x \in \mathbb{R}$ and $n \in \mathbb{Z}$. Show that the only equivariant analytic map $f : \mathbb{R} \to \mathbb{R}$ for this action are the translations $f(x) = x + \alpha$ for some $\alpha \in \mathbb{R}$.

---

---

**Exercise** *Equivariant Maps For The Fundamental Representation Of* $SU(2)$

Consider the linear action of $SU(2)$ on the Qbit space $\mathbb{C}^2$. Show that any linear equivariant map $T : \mathbb{C}^2 \to \mathbb{C}^2$ must be of the form $T(\vec{z}) = \alpha\vec{z}$ for some complex number $\alpha$.

---

## 5.4 Some Special Relations Of $2 \times 2$ Matrix Groups And Symmetries Of Spacetime

In low dimensions there are some very beautiful special relations between natural actions of $2 \times 2$ matrix groups on $2 \times 2$ matrices and rotation/Lorentz symmetries of natural spacetimes.

### 5.4.1 An Important Homomorphism From $SU(2)$ To $SO(3)$

Let $\mathcal{H}_2^0$ be the vector space of $2 \times 2$ complex matrices which are both Hermitian and traceless. It is not hard to see that this is a real vector space and it is isomorphic to $\mathbb{R}^3$ with isomorphism

$$h : \mathbb{R}^3 \to \mathcal{H}_2^0 \tag{5.33}$$

$$h(\vec{x}) := \vec{x} \cdot \vec{\sigma} = \begin{pmatrix} x^3 & x^1 - \mathrm{i}x^2 \\ x^1 + \mathrm{i}x^2 & -x^3 \end{pmatrix} \tag{5.34}$$

We now define a group action of $SU(2)$ on $\mathcal{H}_2^0$. For $u \in SU(2)$ and $m \in \mathcal{H}_2^0$ we define:

$$\phi(u, m) := umu^{-1} \tag{5.35}$$

For a fixed $u$ let $C_u : \mathcal{H}_2^0 \to \mathcal{H}_2^0$ be defined by:

$$C_u(m) := umu^{-1} \ . \tag{5.36}$$

Note that $C_u(m)$ is traceless if $m$ is. Moreover since $u^{-1} = u^\dagger$, $C_u(m)$ is Hermitian if $m$ is. Therefore $C_u$ maps $\mathcal{H}_2^0$ to $\mathcal{H}_2^0$. Moreover, note that $C_u$ is a linear transformation on $\mathcal{H}_2^0$.

Since $h$ is an isomorphism of vector spaces we can defining a corresponding linear transformation $R(u) : \mathbb{R}^3 \to \mathbb{R}^3$ by

$$\begin{array}{ccc} \mathbb{R}^3 & \xrightarrow{R(u)} & \mathbb{R}^3 \\ {\scriptstyle h}\downarrow & & \downarrow{\scriptstyle h} \\ \mathcal{H}_2 & \xrightarrow[C_u]{} & \mathcal{H}_2 \end{array} \tag{5.37}$$

Since $h$ and $C_u$ are invertible maps $R(u)$ is an invertible linear map, i.e. $R(u) \in GL(3, \mathbb{R})$.

The association $u \mapsto R(u)$ is a map

$$R : SU(2) \to GL(3, \mathbb{R}) \quad . \tag{5.38}$$

It has the property that, for all $\vec{x} \in \mathbb{R}^3$ we have:

$$u\vec{x} \cdot \vec{\sigma} u^{-1} = (R(u)\vec{x}) \cdot \vec{\sigma} \tag{5.39}$$

Put differently, since $\vec{x}$ is real, the $2 \times 2$ matrix $u\vec{x} \cdot \vec{\sigma} u^{-1}$ is hermitian, and traceless, and hence has to be of the form $\vec{y} \cdot \vec{\sigma}$, where $\vec{y} \in \mathbb{R}^3$. We call that vector $\vec{y} = R(u)\vec{x}$. It should be clear from the definition that $R(u_1 u_2) = R(u_1)R(u_2)$, so that $R : SU(2) \to GL(3, \mathbb{R})$ is a homomorphism of groups.

In fact, we claim that $R$ is a homomorphism from $SU(2)$ into the subgroup $SO(3) \subset GL(3, \mathbb{R})$. To see this we note that the Euclidean norm on $\mathbb{R}^3$ is nicely expressed in terms of $\mathcal{H}_2^0$. Note that (see exercise below)

$$(\vec{x} \cdot \vec{\sigma})^2 = \vec{x}^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{5.40}$$

Alternatively, (see exercise below)

$$\det \vec{x} \cdot \vec{\sigma} = -\vec{x}^2 \tag{5.41}$$

Therefore, if we write $u\vec{x} \cdot \vec{\sigma} u^{-1} = \vec{y} \cdot \vec{\sigma}$ for some $\vec{y}$ then it follows from either (5.40) or (5.41) that $\vec{x}^2 = \vec{y}^2$. We therefore conclude that $\vec{y} = R(u)\vec{x}$ with $R(u) \in O(3) \subset GL(3, \mathbb{R})$.

In fact, $R$ maps $SU(2)$ into the subgroup $SO(3) \subset O(3)$. To see this note that from the definition:

$$u\sigma^i u^{-1} = R(u)_{ji} \sigma^j \quad . \tag{5.42}$$

(Repeated indices are summed on the RHS.) Now note that

$$\begin{aligned} 2\mathrm{i} &= \mathrm{tr}\ \left(\sigma^1 \sigma^2 \sigma^3\right) \\ &= \mathrm{tr}\ \left(u\sigma^1 u^{-1} u\sigma^2 u^{-1} u\sigma^3 u^{-1}\right) \\ &= R(u)_{j_1,1} R(u)_{j_2,2} R(u)_{j_3,3} \mathrm{tr}\ \left(\sigma^{j_1} \sigma^{j_2} \sigma^{j_3}\right) \\ &= 2\mathrm{i} \epsilon^{j_1 j_2 j_3} R(u)_{j_1,1} R(u)_{j_2,2} R(u)_{j_3,3} \\ &= 2\mathrm{i} \det R(u) \end{aligned} \tag{5.43}$$

and hence $\det R(u) = 1$. Alternatively, if you know about Lie groups, you can use the fact that $R$ is continuous, and $SU(2)$ is a connected manifold.

### Remarks

1. We will return to this important homomorphism at several points below. We are going to show that
   a.) $Ker(R) = \{\pm 1_{2\times 2}\} \cong \mathbb{Z}_2$
   b.) $R$ is surjective to $SO(3)$.

2. Note that we have just constructed a three-dimensional representation of $SU(2)$. It is a very special representation known as the *adjoint representation.*

---

**Exercise** *Simple Identities For $\vec{x} \cdot \vec{\sigma}$*
a.) Prove (5.40)
b.) Prove (5.41)
c.) Show that both these formulae imply that if $\vec{x} \cdot \vec{\sigma} = \vec{y} \cdot \vec{\sigma}$ then $\vec{x}^2 = \vec{y}^2$.

---

---

**Exercise** *A Formula For The Angle Between Vectors In Terms Of Spherical Angles*
a.) Show that if $\gamma$ is the angle between $\vec{x}$ and $\vec{y}$ then that

$$\mathrm{Tr}(\vec{x} \cdot \vec{\sigma})(\vec{y} \cdot \vec{\sigma}) = 2|\vec{x}||\vec{y}| \cos \gamma \tag{5.44}$$

b.) Now parametrize $\vec{x}$ and $\vec{x}'$ in terms of standard polar angles:

$$\begin{aligned} x^1 &= r \sin\theta \cos\phi \\ x^2 &= r \sin\theta \sin\phi \\ x^3 &= r \cos\theta \end{aligned} \tag{5.45}$$

Show that the angle $\gamma$ between $\vec{x}$ and $\vec{x}'$ is given by

$$\cos\gamma = \cos\theta \cos\theta' + \cos(\phi - \phi') \sin\theta \sin\theta' \tag{5.46}$$

---

### 5.4.2 $SL(2,\mathbb{R})$ And Lorentz Transformations In $2+1$ Dimensions

We now consider <u>*traceless*</u> and <u>*real*</u> matrices $M_2^0(\mathbb{R})$. This is a three-real dimensional vector space. In analogy to the previous section we can parametrize it as:

$$\vec{x} := (t, x, y) \in \mathbb{R}^3 \leftrightarrow \begin{pmatrix} x & y - t \\ y + t & -x \end{pmatrix} := M(\vec{x}) \tag{5.47}$$

Now the group $G = SL(2,\mathbb{R})$ acts on $M_2^0(\mathbb{R})$ via

$$\phi(A, m) := AmA^{-1} \tag{5.48}$$

for $A \in SL(2,\mathbb{R})$ and $m \in M_2^0(\mathbb{R})$. Simply note that $AmA^{-1}$ is real and traceless, if $m$ is.

Now, as before, we can use the identification of $M_2^0(\mathbb{R})$ with $\mathbb{R}^3$ to define a homomorphism $\Lambda : SL(2,\mathbb{R}) \to GL(3,\mathbb{R})$ via

$$AM(\vec{x})A^{-1} = M(\Lambda(A)\vec{x}) \tag{5.49}$$

Now note that

$$\det M_{\vec{x}} = t^2 - x^2 - y^2 = \det M_{\vec{x}'} . \tag{5.50}$$

and therefore, $\Lambda(A)$ is a *Lorentz transformation*. Moreover

$$\Lambda : A \in SL(2, R) \to \Lambda(A) \in O(1, 2) \tag{5.51}$$

is a homomorphism.

**Remarks**

1. Once again it will turn out that

$$\ker\Lambda = \mathbb{Z}_2 \qquad image(\Lambda) = SO_0(1, 2) \tag{5.52}$$

2. Once again, we are discussing the adjoint representation of $SL(2, \mathbb{R})$ on its Lie algebra.

---

**Exercise**

For $\lambda > 0$, $n \in \mathbb{R}$ and a real angle $\theta$ consider the matrix:

$$A = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \cdot \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \tag{5.53}$$

a.) Show that $A \in SL(2, \mathbb{R})$.

Note: It turns out that the general element of $SL(2, \mathbb{R})$ can always be uniquely written in this form. That is a special case of a general parametrization of Lie groups known as the $KAN$ decomposition. For $SL(2, \mathbb{R})$ (and indeed for $SL(n, \mathbb{R})$ for any $n$) it follows simply from the Gram-Schmidt procedure. This shows that as a manifold $SL(2, \mathbb{R})$ is isomorphic to $S^1 \times \mathbb{R}^2$ and provides a global system of coordinates on $SL(2, \mathbb{R})$.

b.) What kind of Lorentz transformations are described by the parameters $\theta, \lambda, n$ in the global parametrization (5.53)? Show that

$\theta$: rotation in $x, y$ plane

$\lambda$: boost along $y$ axis of rapidity $\log\lambda^2$

$n$: Defines a "null boost," best expressed in light cone coordinates:

$$\begin{aligned} t + y &\to t + y \\ x &\to x + n(t + y) \\ t - y &\to t - y + 2nx + n^2(t + y) \end{aligned} \tag{5.54}$$

---

### 5.4.3 $SL(2,\mathbb{C})$ And $O(1,3)$

We now consider an analogous construction using $\mathcal{H}_2$, the vector space of all $2\times 2$ Hermitian matrices. This is a four-dimensional real vector space and it can naturally be identified with four-dimensional Minkowski space $\mathbb{M}^{1,3}$. To see that we denote for $\mathbf{x} = (x^0, x^1, x^2, x^3) \in \mathbb{M}^{1,3}$

$$M(\mathbf{x}) := \begin{pmatrix} x^0 + x^3 & x^1 - \mathrm{i}x^2 \\ x^1 + \mathrm{i}x^2 & x^0 - x^3 \end{pmatrix} = x^0 1_{2\times 2} + \vec{x}\cdot\vec{\sigma} \tag{5.55}$$

There is a group action of $SL(2,\mathbb{C})$ on $\mathcal{H}_2$ defined by

$$\phi(A, M) = AMA^{\dagger} \tag{5.56}$$

for $A \in SL(2,\mathbb{C})$ and $M \in \mathcal{H}_2$. In close analogy to the previous example, this defines a linear action of $SL(2,\mathbb{C})$ on $\mathcal{H}_2$ which is equivalent to a linear action on $\mathbb{M}^{1,3}$ and hence we get a group homomorphism

$$\Lambda : SL(2,\mathbb{C}) \to GL(4,\mathbb{R}) \tag{5.57}$$

defined by

$$AM(\mathbf{x})A^{\dagger} = M(\Lambda(A)\mathbf{x}) \ . \tag{5.58}$$

Now note that

$$det M(\mathbf{x}) = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 \tag{5.59}$$

and since $\mathrm{det}AMA^{\dagger} = \mathrm{det}M$ we conclude that in fact we have a homomorphism:

$$\Lambda : SL(2,\mathbb{C}) \to O(1,3) \tag{5.60}$$

**Remarks**:

1. We will show that $\ker(\Lambda) = \{\pm 1_{2\times 2}\} \cong \mathbb{Z}_2$. It will turn out that $\Lambda$ is NOT surjective, but it is surjective onto the connected component of the identity, denoted $SO_0(1,3)$. See below for more about that.

2. *The Celestial Sphere And Möbius Transformations.* Consider a point $p$ on a light-ray in $\mathbb{M}^{1,3}$ that begins at the origin $x^{\mu} = 0$. Assume this point is in the forward light cone so the coordinates $x^{\mu}(p)$ have $x^0(p) > 0$. Then we know that

$$\mathrm{det}M(x^{\mu}(p)) = 0 \tag{5.61}$$

So long as $M \neq 0$ this means that there exist vectors $v, w \in \mathbb{C}^2$ so that $M$ is the outer-product of the vectors. In terms of matrix elements

$$M(x^{\mu}(p))_{A\dot{B}} = v_A w_{\dot{B}} \tag{5.62}$$

where $A \in \{1,2\}$ and $\dot{B} \in \{1,2\}$. The dot over the $B$ has a definite meaning in terms of "spinor index notation," but we are not going to explain that here. Since

$M$ is Hermitian we know that $w = v^*$. Now note that two points $p, p'$ are on the same light-ray in the forward light cone iff there is a positive real number $\lambda > 0$ so that $x^\mu(p') = \lambda x^\mu(p)$. Next, we claim that two matrices $M'$ and $M$ of the form (5.62) are related by scaling with a positive real number iff $v' = \xi v$ for some nonzero complex number $\xi$. In this way we see that there is a 1-1 correspondence between the light-rays emanating from the origin or $\mathbb{M}^{1,3}$ and points of $\mathbb{CP}^1$. It goes as follows:

Given a light-ray $\ell \subset \mathbb{M}^{1,3}$ from the origin in the forward light-cone we choose a point - any point - $p \in \ell$. Then there exists a vector $v$ (it is not unique) with components $v_A$ so that $M(x^\mu(p))_{A\dot{B}} = v_A v^*_{\dot{B}}$. The vector $v$ defines a <u>point</u> $[v_1 : v_2] \in \mathbb{CP}^1$.

There are a lot of choices here, and you should convince yourself that nevertheless there is a 1-1 correspondence between a light-ray $\ell$ and a point in $\mathbb{CP}^1$. We can say this a bit more formally as follows: Let

$$\mathcal{L}_+ = \{x^\mu | x^\mu x_\mu = 0 \qquad \text{and} \qquad x^0 > 0\} \tag{5.63}$$

denote the forward lightcone in $\mathbb{M}^{1,3}$ and

$$\mathcal{H}_2^{deg} = \{M \in \mathcal{H}_2 | \det M = 0 \qquad \text{and} \qquad M \neq 0\} \tag{5.64}$$

then there is a one-one map $f_1 : \mathcal{L}_+ \to \mathcal{H}_2^{deg}$. On the other hand there is a map $f_2 : \mathbb{C}^2 - \{0\} \to \mathcal{H}_2^{deg}$ given by

$$f_2 : \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \to \begin{pmatrix} |v_1|^2 & v_1 v_2^* \\ v_1^* v_2 & |v_2|^2 \end{pmatrix} \tag{5.65}$$

and $f_3 = f_1^{-1} \circ f_2$ is equivariant for the $\mathbb{C}^*$ action, where $\mathbb{C}^*$ acts on $\mathcal{L}_+$ via its homomorphic image $\mathbb{R}_{>0}$. (See equation (5.15) above. The homomorphism is $\varphi(z) := |z|^2$.) Looking at the space of orbits one can check there is a one-one map $\bar{f}_3 : \mathbb{CP}^1 \to \bar{\mathcal{L}}_+ := \mathcal{L}_+/\mathbb{R}_{>0}$. Here $\bar{\mathcal{L}}_+$ is the moduli space of light rays in the forward cone emanating from the origin.

On the other hand we can also define 1-1 maps

1. $p : \mathbb{CP}^1 \to \widehat{\mathbb{C}}$ defined by $p([v_1 : v_2]) := v_1/v_2$

2. $s : S^2 \to \widehat{\mathbb{C}}$ by stereographic projection

3. $t : \bar{\mathcal{L}}_+ \to S^2$ by the point on the celestial sphere pierced by a light-ray in the forward light cone.

We have the commutative diagram giving us four ways to think about the celestial sphere:

$$\begin{array}{ccc} \mathbb{CP}^1 & \xrightarrow{\bar{f}_3} & \bar{\mathcal{L}}_+ \\ \downarrow p & & \downarrow t \\ \widehat{\mathbb{C}} & \xleftarrow{s} & S^2_\infty \end{array} \tag{5.66}$$

The $SL(2,\mathbb{C})$ action on $\mathbb{C}^2 - \{0\}$ and that by its homomorphic image under $\Lambda$ on $\mathcal{L}_+$ commutes with the $\mathbb{C}^*$ action above. (Again we make use of equation (5.15).) It commutes with the $\mathbb{C}^*$ action and hence descends to an action on the four spaces in (5.66). The action of $SL(2,\mathbb{C})$ on the space of matrices with $\det M = 0$ is $M \to AMA^\dagger$ and, as we have discussed is equivalent to the action of the Lorentz group $O(1,3)$ on $\mathcal{L}_+$. From (5.62) this action is equivalent to $v \to Av$. But under the identification $\mathbb{CP}^1 \cong \widehat{\mathbb{C}}$ given by $p$ the action of $SL(2,\mathbb{C})$ becomes the Mobius action of $SL(2,\mathbb{C})$ on $\widehat{\mathbb{C}}$. In this sense, the action of the Lorentz group on the celestial sphere equivalent to the Mobius action of $SL(2,\mathbb{C})$ on $\widehat{\mathbb{C}}$.

Pursuing these ideas leads to the idea of the "twistor correspondence" which relates aspects of wave equations in (complex) Minkowski space to projective geometry

---

**Exercise** *Explicit Lorentz transformations*

a.) Show that the natural $SU(2)$ subgroup $SU(2) \subset SL(2,\mathbb{C})$ acts as rotations in $x^1, x^2, x^3$, leaving $x^0$ invariant.

b.) Show that

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \qquad \lambda \in \mathbb{C}^* \tag{5.67}$$

acts as a boost along the $x_3$ axis of rapidity $\log|\lambda|^2$ and a rotation of angle $arg(\lambda^*/\lambda)$ in the $x^1 x^2$ plane.

c.) Find a physical interpretation of the Lorentz transformation associated to

$$A = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \qquad n \in \mathbb{C} \tag{5.68}$$

d.) Find a representation for the matrix $\Lambda_{ij}$ in terms of the components of the boost velocity $v_i$.

---

---

**Exercise** *3-dimensional Euclidean hyperbolic space*

We can now understand more clearly the isometries of 3-dimensional hyperbolic space mentioned above.

Euclidean $AdS_3$ can be defined as the space of matrices.

$$X = \begin{pmatrix} X_- & X_1 + iX_2 \\ X_1 - iX_2 & X_+ \end{pmatrix} \tag{5.69}$$

where $X_\pm, X_1, X_2$ are real and:

$$\det X = X_- X_+ - |X_1 + iX_2|^2 = 1. \tag{5.70}$$

This defines a hyperboloid in $\mathbb{R}^{1,3}$.

Poincare coordinates: $X_+ = 1/y, X_- = y + |z|^2/y, X_1 + iX_2 = z/y$ are introduced by the Gauss decomposition, which now covers the manifold:

$$X = \begin{pmatrix} 1 & \bar{z} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y & 0 \\ 0 & 1/y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix} \tag{5.71}$$

From this, one easily computes the metric is $(dy^2 + |dz|^2)/y^2$.

Now, isometries are plainly Lorentz transformations:

$$X \to \begin{pmatrix} a & b \\ c & d \end{pmatrix} X \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} \tag{5.72}$$

with $ad - bc = 1$. Acting on the Gauss decomposition one easily shows this is:

$$\gamma \cdot (z, y) = \left( \frac{(az+b)\overline{(cz+d)} + a\bar{c}y^2}{|cz+d|^2 + |c|^2y^2}, \frac{y|ad-bc|}{|cz+d|^2 + |c|^2y^2} \right) \tag{5.73}$$

Note that one example of such a hyperbolic space is the mass shell of a massive particle in $3+1$ dimensions. If we interpret $\mathbb{R}^{1,3}$ as momentum space with

$$P = \begin{pmatrix} P_- & P_1 + iP_2 \\ P_1 - iP_2 & P_+ \end{pmatrix} \tag{5.74}$$

Then

$$\det P = M^2 \tag{5.75}$$

is the mass-shell hyperboloid. We now see that this is a copy of 3-dimensional hyperbolic space.

---

### 5.4.4 $SU(2) \times SU(2)$ and Rotations in 4 Euclidean Dimensions

Finally, we identify 4-dimensional Euclidean space with $2 \times 2$ complex matrices of the form

$$\begin{aligned} M_x &:= \begin{pmatrix} x^4 + ix^3 & ix^1 + x^2 \\ ix^1 - x^2 & x^4 - ix^3 \end{pmatrix} \\ &= x^\mu \tau_\mu \\ &= x_4 1 + i\vec{x} \cdot \vec{\sigma} \end{aligned} \tag{5.76}$$

We will denote this space of matrices with *real* $x^\mu$ by $\mathbb{H}$. Actually, this will only be a provisional definition of the quaternions, as it is really a matrix representation of the quaternions, which will be properly defined below. Another way to think about it is:

$$\mathbb{H} = \{x^\mu \tau_\mu : x^\mu \in \mathbb{R}\} = \{M \in M_2(\mathbb{C}) : M^* = \sigma_2 M \sigma_2\} \tag{5.77}$$

Now

$$\det M_x = +x^2 \tag{5.78}$$

is the Euclidean metric in 4 dimensions. Just as above, there is a left-action of $SU(2) \times SU(2)$ on $\mathbb{H}$

$$M \to u_1 M u_2^{-1} \tag{5.79}$$

The reader should check that this really is a map of $\mathbb{H}$ to $\mathbb{H}$. Morever, the transformation considered as a linear transformation on $\mathbb{R}^4$ preserves the metric thanks to (5.78).

Thus, in the same way as above we define a homomorphism

$$R : SU(2) \times SU(2) \to SO(4) \tag{5.80}$$

Now the kernel of $R$ is again a copy of $\mathbb{Z}_2$ with the nontrivial element being $(-1, -1) \in SU(2) \times SU(2)$

---

**Exercise**

a.) Show that the set of matrices $M$ satisfying (5.76) may be identified with the set of matrices $M$ satisfying:

$$M^* = \sigma_2 M \sigma_2 \tag{5.81}$$

b.) Check that every matrix in $SU(2)$ satisfies (5.81).

---

### 5.4.5 $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ and symmetries of Anti-DeSitter space

Finally, let us consider signature $(2,2)$. The space of $2 \times 2$ real matrices $M_2(\mathbb{R})$ can be identified with $\mathbb{R}^{2,2}$ with the metric being identified with the determinant, as we saw above. Identifying $x = (T_1, T_2, X_1, X_2) \in \mathbb{R}^{2,2}$ with such a matrix via

$$M_x = \begin{pmatrix} T_1 + X_1 & X_2 + T_2 \\ X_2 - T_2 & T_1 - X_1 \end{pmatrix} \tag{5.82}$$

we have a left-action of $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$:

$$M_x \to A_1 M_x A_2^{-1} = M_{\Lambda \cdot x} \tag{5.83}$$

for $(A_1, A_2) \in SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$.

The considerations analogous to the above show that we have an exact sequence: ♣EXACT SEQUENCE HERE IS OUT OF PLACE BELONGS BELOW ♣

$$1 \to \mathbb{Z}_2 \to SL(2,\mathbb{R}) \times SL(2,\mathbb{R}) \to SO_0(2,2;\mathbb{R}) \to 1 \tag{5.84}$$

where $\mathbb{Z}_2$ in the quotient is embedded diagonally.

The subspace of matrices $\det M_x = 1$ is a copy of $SL(2,\mathbb{R})$ itself. ***** EXPLAIN MORE **** Since the equation $\det M_x = 1$ is $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ invariant we learn that $SL(2,\mathbb{R}) \times SL(2,\mathbb{R})$ is also the group of isometries of $SL(2,\mathbb{R})$. Therefore the same holds for the universal cover - three-dimensional anti-deSitter space.

## 5.5 Group Actions On Sets Induce Group Actions On Associated Function Spaces

The following general abstract idea is of great importance in both mathematics and physics: Suppose $X$ and $Y$ are any two sets and $\mathcal{F}[X \to Y]$ is the set of functions from $X$ to $Y$. Now suppose that there is a left $G$-action on $X$ defined by $\phi : G \times X \to X$. Then, automatically, there is also a $G$ action $\tilde{\phi}$ on $\mathcal{F}[X \to Y]$. To define it, suppose $F \in \mathcal{F}[X \to Y]$ and $g \in G$. Then we need to define $\tilde{\phi}(g, F) \in \mathcal{F}[X \to Y]$. We do this by setting $\tilde{\phi}(g, F)$ to be that specific function whose values are defined by:

$$\tilde{\phi}(g, F)(x) := F(\phi(g^{-1}, x)). \tag{5.85}$$

Note the inverse of $g$ on the RHS. It is there so that the group law works out:

$$\begin{aligned} \tilde{\phi}(g_1, \tilde{\phi}(g_2, F))(x) &= \tilde{\phi}(g_2, F)(\phi(g_1^{-1}, x)) \\ &= F(\phi(g_2^{-1}, \phi(g_1^{-1}, x)) \\ &= F(\phi(g_2^{-1} g_1^{-1}, x)) \\ &= F(\phi((g_1 g_2)^{-1}, x)) \\ &= \tilde{\phi}(g_1 g_2, F)(x) \end{aligned} \tag{5.86}$$

and hence $\tilde{\phi}(g_1, \tilde{\phi}(g_2, F)) = \tilde{\phi}(g_1 g_2, F)$ as required for a group action. It should also be clear that $\tilde{\phi}(1_G, F) = F$.

In less cumbersome notation we would simply write

$$(g \cdot F)(x) := F(g^{-1} \cdot x) \tag{5.87}$$

In the above discussion we could impose various conditions, on the functions in $\mathcal{F}[X \to Y]$. For example, if $X$ and $Y$ are manifolds we could ask our maps to be continuous, differentiable, etc. The above discussion would be unchanged.

As just one (important) example of this general idea: In field theory if we have fields on a spacetime, and a group of symmetries acting on that spacetime, such as the ones we studied in the previous section, then that group also acts on the space of fields.

---

**Exercise** *When Y Is A G-Set*

Suppose there is a left $G$-action on a set $Y$ and $X$ is any set. Show that there is a natural left $G$-action on $\mathcal{F}[X \to Y]$.

---

# 6. The Symmetric Group.

The symmetric group is an important example of a finite group. As we shall soon see, all finite groups are isomorphic to subgroups of the symmetric group.

Recall from section 3 above that for any set $X$ we can define a group $S_X$ of all permutations of the set $X$. If $n$ is a positive integer the symmetric group on $n$ elements, denoted $S_n$, is defined as the group of permutations of the set $X = \{1, 2, \ldots, n\}$.

In group theory, as in politics, there are leftists and rightists and we can actually define *two* group operations:

$$\begin{aligned} (\phi_1 \cdot_L \phi_2)(i) &:= \phi_2(\phi_1(i)) \\ (\phi_1 \cdot_R \phi_2)(i) &:= \phi_1(\phi_2(i)) \end{aligned} \tag{6.1}$$

That is, with $\cdot_L$ we read the operations from left to right and first apply the permutation $\phi_1$, and then the next permutation to the right, namely $\phi_2$. With $\cdot_R$ it is the other way around: We apply the maps in the order of reading from right to left. Each convention has its own advantages and both are frequently used. For example, if we think of the permutation action as an arrow taking $i$ to some $i'$ then the subsequent composition of arrows is most naturally represented using $\cdot_L$. However, if we think of a permutation as a function then $\cdot_R$ is more natural because we are composition functions.

In these notes we will adopt the $\cdot_R$ convention and henceforth simply write $\phi_1\phi_2$ for the product. It is

We can write a permutation symbolically as

$$\phi = \begin{pmatrix} 1 & 2 & \cdots & n \\ p_1 & p_2 & \cdots & p_n \end{pmatrix} \tag{6.2}$$

meaning: $\phi(1) = p_1, \phi(2) = p_2, \ldots, \phi(n) = p_n$. Note that we could equally well write the same permutation as:

$$\phi = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ p_{a_1} & p_{a_2} & \cdots & p_{a_n} \end{pmatrix} \tag{6.3}$$

where $a_1, \ldots, a_n$ is any permutation of $1, \ldots, n$. With this understood, suppose we want to compute $\phi_1 \cdot_L \phi_2$. We should first see what $\phi_1$ does to the ordered elements $1, \ldots, n$, and then see what $\phi_2$ does to the ordered output from $\phi_1$. So, if we write:

$$\begin{aligned} \phi_1 &= \begin{pmatrix} 1 & \cdots & n \\ q_1 & \cdots & q_n \end{pmatrix} \\ \phi_2 &= \begin{pmatrix} q_1 & \cdots & q_n \\ p_1 & \cdots & p_n \end{pmatrix} \end{aligned} \tag{6.4}$$

Then

$$\phi_1 \cdot_L \phi_2 = \begin{pmatrix} 1 & \cdots & n \\ p_1 & \cdots & p_n \end{pmatrix} \tag{6.5}$$

On the other hand, to compute $\phi_1 \cdot_R \phi_2$ we should first see what $\phi_2$ does to $1, \ldots, n$ and then see what $\phi_1$ does to that output. We could write represent this as:

$$\phi_2 = \begin{pmatrix} 1 & \cdots & n \\ q'_1 & \cdots & q'_n \end{pmatrix}$$
$$\phi_1 = \begin{pmatrix} q'_1 & \cdots & q'_n \\ p'_1 & \cdots & p'_n \end{pmatrix} \tag{6.6}$$

and then

$$\phi_1 \cdot_R \phi_2 = \begin{pmatrix} 1 & \cdots & n \\ p'_1 & \cdots & p'_n \end{pmatrix} \tag{6.7}$$

---

**Exercise**

a.) Show that the order of the group is $|S_n| = n!$.

b.) Show that if $n_1 \leq n_2$ then we can consider $S_{n_1}$ as a subgroup of $S_{n_2}$.

c.) In how many ways can you consider $S_2$ to be a subgroup of $S_3$? [36]

d.) In how many ways can you consider $S_{n_1}$ to be a subgroup of $S_{n_2}$ when $n_1 \leq n_2$ ? [37]

---

---

**Exercise** Show that the inverse of (6.2) is the permutation:

$$\phi = \begin{pmatrix} p_1 & p_2 & \cdots & p_n \\ 1 & 2 & \cdots & n \end{pmatrix} \tag{6.8}$$

---

It is often useful to visualize a permutation in terms of "time evolution" (going up) as shown in 1.

♣SHOW THE PERMUTATIONS $\phi_1, \phi_2$ MORE CLEARLY IN THE PICTURE ♣

---

**Exercise** *Left versus right*

a.) Show that in the pictorial interpretation the inverse is obtained by running arrows backwards in time.

b.) Show that the left- and right- group operation conventions are related by

$$\phi_1 \cdot_L \phi_2 = (\phi_1^{-1} \cdot_R \phi_2^{-1})^{-1} \tag{6.9}$$

[36] *Answer*: There are three subgroups of $S_3$ isomorphic to $S_2$. They are the subgroups that leave one element unchanged. That is the permutations that leave $1, 2$ or $3$ unchanged.

[37] *Answer*: for any subset $T \subset \{1, \dots, n_2\}$ of cardinality $n_2 - n_1$ we can consider the subset of permutations that leave all elements of $T$ unchanged. This subset of permutations will be a subgroup isomorphic to $S_{n_1}$. So there are $\binom{n_2}{n_1}$ distinct subgroups isomorphic to $S_{n_1}$.

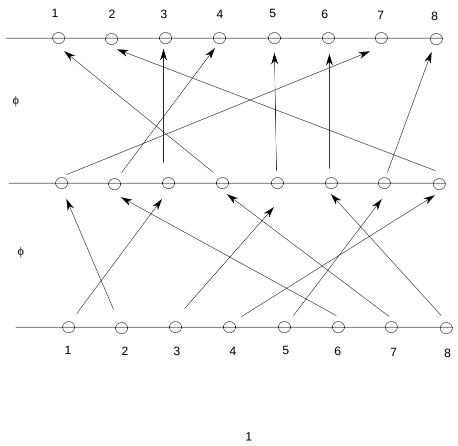



**Figure 1:** A pictorial view of the composition of two permutations $\phi_1, \phi_2$ in $S_8$. Thus $1 \to 3, 2 \to 7$ etc. for the group product $\phi_2 \cdot \phi_1$.

c.) Interpret (6.9) as the simple statement that $\phi_1 \cdot_R \phi_2$ puts $\phi_2$ in the past while $\phi_1 \cdot_L \phi_2$ puts $\phi_1$ in the past.

---

The next two exercises assume some familiarity with concepts from linear algebra. See Chapter 2 below if they are not familiar.

---

**Exercise** *The Canonical Permutation Representation Of* $S_n$

Consider the standard Euclidean vector space $\mathbb{R}^n$ (or $\mathbb{C}^n$ or $\kappa^n$) with basis vectors $\vec{e}_1, \dots, \vec{e}_n$ where $\vec{e}_i$ has component 1 in the $i^{th}$ position and zero else. Note that the symmetric group permutes these vectors in an obvious way:

$$T(\phi) : \vec{e}_i \to \vec{e}_{\phi(i)} \ , \tag{6.10}$$

and now extend by linearity so that

$$T(\phi) : \sum_{i=1}^{n} x_i e_i \mapsto \sum_{i=1}^{n} x_i e_{\phi(i)} = \sum_{i=1}^{n} x_{\phi^{-1}(i)} e_i \tag{6.11}$$

Thus to any permutation $\phi \in S_n$ we can associate a linear transformation $T(\phi)$ on $\kappa^n$.

a.) Show that

$$T(\phi_1) \circ T(\phi_2) = T(\phi_1 \circ \phi_2) = T(\phi_1 \cdot_R \phi_2) \tag{6.12}$$

This means we have a linear representation of the group $S_n$.

b.) The matrix $A(\phi)$ of $T(\phi)$ defined by $T(\phi)$ and the ordered basis $\{e_1, \dots, e_n\}$ is defined by:

$$T(\phi)\vec{e}_i = \sum_{j=1}^{n} A(\phi)_{ji} \vec{e}_j \tag{6.13}$$

b.) Show that

$$A(\phi_1)A(\phi_2) = A(\phi_1 \circ \phi_2) \tag{6.14}$$

and in particular that $A(\phi^{-1}) = A(\phi)^{-1}$. Thus, $\phi \to A(\phi)$ is a matrix representation of $S_n$.

c.) Write out $A(\phi)$ for small values of $n$ and some simple permutations $\phi$.

d.) Write a general formula for the matrix elements of $A(\phi)$. [38]

e.) The matrices $A(\phi)$ are called *permutation matrices*. In each row and column there is only one nonzero matrix element, and that nonzero element is 1. If $B$ is any other $n \times n$ matrix show that

$$\left(A(\phi)^{-1} B A(\phi)\right)_{i,j} = B_{\phi(i),\phi(j)} \tag{6.15}$$

In general, if we have a representation of a group $T : G \to GL(V)$ and a nontrivial subspace $W \subset V$ such that $T(g)$ takes vectors in $W$ to vectors in $W$ for all $g \in G$, we say that the representation is *reducible*. See section 11.7 for more information.

f.) Show that the natural permutation representation of $S_n$ on $\mathbb{R}^n$ is reducible. [39]

---

---

**Exercise** *Signed Permutation Matrices*

Define *signed permutation matrices* to be invertible matrices such that in each row and column there is only one nonzero matrix element, and the nonzero matrix element can be either $+1$ or $-1$. Finally, require the matrix to be invertible.

a.) Show that the set of $n \times n$ signed permutation matrices form a group. We will call it $W(B_n)$ for reasons that will not be obvious for a while.

b.) Define a group homomorphism $W(B_n) \to S_n$.

---

### 6.1 Cayley's Theorem

As a nice illustration of some of the concepts we have introduced we now prove Cayley's theorem. This theorem states that *any* finite group is isomorphic to a subgroup of a permutation group $S_N$ for some $N$.

Recall the notion of a group action of a group $G$ on a set $X$. In this case we take $X = G$ and $G$ is acting on itself by *left-multiplication*, defined as follows:

For $h \in G$, define the map $L(h) : G \to G$ by the rule:

$$L(h) : g \mapsto h \cdot g \qquad \forall g \in G. \tag{6.16}$$

[38] *Answer*: $A(\phi)_{i,j} = \delta_{i,\phi(j)} = \delta_{\phi^{-1}(i),j}$.

[39] *Answer*: Show that the linear subspace spanned by the "all ones vector" $v_0 = e_1 + \cdots + e_n$ is preserved under the action of $T(\phi)$ for all $\phi \in S_n$: $T(\phi)(\lambda v_0) = \lambda v_0$, for all $\lambda \in \mathbb{R}$.

This map is one-one and invertible so $L(h) \in S_G$, the group of permutations of the set $G$. Now note that

$$L(h_1) \circ L(h_2) = L(h_1 \cdot h_2) \tag{6.17}$$

so the map $\mathcal{L}$ defined by $\mathcal{L} : h \mapsto L(h)$ is a homomorphism

$$\mathcal{L} : G \to S_G \tag{6.18}$$

$\mathcal{L}$ is the quantity denoted by $\Phi$ above in our general discussion of group actions. Furthermore, if $L(h_1) = L(h_2)$ then $h_1 = h_2$. Therefore $\mathcal{L}$ is an isomorphism of $G$ with its image in $S_G$.

The above remarks apply to any group. However, now consider any finite group $G$ with $N = |G|$ then $S_G$ is isomorphic to $S_N$. Therefore, any finite group is isomorphic to a subgroup of a symmetric group $S_N$ for some $N$. This is Cayley's theorem. Note that which subgroup of $S_N$ we obtain depends on how we choose to order $G$, that is, it depends on the choice of isomorphism $S_G \cong S_N$. To produce an isomorphism we need to choose a total ordering on $G$, but in general there is no natural ordering on a finite group $G$.

---

**Exercise** *Concrete Example*

By Cayley's theorem the cyclic group $\mathbb{Z}_n$ of order $n$ is isomorphic to a subgroup of a permutation group. Exhibit such an isomorphic subgoup. [40]

---

---

**Exercise** *Right Action*

♣This is redundant with some material on group actions below. ♣

There are other ways $G$ can act on itself. For example we can define

$$R(h) : g \mapsto g \cdot h \tag{6.19}$$

a.) Show that $R(h)$ permutes the elements of $G$.

b.) Show that $R(h_1) \circ R(h_2) = R(h_2 h_1)$. Thus, $h \mapsto R(h)$ is not a homomorphism of $G$ into the group $S_G$ of permutations of $G$.

c.) Show that $h \mapsto R(h^{-1})$ is a homomorphism of $G$ into $S_G$.

---

### 6.2 Cyclic Permutations And Cycle Decomposition

A very important class of permutations are the *cyclic permutations of length* $\ell$. Choose $\ell$ distinct numbers, $a_1, \dots, a_\ell$ between 1 and $n$ and permute:

$$a_1 \to a_2 \to \cdots \to a_\ell \to a_1 \tag{6.20}$$

[40] *Answer*: Choose any cyclic permutation of length $n$ (Cyclic permutations are defined in Section 6.2 below.) Then it generates a subgroup of $S_N$ of length $n$ for any $N \geq n$.

holding all other $n-\ell$ elements fixed. Such a permutation is called *a cycle of length $\ell$*. We will denote such permutations as:

$$\phi = (a_1 a_2 \dots a_\ell). \tag{6.21}$$

Bear in mind that with this notation the same permutation can be written in $\ell$ different ways:

$$(a_1 a_2 \dots a_\ell) = (a_2 a_3 \dots a_\ell a_1) = (a_3 \dots a_\ell a_1 a_2) = \cdots = (a_\ell a_1 a_2 \dots a_{\ell-1}) \tag{6.22}$$

Let us write out the elements of the first few symmetric groups in this notation:

$$S_2 = \{1, (12)\} \tag{6.23}$$

$$S_3 = \{1, (12), (13), (23), (123), (132)\} \tag{6.24}$$

**Remarks**

1. $S_2$ is abelian.
2. $S_3$ is NOT ABELIAN [41]

$$\begin{aligned} (12)\cdot(13) &= (132) \\ (13)\cdot(12) &= (123) \end{aligned} \tag{6.25}$$

   and therefore so is $S_n$ for $n > 2$.

It is not true that all permutations are just cyclic permutations, as we first see by considering $S_4$:

$$\begin{aligned} S_4 = \{&1, (12), (13), (14), (23), (24), (34), (12)(34), (13)(24), (14)(23), \\ &(123), (132), (124), (142), (134), (143), (234), (243) \\ &(1234), (1243), (1324), (1342), (1423), (1432)\} \end{aligned} \tag{6.26}$$

Now a key observation is:

*Any permutation $\sigma \in S_n$ can be uniquely written as a product of disjoint cycles.* This is called the cycle decomposition of $\sigma$.

For example

$$\sigma = (12)(34)(10,11)(56789) \tag{6.27}$$

is a cycle decomposition in $S_{11}$. There are 3 cycles of length 2 and 1 of length 5.

The decomposition into products of disjoint cycles is known as the *cycle decomposition.*

---

[41]Note that $(12)\cdot_L(13) = (123)$. But we use the $\cdot_R$ convention.

**Exercise** *Decomposition as a product of disjoint cyclic permutations*

Prove the above claim: every permutation above is a product of cyclic permutations on disjoint sets of integers. [42]

---

---

**Exercise**

a.) Let $\phi$ be a cyclic permutation of order $\ell$. Suppose we compose $\phi$ with itself $N$ times. Show that the result is the identity transformation iff $\ell$ divides $N$.

b.) Suppose $\phi$ has a cycle decomposition with cycles of length $k_1, \dots, k_s$. What is the smallest number $N$ so that if we compose $\phi$ with itself $\phi \circ \cdots \circ \phi$ for $N$ times that we get the identity transformation?

---

### 6.3 Transpositions

A *transposition* is a permutation of the form: $(ij)$. These satisfy some nice properties: Suppose $i, j, k$ are distinct. You can check as an exercise that transpositions obey the following identities:

$$\begin{aligned} (ij)\cdot(jk)\cdot(ij) = (ik) &= (jk)\cdot(ij)\cdot(jk) \\ (ij)^2 &= 1 \\ (ij)\cdot(kl) &= (kl)\cdot(ij) \qquad \{i,j\}\cap\{k,l\} = \emptyset \end{aligned} \tag{6.28}$$

The first identity is illustrated in Figure 2. Draw the other two.

We observed above that there is a cycle decomposition of permutations. Now note that

*Any cycle* $(a_1, \cdots, a_k)$ *can be written as a product of transpositions.*

The explicit formula is (nota bene: all the $a_j$ are distinct!):

$$(a_1, a_k)(a_1, a_{k-1})\cdots(a_1, a_2) = (a_1, a_2, a_3, \dots, a_k) \tag{6.29}$$

Therefore, *every element of* $S_n$ *can be written as a product of transpositions, generalizing* (6.25). We say that the transpositions *generate* the permutation group. Taking products of various transpositions – what we might call a "word" whose "letters" are the transpositions – we can produce any element of the symmetric group. We will return to this notion in §10 below.

Of course, a given permutation can be written as a product of transpositions in many ways. This clearly follows because of the identities (6.28). A nontrivial fact is that the

[42] *Answer*: Use induction: Consider any element, say $x \in \{1, \dots, n\}$ and let $\phi$ be a permutation. Consider the elements $x, \phi(x), \phi(\phi(x)), \dots$. This must be a finite set $C$, so we get a cyclic permutation of the elements in $C$. Then $\phi$ must permute all the elements in $\{1, \dots, n\} - C$. But this has cardinality strictly smaller than $n$. So, use the inductive hypothesis.

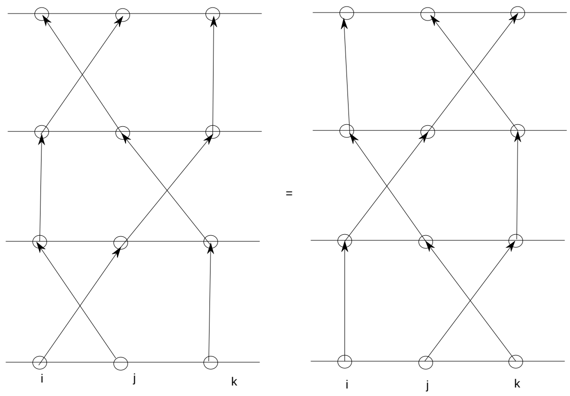



**Figure 2:** Pictorial illustration of equation (4.21) line one for transpositions where $i < j < k$. Note that the identity is suggested by "moving the time lines" holding the endpoints fixed. Reading time from bottom to top corresponds to reading the composition from left to right in the $\cdot_R$ convention.

transpositions together with the above relations generate precisely the symmetric group. [43] It therefore follows that all possible nontrivial identities made out of transpositions follow from repeated use of these identities.

Although permutations can be written as products of transpositions in different ways, the number of transpositions in a word *modulo 2* is always the same, because the identities (6.28) have the same number of transpositions, modulo two, on the LHS and RHS. Thus we can define *even, resp. odd, permutations* to be products of even, resp. odd numbers of transpositions.

**Definition:** The *alternating group* $A_n \subset S_n$ is the subgroup of $S_n$ of even permutations.

---

**Exercise**

a.) What is the order of $A_n$ ? [44]

b.) Write out $A_2$, $A_3$, and $A_4$. Show that $A_3$ is isomorphic to $\mathbb{Z}_3$. [45]

c.) $A_3$ is Abelian. Is $A_4$ Abelian? [46]

[43] This follows once one has shown that the Coxeter presentation given below gives precisely the symmetric group, and not some larger group (requiring the imposition of further relations) since the above relations all follow from the Coxeter relations.

[44] *Answer*: $\frac{1}{2}n!$ for $n > 1$. To prove this note that the transformation $\phi \to \phi \circ (12)$ is an invertible transformation $S_n \to S_n$ that squares to the identity. On the other hand, it exchanges even and odd permutations.

[45] *Answer*: $A_2 = \{1\}$. $A_3 = \{1, (123), (132)\}$.

$$A_4 = \{1, (123), (132), (124), (142), (134), (143), (234), (243), (12)(34), (13)(24), (14)(23)\}.$$

[46] *Answer*: No. Just multiply a few elements to find a counterexample. For example $(123)(134) = (234)$ but $(134)(123) = (124)$.

**Exercise** *Alternative Proof Of* (6.29)
Give an alternative proof of (6.29) along the following lines:
a.) First show that

$$(1,k)(1,k-1)\cdots(1,4)(1,3)(1,2) = (1,2,3,4,\dots,k) \tag{6.30}$$

b.) Now, consider a permutation that takes

$$1 \to a_1, \quad 2 \to a_2, \quad 3 \to a_3, \cdots, k \to a_k \tag{6.31}$$

For our purposes, it won't really matter what it does to the other integers greater than $k$. Choose any such permutation and call it $\phi$. Note that

$$\phi \circ (1\ 2\ \cdots\ k) \circ \phi^{-1} = (a_1\ a_2\ \cdots\ a_k) \tag{6.32}$$

c.) Now multiply the above identity by $\phi$ on the left and $\phi^{-1}$ on the right to get:

$$\phi(1,k)\phi^{-1}\phi(1,k-1)\phi^{-1}\cdots\phi(1,4)\phi^{-1}\phi(1,3)\phi^{-1}\phi(1,2)\phi^{-1} = (a_1,a_2,\dots,a_k) \tag{6.33}$$

but $\phi(1,j)\phi^{-1} = (a_1,a_j)$. So we get a decomposition of $(a_1\ a_2\ \cdots\ a_k)$ as a product of transpositions.

In general, a group element of the form $ghg^{-1}$ is called a *conjugate of* $h$. See Section 7.2 below.

---

---

**Exercise**

When do two cyclic permutations commute? Illustrate the answer with pictures, as above.

---

---

**Exercise** *A Smaller Set Of Generators*

Show that from the transpositions $\sigma_i := (i,i+1)$, $1 \le i \le n-1$ we can generate all other transpositions in $S_n$. These are sometimes called the elementary generators.

---

---

**Exercise** *An Even Smaller Set Of Generators*

Show that, in fact, $S_n$ can be generated by just two elements: (12) and $(1\ 2\ \cdots\ n)$. [47]

---

---

**Exercise** *Center of $S_n$*
What is the center of $S_n$? [48]

---

---

**Exercise** *Decomposing the reverse shuffle*
Consider the permutation which takes $1, 2, \ldots, n$ to $n, n-1, \ldots, 1$.
a.) Write the cycle decomposition.
b.) Write a decomposition of this permutation in terms of the *elementary generators* $\sigma_i$. [49]

---

**Example 4.2** *The sign homomorphism.*
This is a very important example of a homomorphism:

$$\epsilon : S_n \to \mathbb{Z}_2 \tag{6.34}$$

where we identify $\mathbb{Z}_2$ as the multiplicative group $\{\pm 1\}$ of square roots of 1. The rule is:
$\epsilon : \sigma \to +1$ if $\sigma$ is a product of an *even* number of transpositions.
$\epsilon : \sigma \to -1$ if $\sigma$ is a product of an *odd* number of transpositions.

Put differently, we could define $\epsilon(ij) = -1$ for any transposition. This is compatible with the words defining the relations on transpositions. Since the transpositions generate the group the homomorphism is well-defined and completely determined.

In physics one often encounters the sign homomorphism in the guise of the "epsilon tensor" denoted:

$$\epsilon_{i_1 \cdots i_n} \tag{6.35}$$

Its value is:

1. $\epsilon_{i_1 \cdots i_n} = +1$ if

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \tag{6.36}$$

is an even permutation.

[47] *Answer*: Conjugate (12) by the $n$-cycle to get (23). Then conjugate again to get (34) and so forth. Now we have the set of generators of the previous exercise.

[48] *Answer*: If $n = 2$ then $S_n$ is Abelian and the center is all of $S_2$. If $n > 2$ then the center is the trivial group. To prove this suppose $z \in Z(S_n)$. If $z$ is not the trivial element then it moves some $i$ to some $j$. WLOG we can say it moves 1 to $i \neq 1$. Then $z(i) \neq i$. If $z(i) = 1$ then $z$ is the transposition $(1, i)$. If $n > 2$ there will be some other $j \neq 1, i$ and $z$ will not commute with $(1, j)$. If $z(i) = j$ with $j \neq 1, i$ then $\phi = (1, i)$ does not commute with $z$ because $z\phi$ takes $1 \to j$ and $\phi z$ takes $1 \to 1$.

[49] *Hint*: Use the pictorial interpretation mentioned above.

2. $\epsilon_{i_1\cdots i_n} = -1$ if

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix} \tag{6.37}$$

is an odd permutation.

3. $\epsilon_{i_1\cdots i_n} = 0$ if two indices are repeated. (This goes a bit beyond what we said above since in that case we are not discussing a permutation.)

So, e.g. among the 27 entries of $\epsilon_{ijk}$, $1 \le i,j,k \le 3$ we have

$$\begin{aligned} \epsilon_{123} &= 1 \\ \epsilon_{132} &= -1 \\ \epsilon_{231} &= +1 \\ \epsilon_{221} &= 0 \end{aligned} \tag{6.38}$$

and so forth.

---

**Exercise**
Show that

$$\epsilon_{i_1 i_2\cdots i_n}\epsilon_{j_1 j_2\cdots j_n} = \sum_{\sigma\in S_n} \epsilon(\sigma)\delta_{i_1 j_{\sigma(1)}}\delta_{i_2 j_{\sigma(2)}}\cdots\delta_{i_n j_{\sigma(n)}} \tag{6.39}$$

This formula is often useful when proving identities involving determinants. An important special case occurs for $n = 3$ where it is equivalent to the rule for the cross-product of 3 vectors in $\mathbb{R}^3$:

$$\vec{A} \times (\vec{B} \times \vec{C}) = \vec{B}(\vec{A}\cdot\vec{C}) - \vec{C}(\vec{A}\cdot\vec{B}) \tag{6.40}$$

---

### 6.4 Diversion and Example: Card shuffling

One way we commonly encounter permutation groups is in shuffling a deck of cards.

A deck of cards is equivalent to an ordered set of 52 elements. Some aspects of card shuffling and card tricks can be understood nicely in terms of group theory.

Mathematicians often use the *perfect shuffle* or the *Faro shuffle*. Suppose we have a deck of $2n$ cards, so $n = 26$ is the usual case. There are actually two kinds of perfect shuffles: the In-shuffle and the Out-shuffle.

In either case we begin by splitting the deck into two equal parts, and then we interleave the two parts perfectly.

Let us call the top half of the deck the left half-deck and the bottom half of the deck the right half-deck. Then, to define the *Out-shuffle* we put the top card of the left deck on top, followed by the top card of the right deck underneath, and then proceed to interleave them perfectly. The bottom and top cards stay the same.

If we number the cards $0, 1, \dots, 2n-1$ from top to bottom then the top (i.e. left) half-deck consists of the cards numbered $0, 1, \dots, n-1$ while the bottom (i.e. right) half-deck consists of the cards $n, n+1, \dots, 2n-1$. Then the Out-shuffle gives the cards in the new order

$$0, n, 1, n+1, 2, n+2, \dots, n+2, 2n-2, n-1, 2n-1 \tag{6.41}$$

Another way to express this is that the Out-shuffle defines a permutation of $\{0, 1, \dots, 2n-1\}$. If we let $C_x$, $0 \le x \le 2n-1$ denote the cards in the original order then the new ordered set of cards $C'_x$ are related to the old ones by:

$$C'_{\mathcal{O}(x)} = C_x \tag{6.42}$$

where

$$\mathcal{O}(x) = \begin{cases} 2x & x \le n-1 \\ 2x - (2n-1) & n \le x \le 2n-1 \end{cases} \tag{6.43}$$

Note that this already leads to a card trick: Modulo $(2n-1)$ the operation is just $x \to 2x$, so if $k$ is the smallest number with $2^k = 1 \mathrm{mod}(2n-1)$ then $k$ Out-shuffles will restore the deck perfectly.

For example: For a standard deck of 52 cards, $2^8 = 5 \times 51 + 1$ so 8 perfect Out-shuffles restores the deck!

We can also see this by working out the cycle presentation of the Out-shuffle:

$$\begin{aligned} \mathcal{O} = {} & (0)(1,2,4,8,16,32,13,26)(3,6,12,24,48,45,39,27) \\ & (5,10,20,40,29,7,14,28)(9,18,36,21,42,33,15,30) \\ & (11,22,44,37,23,46,41,31)(17,34)(19,38,25,50,49,47,43,35)(51) \end{aligned} \tag{6.44}$$

Clearly, the $8^{th}$ power gives the identity permutation.

Now, to define the *In-shuffle* we put the top card of the right half-deck on top, then the top card of the left half-deck underneath, and then proceed to interleave them.

Now observe that if we have a deck with $2n$ cards $\mathcal{D}(2n) := \{0, 1, \dots, 2n-1\}$ and we embed it in a Deck with $2n+2$ cards

$$\mathcal{D}(2n) \to \mathcal{D}(2n+2) \tag{6.45}$$

by the map $x \to x+1$ then *the Out-shuffle on the deck $\mathcal{D}(2n+2)$ permutes the cards $1, \dots, 2n$ amongst themselves and acts as an In-shuffle on these cards!*

♣Explain this some more, e.g. by illustrating with a pack of 6 cards. ♣

Therefore, applying our formula for the Out-shuffle we find that the In-shuffle is given by the formula

$$\mathcal{I}(x) = \begin{cases} 2(x+1) - 1 & x+1 \le n \\ 2(x+1) - (2n+1) - 1 & n \le x \le 2n-1 \end{cases} \tag{6.46}$$

One can check that this is given by the uniform formula

$$\mathcal{I}(x) = (2x+1) \,\mathrm{mod}(2n+1) \tag{6.47}$$

for $x \in \mathcal{D}(2n)$.

For $2n = 52$ this turns out to be one big cycle!

$$\begin{aligned}(0, 1, 3, 7, 15, 31, 10, 21, 43, 34,16, 33, 14, 29, 6, 13, 27, 2, 5,\\ 11, 23, 47, 42, 32, 12, 25, 51, 50,48, 44, 36, 20, 41, 30, 8, 17,\\ 35, 18, 37, 22, 45, 38, 24,49, 46, 40, 28, 4, 9, 19, 26)\end{aligned} \tag{6.48}$$

so it takes 52 consecutive perfect In-shuffles to restore the deck.

One can do further magic tricks with In- and Out-shuffles. As one example there is a simple prescription for bringing the top card to any desired position, say, position $\ell$ by doing In- and Out-shuffles.

To do this we write $\ell$ in its binary expansion:

$$\ell = 2^k + a_{k-1}2^{k-1} + \cdots + a_1 2^1 + a_0 \tag{6.49}$$

where $a_j \in \{0, 1\}$. Interpret the coefficients 1 as In-shuffles and the coefficients 0 as Out-shuffles. Then, reading from left to right, perform the sequence of shuffles given by the binary expression: $1a_{k-1}a_{k-2}\cdots a_1a_0$.

To see why this is true consider iterating the functions $o(x) = 2x$ and $i(x) = 2x + 1$. Notice that the sequence of operations given by the binary expansion of $\ell$ are

$$\begin{aligned} 0 &\to 1\\ &\to 2 \cdot 1 + a_{k-1}\\ &\to 2 \cdot (2 \cdot 1 + a_{k-1}) + a_{k-2} = 2^2 + 2a_{k-1} + a_{k-2}\\ &\to 2 \cdot (2^2 + 2a_{k-1} + a_{k-2}) + a_{k-3} = 2^3 + 2^2 a_{k-1} + 2a_{k-2} + a_{k-3}\\ &\vdots \quad \vdots\\ &\to 2^k + a_{k-1}2^{k-1} + \cdots + a_1 2^1 + a_0 = \ell \end{aligned} \tag{6.50}$$

For an even ordered set we can define a notion of permutations preserving *central symmetry.* For $x \in D_{2n}$ let $\bar{x} = 2n - 1 - x$. Then we define the group $W(B_n) \subset S_{2n}$ to be the subgroup of permutations which permutes the pairs $\{x, \bar{x}\}$ amongst themselves.

Note that there is clearly a homomorphism

$$\phi : W(B_n) \to S_n \tag{6.51}$$

Moreover, both $\mathcal{O}$ and $\mathcal{I}$ are elements of $W(B_n)$. Therefore the *shuffle group*, the group generated by these is a subgroup of $W(B_n)$. Using this one can say some nice things about the structure of the group generated by the in-shuffle and the out-shuffle. It was completely determined in a beautiful paper (the source of the above material):

"The mathematics of perfect shuffles," P. Diaconis, R.L. Graham, W.M. Kantor, Adv. Appl. Math. **4** pp. 175-193 (1983)

It turns out that shuffles of decks of 12 and 24 cards have some special properties. In particular, special shuffles of a deck of 12 cards can be used to generate a very interesting group known as the Mathieu group $M_{12}$. It was, historically, the first "sporadic" finite simple group. See section §16.4 below.

To describe $M_{12}$ we need to introduce a *Mongean shuffle*. Here we take the deck of cards put the top card on the right. Then from the deck on the left alternatively put cards on the top or the bottom. So the second card from of the deck on the left goes on top of the first card, the third card from the deck on the left goes under the first card, and so on. If we label our deck as cards $1, 2, \ldots, 2n$ then the Mongean shuffle is:

$$m : \{1, 2, \ldots, 2n\} \to \{2n, 2n-2, \ldots, 4, 2, 1, 3, 5, \ldots, 2n-3, 2n-1\} \tag{6.52}$$

In formulae, acting on $\mathcal{D}(2n)$

$$m(x) = \mathrm{Min}[2x, 2n+1-2x] \tag{6.53}$$

In particular for $2n = 12$ we have

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} \to \{12, 10, 8, 6, 4, 2, 1, 3, 5, 7, 9, 11\} \tag{6.54}$$

which has cycle decomposition (check!)

$$(3\ 8) \cdot (1\ 12\ 11\ 9\ 5\ 4\ 6\ 2\ 10\ 7) \tag{6.55}$$

Now consider the *reverse shuffle* that simply orders the cards backwards. In general for a deck $\mathcal{D}(2n)$ with $n = 2\mathrm{mod}4$ Diaconis et. al. show that $r$ and $m$ generate the entire symmetric group. However, for a pack of 12 cards $r$ and $m$ generate the Mathieu group $M_{12}$. It turns out to have order

$$|M_{12}| = 2^6 \cdot 3^3 \cdot 5 \cdot 11 = 95040 \tag{6.56}$$

Compare this with the order of $S_{12}$:

$$12! = 2^{10} \cdot 3^5 \cdot 5^2 \cdot 7 \cdot 11 = 479001600 \tag{6.57}$$

So with the uniform probability distribution on $S_{12}$, the probability of finding a Mathieu permutation is $\frac{1}{5040} \sim 2 \times 10^{-4}$.

We mention some final loosely related facts:

1. There are indications that the Mathieu groups have some intriguing relations to string theory, conformal field theory, and K3 surfaces.

2. In the theory of $L_\infty$ algebras and associated topics, which are closely related to string field theory one encounters the concept of the $k$-shuffle...

   FILL IN.

---

**Exercise** *Cycle structure for the Mongean shuffle*

Write the cycle structure for the Mongean shuffle of a deck with 52 cards. How many Mongean shuffles of such a deck will restore the original order?

---

# 7. Cosets and conjugacy

## 7.1 Lagrange Theorem

The reader should refresh her/his memory about equivalence relations and group actions above.

**Definition 7.1.1**: Let $H \subseteq G$ be a subgroup. The set

$$gH \equiv \{gh | h \in H\} \subset G \tag{7.1}$$

is called a *left-coset* of H.

**Example 1**: $G = \mathbb{Z}, H = 2\mathbb{Z}$. There are two cosets: $H$ and $H + 1$.

**Example 2**: $G = S_3$, $H = \{1, (12)\} \cong S_2$. Cosets:

$$\begin{aligned}
1 \cdot H &= \{1, (12)\} \\
(12) \cdot H &= \{(12), 1\} = \{1, (12)\} = H \\
(13) \cdot H &= \{(13), (123)\} \\
(23) \cdot H &= \{(23), (132)\} \\
(123) \cdot H &= \{(123), (13)\} = \{(13), (123)\} = (13) \cdot H \\
(132) \cdot H &= \{(132), (23)\} = \{(23), (132)\} = (23) \cdot H
\end{aligned} \tag{7.2}$$

**Remarks**

1. Two left cosets are either *identical* or *disjoint*. Moreover, every element $g \in G$ lies in some coset. That is, the cosets define an equivalence relation by saying $g_1 \sim g_2$ if there is an $h \in H$ such that $g_1 = g_2 h$. The reader should give a direct proof of this. It is a very important exercise.

2. Cosets have a natural interpretation in terms of group actions. If $H \subset G$ is a subgroup then there is a left $H$-action on $G$ defined by

$$\phi(h, g) := gh^{-1} \tag{7.3}$$

The orbits of this action are precisely the left $H$-cosets $gH$. As we have seen: Orbits of a group action are the equivalence classes of an equivalence relation: This is another way to verify the claim in the previous remark.

---

**Exercise** *Left Cosets Define An Equivalence Relation*

Show that the relation $g_1 \sim g_2$ if there is an $h \in H$ such that $g_2 = g_1 h$ is an equivalence relation. [50]

---

The basic principle above leads to a fundamental theorem:

**Theorem 7.1.1** (Lagrange) If $H$ is a subgroup of a finite group $G$ then the order of $H$ divides the order of $G$:

$$|G|/|H| \in \mathbb{Z}_+ \tag{7.4}$$

*Proof* : If $G$ is finite $G = \amalg_1^m g_i H$ for some set of $g_i$, leading to *distinct* cosets. Now note that the order of any coset is the order of $H$, because the invertible action of left-multiplication by $g$ sets up a 1-1 correspondence between the elements of $H$ and those of $gH$. Therefore

$$|g_i H| = |H| \tag{7.5}$$

So $|G|/|H| = m$, where $m$ is the number of distinct cosets. ♠

This theorem is simple, but powerful, as we will see.

**Definition**: Thus far we have repeatedly spoken of the "order of a group $G$" and of various subsets of $G$, meaning simply the cardinality of the various sets. In addition a common terminology is to say that an element $g \in G$ *has order* $n$ if $n$ is the smallest natural number such that $g^n = 1$. If there is no such integer $n$ then $g$ is said to be of "infinite order."

Note carefully that if $g$ has order $n$ and $k$ is a natural number then $(g^n)^k = g^{nk} = 1$ and hence if $g^m = 1$ for some natural number $m$ it does not necessarily follow that $g$ has order $m$. However, as an application of Lagrange's theorem we can say the following: *If $G$ is a finite group then the order of $g$ must divide $|G|$, and in particular $g^{|G|} = 1$.* The proof is simple: Consider the subgroup generated by $g$, i.e. $\{1, g, g^2, \dots\}$. The order of this subgroup is the same as the order of $g$.

Using the same idea we can establish the following beautiful

**Corollary**: Any finite group of prime order $p$ is isomorphic to $\mu_p \cong \mathbb{Z}_p$. Moreover, such groups have exactly two subgroups: The trivial group and itself.

*Proof*: Choose a nonidentity element $g \in G$ and consider the subgroup generated by $g$ i.e,

$$\{1, g, g^2, g^3, \dots\} \tag{7.6}$$

The order of this group must divide $|G|$ so if $|G| = p$ is prime it must be the entire group. ♠

---

[50] *Answer* First, $g$ is in $gH$, so every element is in *some* coset. Second, suppose $g \in g_1 H \cap g_2 H$. Then $g = g_1 h_1$ and $g = g_2 h_2$ for some $h_1, h_2 \in H$. This implies $g_1 = g_2(h_2 h_1^{-1})$ so $g_1 = g_2 h$ for an element $h \in H$. (Indeed $h = h_2 h_1^{-1}$, but the detailed form is not important.) Now note that, as sets, $hH = H$. Indeed, as we saw in the proof of Cayley's theorem, left-multiplication by $h \in H$ just leads to a permutation of the set $H$. Therefore $g_1 H = g_2 H$.

**Definition 7.1.2**: If $G$ is any group and $H$ any subgroup then the *set of left cosets of $H$ in $G$* is denoted $G/H$. It is the set of orbits under left $H$ action on $G$ defined in (7.3). A set of the form $G/H$ is an important example of the concept of a *homogeneous space*. The cardinality of $G/H$ set is the *index of $H$ in $G$*, and denoted $[G : H]$.

**Example 1**: If $G = S_3, H = \{1, (12)\} \cong S_2$, then $G/H = \{H, (13) \cdot H, (23) \cdot H\}$, and $[G : H] = 3$.

**Example 2**: Let $G = \{1, \omega, \omega^2, \ldots, \omega^{2N-1}\} = \mu_{2N}$ where $\omega$ is a primitive $(2N)^{th}$ root of 1. Let $H = \{1, \omega^2, \omega^4, \ldots, \omega^{2N-2}\} = \mu_N$. Then $[G : H] = 2$ and $G/H = \{H, \omega H\}$.

**Example 3**: Let $G = A_4$ and $H = \{1, (12)(34)\} \cong \mathbb{Z}_2$. Then $[G : H] = 6$ and

$$G/H = \{H, (13)(24) \cdot H, (123) \cdot H, (132) \cdot H, (124) \cdot H, (142) \cdot H\} \tag{7.7}$$

**Remark** Note well! If $H \subset G$ is a subgroup and $g_1 H = g_2 H$ it does not follow that $g_1 = g_2$. All you can conclude is that there is some $h \in H$ with $g_1 = g_2 h$.

---

**Exercise** *Subgroups of $\mathbb{Z}_N$*
a.) Show that the subgroups of $\mathbb{Z}_N$ are isomorphic to the groups $\mathbb{Z}_M$ for $M|N$.
b.) For $N = 8, M = 4$ write out $H$.

---

---

**Exercise** *Is there a converse to Lagrange's theorem?*
Suppose $n||G|$, does there then exist a subgroup of $G$ of order $n$? Not necessarily! Find a counterexample. That is, find a group $G$ and an $n$ such that $n$ divides $|G|$, but $G$ has no subgroup of order $n$. [51]

---

Nevertheless, there is a very powerful theorem in group theory known as

**Theorem 7.1.2**: (Sylow's (first) theorem). Suppose $p$ is prime and $p^k$ divides $|G|$ for a nonnegative integer $k$. Then there is a subgroup $H \subset G$ of order $p^k$.

[51] *Answer:* One possible example is $A_4$, which has order 12, but no subgroup of order 6. By examining the table of groups below we can see that this is the example with the smallest value of $|G|$. Sylow's theorem (discussed below) states that if a prime power $p^k$ divides $|G|$ then there is in fact a subgroup of order $p^k$. This fails for composite numbers - products of more than one prime. Indeed, the smallest composite number is $6 = 2 \cdot 3$. Thus, in regard to a hypothetical converse to Lagrange's theorem, as soon as things can go wrong, they do go wrong.

Herstein's book, sec. 2.12, waxes poetic on the Sylow theorems and gives three proofs. We'll give a proof as an application of the class equation in section 9 below. Actually, Sylow has a bit more to say. We will explain some more about this in the next section.

---

**Exercise** *Subgroups of $A_4$*

Write down all the subgroups of $A_4$. Draw a diagram indicating how these are subgroups of each other.

---

---

**Exercise** *Orders of group elements in infinite groups*

a.) Give an example of an infinite group in which all elements, other than the identity, have infinite order. (This should be quite easy for you.) [52]

b.) Give an example of an infinite group where some group elements have finite order and some have infinite order. [53]

c.) Give an example of an infinite group where all elements have finite order. [54]

---

---

**Exercise**

Suppose a finite group $G$ has subgroups $H_i$, $i = 1, \dots, s$ of order $h_i$. Show that the least common multiple of the $h_i$ divides $|G|$. In particular if the $h_i$ are relatively prime then $\prod_i h_i$ divides $|G|$.

---

### 7.2 Conjugacy

Now introduce a notion generalizing the idea of similarity of matrices:

**Definition 7.2.1 :**

a.) A group element $h$ is *conjugate* to $h'$ if $\exists g \in G \qquad h' = ghg^{-1}$.

b.) Conjugacy defines an equivalence relation and the *conjugacy class of $h$* is the equivalence class under this relation:

$$C(h) := \{ghg^{-1} : g \in G\} \tag{7.8}$$

[52] *One possible answer*: Take $\mathbb{Z}$ or $\mathbb{Z}^n$ or ....

[53] *One possible answer*: $\mathbb{Z} \times \mathbb{Z}_N$. A more interesting example is $G = U(1)$. The roots of unity have finite order while elements of the form $\exp[2\pi i\alpha]$ with $\alpha$ an irrational real number have infinite order.

[54] *One possible answer*: Regard $U(1)$ as the group of complex numbers of modulus one. Let $G$ be the subgroup of complex numbers so that $z^N = 1$ for some integer $N$. This is the group of all roots of unity of any order. It is clearly an infinite group, and by its very definition every element has finite order. Using the notation of the next section, this group is isomorphic to $\mathbb{Q}/\mathbb{Z}$.

c.) Let $H \subseteq G, K \subseteq G$ be two subgroups. We say "$H$ is conjugate to $K$" if $\exists g \in G$ such that we have an equality of sets:

$$K = gHg^{-1} := \{ghg^{-1} : h \in H\} \tag{7.9}$$

**Remark**: In a similar manner to the case of left-cosets we can understand conjugacy classes in terms of a group action. In this case, the group $G$ acts on the set $X = G$ via

$$\phi(g, g') := gg'g^{-1} \tag{7.10}$$

The reader should check this is a well-defined group action. The orbits of this group action are the conjugacy classes.

**Example 7.2.1 :** We showed above that all cyclic permutations in $S_n$ are conjugate: The conjugacy class of any cyclic permutation of length $\ell$ is the set of all cyclic permutations of length $\ell$.

**Example 7.2.2 :** For $G = GL(n, \kappa)$, the notion of conjugacy is equivalent to similarity of matrices: $g_1, g_2 \in GL(n, \kappa)$ are conjugate if there is an $s \in GL(n, \kappa)$ so that $g_2 = sg_1s^{-1}$ are similar matrices. The next few examples examine conjugacy for some matrix subgroups of $GL(n, \kappa)$:

**Example 7.2.3 :** Consider $G = U(N)$. Then conjugacy within $U(N)$ is the same as unitary equivalence: $u_1, u_2 \in U(N)$ are conjugate in $U(N)$ if there is a $g \in U(N)$ with $u_2 = gu_1g^\dagger = gu_1g^{-1}$.

An important theorem known as the Spectral theorem tells us about the conjugacy classes of elements of the unitary group. The spectral theorem for $U(N)$ is proven by induction on $N$ in section 17 of the Linear Algebra User's Manuel. The statement of the theorem is that for every $u \in U(N)$ there is a $g \in U(N)$ with $gug^{-1} = Diag\{z_1, \dots, z_N\}$ where $|z_i| = 1$. One might think that the spectral theorem implies that the conjugacy classes are in one-one correspondence with $U(1)^N$: Indeed, given $u$ we diagonalize it using unitary matrices so there is $g \in U(N)$ with

$$gug^{-1} = Diag\{z_1, \dots, z_N\} \tag{7.11}$$

with phases $z_i$. We might be tempted conclude that the conjugacy class is parametrized by $(z_1, \dots, z_N) \in U(1)^N$. Be careful here: This is not quite correct! We need to consider the possibility that we can still conjugate $Diag\{z_1, \dots, z_N\}$ preserving the diagonal structure. In other words, we should consider the possibility that we could diagonalize $u$ to two different diagonal matrices. This is indeed possible.

Consider, for example, the permutation matrix $A(\phi)$ for $\phi \in S_N$ defined by the permutation representation of $S_N$ on $\mathbb{R}^N$. It is not hard to show that $A(\phi)$ is in $U(N)$, and, by (6.15) above:

$$A(\phi)^{-1} Diag\{z_1, \dots, z_N\} A(\phi) = Diag\{z_{\phi(1)}, \dots, z_{\phi(N)}\} \tag{7.12}$$

So if $g$ diagonalizes $u$ then so does $A(\phi)^{-1}g$ for <u>any</u> permutation $\phi$.

However, once we have taken this into account we are done: Up to conjugacy, a unitary matrix is completely characterized by its unordered set of eigenvalues. *The set of conjugacy classes in $U(N)$ can be naturally identified with set of unordered $N$-tuples of phases..*

Put differently, the symmetric group acts on the subgroup $\mathcal{D} \subset U(N)$ of diagonal matrices by conjugation by $A(\phi)$. Then since $\mathcal{D} \cong U(1)^N$ the set of conjugacy classes is therefore a space of orbits $U(1)^N/S_N$.

**Remark:** *The maximal torus.* Now suppose we have two commuting unitary matrices. Again, basic linear algebra (explained in LAUM) shows that they can be simultaneously diagonalized. That is, if $u_1, u_2 \in U(N)$ and $[u_1, u_2] = 1$ (group commutator) then there is a single $g \in U(N)$ with

♣Improve this remark ♣

$$g u_i g^{-1} = D_i \tag{7.13}$$

with $D_i$ diagonal. To prove this diagonalize $u_1$ so that it has blocks corresponding to the distinct eigenvalues:

$$g u_1 g^{-1} = z_1 1_{k_1 \times k_1} \oplus \cdots \oplus z_s 1_{k_s \times k_s} \tag{7.14}$$

Then since the $z_i$ are distinct and $u_2$, we know that $u_2$ must also be block diagonal. But then we can use unitary matrices in the different blocks to diagonalize $u_2$. This argument generalizes to show that any Abelian subgroup of $U(N)$ can be simultaneously diagonalized. In particular, any maximal Abelian subgroup will be conjugate to $\mathcal{D}$, the subgroup of diagonal unitary matrices.

The statements above for $U(N)$ generalize to all compact connected Lie groups in the following sense. We note that in $U(N)$ the subgroup $\mathcal{D}$ of diagonal matrices is isomorphic to $U(1)^N$. Since $U(1)$ is the circle, as a manifold, the product $U(1)^N$ can be viewed as an $N$-dimensional <u>torus</u>, as a manifold. In general, for any compact Lie group $G$ on can consider Abelian subgroups isomorphic to $U(1)^n$ for some $n$. Such subgroups are called *torus subgroups*. In general they will not be subgroups of diagonal matrices. We can consider subgroups of this form which are <u>maximal</u>: They are not contained in any torus subgroup of larger dimension. One can prove that maximal tori are all conjugate to each other and hence all isomorphic to $U(1)^r$ for an integer $r$ known as the *rank of the group $G$*. By a slight abuse of language one often finds such a maximal torus subgroup referred to as "the" *maximal torus.*

Note that a maximal torus might not be conjugate to a subgroup of diagonal matrices. Consider, for example, the group $G = SO(3)$. The subgroup of diagonal matrices is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. (This is an easy exercise!) A maximal torus of $SO(3)$ is isomorphic to $SO(2)$. One possible choice of maximal torus for $SO(3)$ is the subgroup of elements of the form

$$\begin{pmatrix} R(\phi) & 0 \\ 0 & 1 \end{pmatrix} \tag{7.15}$$

and any subgroup conjugate to this is also a maximal torus. Note that the maximal Abelian subgroup of diagonal matrices cannot be conjugated into a maximal torus.

**Example 7.2.4 :** Characterizing conjugacy classes in $GL(n,\kappa)$ is more complicated, because in these cases not all elements are diagonalizable. For definiteness let us consider the case $G = GL(n,\mathbb{C})$.

When discussing conjugacy classes it is very useful to introduce the tool of the characteristic polynomial. For any matrix $A \in M_n(\mathbb{C})$ we can define its *characteristic polynomial*

$$p_A(x) := \det(x1 - A) \tag{7.16}$$

Note that $p_A$ only depends on the conjugacy class of $A$:

$$p_{gAg^{-1}}(x) = p_A(x) \tag{7.17}$$

If $r$ is a root of the polynomial of $A$ then the matrix $r1 - A$ has zero determinant, so it has a nontrivial kernel (see Chapter 2) and therefore there is an eigenvector $v$ of $A$ with eigenvalue $r$:

$$Av = rv \tag{7.18}$$

Conversely, any eigenvalue of $A$ must be a root of the polynomial equation $p_A(x) = 0$. Now, if we are working over the complex numbers then the polynomial $p_A(x)$ has at least one root. Therefore, over the complex numbers, every matrix $A \in M_n(\mathbb{C})$ has at least one eigenvalue and one eigenvector.

It is very important to note that the eigenvectors of a complex matrix $A$ might not form a basis,. Here is a simple and basic example:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{7.19}$$

Then one easily checks $p_A(x) = x^2$. The only possible value for an eigenvalue of $A$ is a root of $x^2 = 0$, that is, the only possible value of an eigenvalue is 0. So an eigenvector would satisfy $Av = 0$, and indeed

$$v = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{7.20}$$

is a nontrivial eigenvector of eigenvalue 0. However, if there were a full basis of eigenvectors then since the only eigenvalue is zero we would have $A = 0$, which is a contradiction. It is not difficult to show that any matrix $A \in M_2(\mathbb{C})$ is either diagonalizable or is conjugate to a matrix of the form

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \tag{7.21}$$

The generalization of the previous example to arbitrary elements of $A \in M_n(\mathbb{C})$ for arbitrary $n$ states that any complex matrix is conjugate to its *Jordan canonical form.* See section 10.4 of LAUM for more extensive discussion. Here is an abbreviated discussion:

A Jordan canonical form is a block diagonal matrix where the blocks on the diagonals are *Jordan blocks.* A Jordan block is a matrix characterized by a complex number $\lambda$ and a positive integer $k$. $J_\lambda^{(k)}$ for $k=1$ is the $1\times 1$ matrix $(\lambda)$. For $k>1$ we define the Jordan block with eigenvalue $\lambda$ to be:

$$J_\lambda^{(k)} = \lambda 1 + N^{(k)} \tag{7.22}$$

where

$$N^{(k)} = e_{1,2} + e_{2,3} + \cdots + e_{k-1,k} \tag{7.23}$$

Choosing the standard ordered basis $\{e_1, \dots, e_k\}$ for $\kappa^k$ the matrix corresponds to the linear transformation taking

$$e_k \to e_{k-1} \to \cdots \to e_2 \to e_1 \to 0 \tag{7.24}$$

is a matrix with all matrix elements zero, except for 1's just above the diagonal. Thus

$$N^{(2)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{7.25}$$

$$N^{(3)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{7.26}$$

and so on. We argued above that $N^{(2)}$ cannot be diagonalized and a similar argument applies to $N^{(k)}$ for all $k>1$. Note that $p_{N^{(k)}}(x) = x^k$. So the eigenvalues would have to be zero, but $N^{(k)}$ is not the zero matrix. Also note that $(N^{(k)})^{k-1} \neq 0$ but $(N^{(k)})^k = 0$.

Let $A$ be any complex $N \times N$ matrix and $p_A(x) = \prod_{i=1}^s (x-\lambda_i)^{k_i}$ where $\lambda_i$ are the <u>distinct</u> roots of multiplicity $k_i$ where $k_i$ is a positive integer. Then $A$ is conjugate to a block form

$$SAS^{-1} = A_{\lambda_1} \oplus \cdots \oplus A_{\lambda_s}, \tag{7.27}$$

where each $A_{\lambda_i}$ is a $k_i \times k_i$ matrix so that $\det(x1 - A_{\lambda_i}) = (x-\lambda_i)^{k_i}$. Moreover we can conjugate each $A_{\lambda_i}$ itself into a block diagonal matrix of Jordan blocks of sizes $n_{a,i}$

$$A_{\lambda_i} = \oplus_{a=1}^{t_i} J_{\lambda_i}^{(n_{a,i})} \tag{7.28}$$

where

$$\sum_{a=1}^{t_i} n_{a,i} = k_i \tag{7.29}$$

The ordering of the blocks is not canonical and can be changed by the action of permutation matrices. Thus, the conjugacy class is characterized by the unordered set of pairs $\{(\lambda_i, \{n_{a,i}\})\}$ where $\lambda_i$ are distinct complex numbers and for each $\lambda_i$ we have a collection of positive integers such that

$$\sum n_{a,i} = n \tag{7.30}$$

To get the conjugacy classes in $GL(n, \mathbb{C})$ we require that all the $\lambda_i$ are nonzero. Note that a matrix is diagonalizable iff all its Jordan blocks are $1 \times 1$.

**Remark/Definitions**: We say that two homomorphisms $\varphi_i : G_1 \to G_2$ are conjugate if there is an element $g_2 \in G_2$ such that

$$\varphi_2(g_1) = g_2 \varphi_1(g_1) g_2^{-1} \tag{7.31}$$

for all $g_1 \in G_1$. Recall that a matrix representation of a group $G$ is a homomorphism

$$\varphi : G \to GL(n, \kappa) \tag{7.32}$$

We say two matrix representations are *equivalent representations* if the two homomorphisms are conjugate.

Put differently, two representations $T_1 : G \to GL(V_1)$ and $T_2 : G \to GL(V_2)$ are said to be equivalent if there is an equivariant map $S : V_1 \to V_2$ which is an isomorphism. See the discussion of "intertwiners" in the section on representation theory below.

**Definition**: A *class function* on a group is a function $f$ on $G$ (it can be valued in any set) such that $f$ takes the same values on conjugate group elements:

$$f(g g_0 g^{-1}) = f(g_0) \tag{7.33}$$

for all $g_0, g \in G$. Note particularly that if $\varphi$ is a matrix representation then

$$\chi_\varphi(g) := \mathrm{Tr}\varphi(g) \tag{7.34}$$

is an example of a class function. This function is called the *character of the representation.* Note that two equivalent representations must have the same character. Another, closely related class function on matrix groups is the characteristic polynomial $A \mapsto p_A(x)$. (The "function" is valued in the set of degree $n$ polynomials.)

---

**Exercise** *Conjugacy Is An Equivalence Relation*

a.) Show that conjugacy is an equivalence relation

b.) Prove that if $H$ is a subgroup of $G$ then $gHg^{-1}$ is also a subgroup of $G$. [55]

---

---

**Exercise** *Rotations In* $SO(3)$ . Consider a subgroup of $SO(3)$ defined by rotations around some axis in $\mathbb{R}^3$. Show that all such subgroups are conjugate subgroups.

---

[55] *Answer*: Note that $(gh_1g^{-1})(gh_2g^{-1}) = g(h_1h_2)g^{-1}$.

**Exercise** *Conjugacy Classes In $SU(2)$*

a.) Using the spectral theorem show that the conjugacy class of a matrix $u \in SU(2)$ is uniquely determined by its trace. [56]

b.) Show that the set of conjugacy classes in $SU(2)$ can be identified with $S^1/\mathbb{Z}_2 = [0, \pi]$.

c.) Show that the most general continuous homomorphism $\varphi : U(1) \to SU(2)$ looks like

$$\varphi : z \to \begin{pmatrix} |\alpha|^2 z + |\beta|^2 z^{-1} & \alpha\beta(z^{-1} - z) \\ \alpha^*\beta^*(z^{-1} - z) & |\alpha|^2 z^{-1} + |\beta|^2 z \end{pmatrix} \tag{7.35}$$

where $(\alpha, \beta) \in \mathbb{C}^2$ satisfy $|\alpha|^2 + |\beta|^2 = 1$ and $z \in U(1)$. (Hint: Show that all continuous homomorphisms $\varphi : U(1) \to SU(2)$ are conjugate.)

---

**Exercise** *Conjugacy Classes Of Hermitian Matrices*

There is a version of the spectral theorem for finite-dimensional Hermitian matrices. (For a proof see LAUM.) Every $n \times n$ Hermitian matrix is unitarily conjugate to a diagonal matrix of real numbers.

Consider the conjugation action of $U(n)$ on the set $H_n$ of $n \times n$ Hermitian matrices. Show that the orbits are in 1-1 correspondence with <u>unordered</u> $n$-tuples of real numbers.

---

**Exercise** *Jordan Form*

a.) Check that $(N^{(k)})^{k-1} \neq 0$ but $(N^{(k)})^k = 0$.

b.) Given the existence of Jordan canonical form show that if all the roots of the characteristic polynomial $p_A(x)$ are distinct then $A$ is diagonalizable. Give a counterexample to the converse statement.

c.) Show that the conjugacy class of a matrix $A \in M_2(\mathbb{C})$ is <u>not</u> uniquely determined by the values of $\text{Tr}(A^k)$, but it is if $A$ is diagonalizable. ($k = 1, 2$ will suffice.)

---

[56] *Answer*: By the spectral theorem $u$ is conjugate to

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}$$

Of course $\theta \sim \theta + 2\pi$, so we can parametrize the diagonal matrices uniquely by choosing $\theta \in [-\pi, \pi]$, with $\pm\pi$ identified. However conjugation by $i\sigma^1$ takes $\theta \to -\theta$. So, up to conjugacy we can take $\theta \in [0, \pi]$. But the value of $Tr(u) = 2\cos\theta$ determines a unique $\theta$ in the interval $[0, \pi]$.

**Exercise** *Nilpotent Cone*

Show that the most general traceless $2 \times 2$ complex matrix with nontrivial Jordan form is of the form

$$\begin{pmatrix} Z & X \\ Y & -Z \end{pmatrix} \tag{7.36}$$

with

$$XY + Z^2 = 0 \tag{7.37}$$

and $X \neq 0$ or $Y \neq 0$.

---

---

**Exercise** *Characterizing A Conjugacy Class By Traces*

a.) Show that the conjugacy class of a unitary $N \times N$ matrix is $u$ uniquely characterized by the set of traces $\mathrm{Tr}(u^j)$ for $1 \leq j \leq N$.

b.) Consider $A \in GL(N, \mathbb{C})$. Is the conjugacy class of $A$ uniquely determined by the set of traces $\mathrm{Tr}(A^j)$, $j \in \mathbb{Z}_+$ ?

♣There is some redundancy with previous exercise. And perhaps more should be explained about the relation between elementary symmetric functions and power symmetric functions. ♣

---

---

**Exercise** *The Complex Conjugate Representation*

a.) Consider the two 2-dimensional representations of $U(2)$ where $\varphi_1$ is the identity and $\varphi_2(u) = u^*$. Are these representations equivalent or inequivalent? [57]

b.) Consider the two 2-dimensional representations of $SU(2)$ where $\varphi_1$ is the identity and $\varphi_2(u) = u^*$. Are these representations equivalent or inequivalent? [58]

c.) Now take $N > 2$ and consider the $N$ dimensional representation of $SU(N)$ given by $\varphi_1(u) = u$ and $\varphi_2(u) = u^*$. Are these representations equivalent or inequivalent? [59]

---

---

**Exercise** *An Example Of Inequivalent Representations*

[57] *Answer*: They are inequivalent. By the spectral theorem $u \in U(2)$ can be conjugated to a diagonal matrix $Diag\{z_1, z_2\}$ with $z_1, z_2 \in U(1)$. The character of $\varphi_1$ is $\chi_1(u) = z_1 + z_2$. The character of $\varphi_2$ is $\chi_2(u) = z_1^{-1} + z_2^{-1}$. For general elements of $U(2)$ these are different so the character functions are different.

[58] *Answer*: Specializing from the previous characters to $SU(2)$ we must take $z_1 = z_2^{-1}$. So the characters of $\varphi_1$ and $\varphi_2$ are now the same. A priori, more work is needed to see if the representations are actually equivalent. In fact they are: Conjugation by $\mathrm{i}\sigma^2$ is equivalent to complex conjugation in $SU(2)$:

$$(\mathrm{i}\sigma^2) u (\mathrm{i}\sigma^2)^{-1} = u^* \ . \tag{7.38}$$

[59] They are inequivalent, as is easily seen by computing the character of diagonal $SU(N)$ matrices.

(To do this exercise you need to understand a little bit about tensor products. See Chapter 2, section 5.3.)

Consider the following four-dimensional representations of $SU(2)$:

$$\varphi_1(u) = \begin{pmatrix} u & 0 \\ 0 & u \end{pmatrix} \tag{7.39}$$

$$\varphi_2(u) = u \otimes u \tag{7.40}$$

Are these representations equivalent or inequivalent? [60]

---

---

**Exercise** *Characters Of A Permutation Representation*

Consider the $n$-dimensional representation of $S_n$ given by $T(\sigma) : e_i \mapsto e_{\sigma(i)}$ where $e_i$ is the standard basis of $\mathbb{R}^n$. Show that the character of this representation is

$$\chi(\sigma) = N(\sigma) = |\{i : \sigma(i) = i\}| \tag{7.41}$$

Below we define the notion of fixed points of a group action. In this case $i$ is a fixed point of the permutation $\sigma$ if $\sigma(i) = i$. So the character is the number of fixed points of $\sigma$.

---

### 7.3 Normal Subgroups And Quotient Groups

Groups which are self-conjugate are very special:

**Definition 7.2.2**: A subgroup $N \subseteq G$ is called a *normal* subgroup, or an *invariant* subgroup if

$$gNg^{-1} = N \qquad \forall g \in G \tag{7.42}$$

Sometimes this is denoted as $N \triangleleft G$.

**Warning!** Equation (7.42) does *not* mean that $gng^{-1} = n$ for all $n \in N$ !

There is a beautiful theorem associated with normal subgroups. In general the set of cosets of a subgroup $H$ in $G$, denoted $G/H$, does not have any natural group structure. [61] However, if $H$ is normal something special happens:

[60] *Answer*: They are inequivalent. One can show this by computing characters of the two representations. If $u$ is in the conjugacy class with $Tr(u) = 2\cos\theta$ then $\chi_1(u) = 4\cos\theta$ while $\chi_2(u) = 4(\cos\theta)^2$.

[61] Note that it might have many unnatural group structures. For example, if $G/H$ is a finite set with $n$ elements then we could choose a group $\tilde{G}$ with exactly $n$ elements. We can always do this, because for every positive integer $n$ there exists a finite group with $n$ elements. Then, having chosen a $\tilde{G}$ we could choose some one-one correspondence between elements of $G/H$ and elements of $\tilde{G}$ to define a group law on $G/H$. We hope the reader can appreciate how incredibly tasteless such a procedure would be. Technically, it is *unnatural* because it makes use of an arbitrary choice of group $\tilde{G}$, since in general there are many groups with a given order, and then it relies on the further extraneous choice of one-one correspondence between the elements of $G/H$ and the elements of $\tilde{G}$.

**Theorem 7.2.1**. If $N \subset G$ is a normal subgroup then the set of left cosets $G/N = \{gN | g \in G\}$ has a natural group structure with group multiplication defined by:

$$(g_1 N) \cdot (g_2 N) := (g_1 \cdot g_2) N \tag{7.43}$$

*Proof- left as an important exercise - see below.*

**Remarks**:

1. All subgroups $N$ of Abelian groups $A$ are normal, and moreover the quotient group $A/N$ is Abelian.
2. Groups of the form $G/N$ are known as *quotient groups*. A very common source of error and confusion is to mix up quotient groups and subgroups. They are very different!
3. As an illustration of the previous remark note that if $T : G \to GL(n, \kappa)$ is a matrix representation of $G$ and if $H \subset G$ is a subgroup then we can also restrict $T$ to $H$ to get an $n$-dimensional representation of $H$. However, if $Q$ is a quotient of $G$ it is not true in general that a representation of $G$ naturally determines a representation of $Q$. Using the data of $T$ the only natural definition of a representation $\tilde{T}$ on $Q$ would be $\tilde{T}(gN) := T(g)$. The problem with such a definition is that it might not be well-defined. Recall that $gN = g'N$ only implies that there is some $n \in N$ with $g' = gn$. But $T(g')$ and $T(g)$ will be different unless $T(n) = 1$. Therefore the above definition of $\tilde{T}$ only makes sense if $T(n) = 1$ for every $n \in N$.

**Example 7.2.1 Cyclic Groups** Since $\mathbb{Z}$ is Abelian $n\mathbb{Z} \subset \mathbb{Z}$ is normal, and the quotient group is $\mathbb{Z}/n\mathbb{Z}$. This is isomorphic to the cyclic group we have previously denoted as $\mu_n$ or $\mathbb{Z}_n$. So $\bar{r}$ is the equivalence class of an integer $r \in \mathbb{Z}$:

$$\bar{r} = r + n\mathbb{Z} \tag{7.44}$$

$$\bar{r} + \bar{s} = (r + s) + n\mathbb{Z} \tag{7.45}$$

**Example 7.2.2 Quotients of $\mathbb{Z}^d$**. Let's try to find a higher-dimensional generalization of cyclic groups. So we replace $G = \mathbb{Z}$ by $G = \mathbb{Z}^d$ for $d > 1$. For what follows it might help to think of $G$ as a subgroup of $\mathbb{R}^d$, but this is not strictly speaking necessary. let $e_i$ be a standard basis, with 1 in the $i^{th}$ row and zeroes elsewhere. How can we describe possible subgroups $H$? Let $A_{ij}$ be a $d \times d$ matrix of integers and consider the elements:

$$f_i := \sum_{j=1}^{d} A_{ij} e_j \tag{7.46}$$

Consider the subgroup $H \subset G$ of all integral linear combinations of $f_i$:

$$H := \{\sum_{i=1}^{d} n_i f_i | n_i \in \mathbb{Z}\} \tag{7.47}$$

$H$ is clearly a subgroup of $G$, so we can form the quotient group $G/H$. If $\det A \neq 0$ then in fact $G/H$ is a finite group. One way to see this easily is to consider $G$ as a subgroup of $\mathbb{Q}^d$, so that we can write

$$e_i = A^{-1}_{ij} f_j \tag{7.48}$$

with $A^{-1} \in GL(d,\mathbb{Q})$. Recall that $A^{-1} = (\det A)^{-1} Cof(A)$ where the cofactor matrix $Cof(A)$ is a matrix of minors, and therefore is a matrix of integers. Therefore $(\det A)e_i \in H$ and hence $\det A[e_i] = 0$ in the quotient group so every element of $G/H$ has a representative of the form $[\sum_i x_i e_i]$ with $|x_i| < |\det A|$.

Actually, if we invoke a nontrivial theorem we can say much more: The matrix $A$ can be put into *Smith normal form.* This means that there are matrices $S, T \in GL(d,\mathbb{Z})$, representing change of generators (i.e. change of basis of the $\mathbb{Z}$-module ) of $H$ and $G$ so that

$$SAT = Diag\{\alpha_1, \dots, \alpha_d\} \tag{7.49}$$

with $\alpha_i = d_i/d_{i-1}$ where $d_0 = 1$ and $d_i$ for $i > 0$ is the g.c.d. of the $i \times i$ minors. Then

$$G/H \cong \mathbb{Z}_{\alpha_1} \times \cdots \mathbb{Z}_{\alpha_d} \tag{7.50}$$

Note that it has order $|G/H| = \det A$. Here is a good example of the difference between a quotient group and a subgroup: No nontrivial finite group will be a subgroup of $\mathbb{Z}^d$.

Note that we set out to find higher-dimensional generalizations of cyclic groups, and we did not really find anything new, but in the process we encountered a very important technical fact: The existence of Smith Normal Form.

**Digression On Smith Normal Form**. We indicate the general idea behind one proof of Smith Normal Form.

First of all recall that row and column operations are invertible transformations: Let $O_{ij}(\alpha) = 1 + \alpha e_{ij}$. Then left-multiplication by $O_{ij}(\alpha)$ adds $\alpha$ times the $j^{th}$ row to the $i^{th}$ row and right-multiplication by $O_{ij}(\alpha)$ adds $\alpha$ times the $i^{th}$ column to the $j^{th}$ column. Note that it is important here that we only use $\alpha$'s which are integral.

Now consider a $2 \times 2$ matrix of integers:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{7.51}$$

Note that if $a = \pm 1$ or $d = \pm 1$ then by row and column operations we can easily transform the off diagonal terms to zero. Of course we can also permute rows and columns by conjugation by permutation matrices. What do we do if $|a| > 1$ and $|d| > 1$?

By multiplying by a diagonal matrix with $\pm 1$ on the diagonal we can make $a$ and $c$ both nonnegative. Using Bezout's theorem (see the Number Theory section **** below) there are relatively prime integers $x, y$ so that

$$ax + cy = e_1 \tag{7.52}$$

where $e_1 = gcd(a,c)$. Note that $0 \le e_1 \le a$. Then there are relatively prime integers $u, v$ so that $xv - uy = 1$ so we can multiply

$$\begin{pmatrix} x & y \\ u & v \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} e_1 & b_1 \\ (au+cv) & d_1 \end{pmatrix} \tag{7.53}$$

But note that $au + cv$ is divisible by $e_1$ so by a row operation (which is left multiplication by an invertible integer matrix) we can bring the RHS to

$$\begin{pmatrix} e_1 & b_1 \\ 0 & d_2 \end{pmatrix} \tag{7.54}$$

Now we repeat this on the right: Let $e_2 = gcd(e_1, b_1)$ so there are integers $x_1 e_1 + y_1 b_1 = e_2$ with $x_1 u_1 - y_1 v_1 = 1$ and multiply

$$\begin{pmatrix} e_1 & b_1 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} x_1 & v_1 \\ y_1 & u_1 \end{pmatrix} = \begin{pmatrix} e_2 & (e_1 v_1 + b_1 u_1) \\ d_2 y_1 & d_3 \end{pmatrix} \tag{7.55}$$

But now a column operation brings this to the form

$$\begin{pmatrix} e_2 & 0 \\ c_2 & d_3 \end{pmatrix} \tag{7.56}$$

Now $e_2$ divides $e_1$ divides $a$. Again using diagonal matrices with $\pm 1$ we can make $e_2$ nonnegative. In this was we get an alternating series of upper and lower triangular matrices with

$$a \ge e_1 \ge e_2 \ge \cdots \ge 0 \tag{7.57}$$

A strictly decreasing sequence

$$a > e_1 > e_2 > \cdots > 0 \tag{7.58}$$

must eventually terminate. So at some point $e_{n+1} = e_n$. When this happens then either a row or column operation allows us to cancel the off-diagonal term in the upper (or lower) triangular matrix, leaving a diagonal matrix.

By a very similar procedure we can use elementary operations to convert an $N \times N$ matrix of integers to the form

$$\begin{pmatrix} a & 0_{1\times(N-1)} \\ 0_{(N-1)\times 1} & A' \end{pmatrix} \tag{7.59}$$

and then we can apply an inductive argument to reduce $A'$.

1. The statement is much more general: It applies to $k \times n$ matrices over an arbitrary principal ideal domain.

2. Moreover, there is an explicit algorithm for successively using row and column operations to reduce to SNF. See the Wikipedia article on Smith Normal Form.

3. And in fact computer algebra programs like Maple and GAP will automatically reduce matrices over the integers to SNF.

**Example 7.2.3 Discriminant Group**. Now consider an embedded lattice in $\mathbb{R}^d$ equipped with Euclidean inner product. This is the integral span of a collection $\{v_i\}$ of vectors. For simplicity we will assume it is full rank, that is, the $\{v_i\}$ form a basis for $\mathbb{R}^d$ over $\mathbb{R}$. We denote it by $\Lambda$, so

$$\Lambda := \{\sum_{i=1}^{d} n_i v_i | n_i \in \mathbb{Z}\} \subset \mathbb{R}^d \tag{7.60}$$

We define the *dual lattice* (closely related to the "reciprocal lattice" in solid state physics) as the set of vectors $w \in \mathbb{R}^d$ such that $w \cdot v \in \mathbb{Z}$ for all $v \in \Lambda$:

$$\Lambda^{\vee} := \{w \in \mathbb{R}^d | \forall v \in \Lambda \quad v \cdot w \in \mathbb{Z}\} \tag{7.61}$$

Now assume that $\Lambda$ is an *integral lattice*. This means that the matrix of inner products $G_{ij} = v_i \cdot v_j$ is a $d \times d$ matrix of integers. (Note it is symmetric and of nonzero determinant.) Then it follows that $\Lambda \subset \Lambda^{\vee}$ is a sublattice. The *discriminant group* of $\Lambda$ is the finite group

$$\mathcal{D} := \Lambda^{\vee}/\Lambda \tag{7.62}$$

Note that $\Lambda^{\vee}$ has a basis $f_i$ with $v_i = G_{ij} f_j$ so one can work out $\mathcal{D}$ as a product of cyclic groups using the Smith normal form of $G_{ij}$.

**Remark**: Discriminant groups play a central role in the theory of integral lattices, which in turn show up in many contexts in physics from condensed matter to Chern-Simons theory to string theory compactification. Related to this, they are important in the theory of theta functions associated to integral lattices. In the context of Abelian Chern-Simons theory with action $\sim \frac{1}{4\pi} \int K_{IJ} A_I dA_J$ the level $K_{IJ}$ is a symmetric matrix of integers and defines an integral lattice. The discriminant group is the fusion group of Abelian anyons.

**Example 7.2.4**. Let us now consider some nonabelian examples.

$$A_3 \equiv \{1, (123), (132)\} \subset S_3 \tag{7.63}$$

is normal. Note that conjugation by any transposition preserves

$$(ij) A_3 (ij)^{-1} = A_3 \tag{7.64}$$

although the conjugation does induce a nontrivial permutation of the set $A_3$. For example

$$(12)(123)(12)^{-1} = (132) \tag{7.65}$$

The group $S_3/A_3$ has order 2 and hence must be isomorphic to $\mathbb{Z}_2$.

**Example 7.2.5**. Of course, in any group $G$ the subgroup $\{1\}$ and $G$ itself are normal subgroups. These are the trivial normal subgroups. It can happen that these are the only normal subgroups of $G$:

**Definition** . A group with no nontrivial normal subgroups is called a *simple group*.

**Remarks**

1. Note that a nonabelian simple group cannot have a nontrivial center.

2. The term "simple group" is a bit of a misnomer: Some "simple groups" are pretty darn complicated. What it means is that there is no means of simplifying it using something called the Jordan-Holder decomposition - discussed below. Simple groups are extremely important in the structure theory of finite groups. One example of simple groups are the cyclic groups $\mathbb{Z}/p\mathbb{Z}$ for $p$ prime. Can you think of others?

3. *Sylow's theorems again.* Recall that Sylow's first theorem says that if $p^k$ divides $|G|$ then $G$ has a subgroup of order $p^k$. If we take the largest prime power dividing $|G|$, that is, if $|G| = p^k m$ with $m$ relatively prime to $p$ then a subgroup of order $p^k$ is called a *$p$-Sylow subgroup*. Sylow's second theorem states that all the $p$-Sylow subgroups are conjugate. (If we do not take the maximal power of $p$ the statement is easily seen to be false. Just consider the product of cyclic groups with many factors of order $p^a$.) The third Sylow theorem says something about how many $p$-Sylow subgroups there are.

4. **WARNING!**: In the theory of Lie groups you will find the term "simple Lie group." A simple Lie group is NOT a simple group in the sense we defined above !! For example $SU(2)$ is a simple Lie group. But it has a nontrivial center namely the two diagonal $SU(2)$ matrices $\{\pm 1_{2\times 2}\}$.

**Example 7.2.6**. Recall that $SL(n,\kappa) \subset GL(n,\kappa)$, $SO(n,\kappa) \subset O(n,\kappa)$, and $SU(n) \subset U(n)$ are all subgroups defined by the condition $\det A = 1$ on a matrix. Note that, since $\det(gAg^{-1}) = \det A$ for any invertible matrix $g$ these are in fact normal subgroups. The quotient groups are

$$\begin{aligned} GL(n,\kappa)/SL(n,\kappa) &\cong \kappa^* \\ O(n,\mathbb{R})/SO(n,\mathbb{R}) &\cong \mathbb{Z}_2 \\ U(n)/SU(n) &\cong U(1) \end{aligned} \tag{7.66}$$

Lines 1 and 3 follow since every element in $GL(n,\kappa)$ can be written as $zA$ with $z \in \kappa^*$ and $A \in SL(n,\kappa)$. For line 2 take $P$ to be any reflection in any hyperplane orthogonal to some vector $v$, then $O(n,\mathbb{R}) = SO(n,\mathbb{R}) \amalg PSO(n,\mathbb{R})$ because $\det P = -1$. Recall that $P_{v_1}P_{v_2}$ is a rotation in the plane spanned by $v_1, v_2$, so it doesn't matter which hyperplane we choose.

♣Material reordered so this proof is in the future... ♣

**Example 7.2.7**. In contrast to the previous example we can consider what happens when we quotient by the center $G/Z(G)$. This always makes sense since $Z(G)$ is always a normal subgroup. As an interesting special case, the center of $U(N)$ consists of matrices proportional to the unit matrix. See the exercise below. Elements in the center of $SU(N)$ must also be diagonal. However, now if $z1_{N\times N}$ is to be in $SU(N)$ then $z^N = 1$ (why?) so $Z(SU(N)) \cong \mu_N \cong \mathbb{Z}_N$. Since this subgroup is normal we can take a quotient and get another group. It is known as

$$PSU(N) := SU(N)/\mathbb{Z}_N \tag{7.67}$$

One can show that $PSU(N) \cong U(N)/Z(U(N)) \cong U(N)/U(1)$. These groups illustrate well the distinction between quotient group and subgroup: There are representations of $SU(N)$ that are not representations of $PSU(N)$ so here is another example where $PSU(N)$ cannot be considered as a subgroup of $SU(N)$ in any sense.

**Example 7.2.8**. Let $G$ be a topological group. Let $G_0$ be the (path-) connected component of the identity element $1_G \in G$. It is not difficult to show that $G_0$ is a subgroup of $G$. (Exercise below.) We claim that $G_0$ is in fact a normal subgroup: If $g_0 \in G_0$ there is a continuous path of group elements $\gamma : [0,1] \to G$ with $\gamma(0) = 1_G$ and $\gamma(1) = g_0$. Then if $g \in G$ is any other group element $g\gamma(t)g^{-1}$ is a continuous path connecting $1_G$ to $gg_0g^{-1}$. The quotient group $G/G_0$ is the *group of components*, sometimes denoted $\pi_0(G)$, because, as a set, it is in 1-1 correspondence with the set of connected components of $G$. In general for a topological space $X$, the set of connected components is denoted by $\pi_0(X)$, but, for general topological spaces $X$, the set $\pi_0(X)$ carries no natural group structure (unlike the higher homotopy groups).

For some examples:

1. $G = \mathbb{R}^*$ under multiplication. The identity element is 1 and clearly $\mathbb{R}^*_{>0}$ is the connected component of the identity. The quotient group is isomorphic to $\mathbb{Z}_2$.

2. The determinant map defines a homomorphism $\det : GL(n,\mathbb{R}) \to \mathbb{R}^*$. Clearly there is no path of $GL(n,\mathbb{R})$ matrices that connects elements with $\det A < 0$ to $\det A > 0$. (Why?) In fact, the subgroup of $GL(n,\mathbb{R})$ of matrices with positive determinant is connected: The sign of the determinant is the only obstruction to deformation to the identity. [62]

3. Very similar considerations hold for $O(n,\mathbb{R})$. One can show that $SO(n,\mathbb{R})$ is the connected component of the identity and $\pi_0 \cong \mathbb{Z}_2$. For example, consider $O(2)$. In an exercise above you showed that as a manifold it has two components, each of which

[62] The theory of fiber bundles shows that if $H$ is a Lie subgroup of $G$ so that $G/H$ is a topological space then if $H$ is connected the set of components of $G$ can be identified with the set of components of $G/H$. We will see later from the stabilizer orbit theorem that $SO(n+1)/SO(n) = S^n$, so we can prove $SO(n)$ is connected by induction on $n$. By Gram-Schmidt procedure $SL(n,\mathbb{R})$ is a product of $SO(n)$ and upper triangular matrices with unit diagonal - and this space is connected. So $SL(n,\mathbb{R})$ is connected. Then the set of components of $GL(n,\mathbb{R})$ is that of $\mathbb{R}^*$, and this is measured by the determinant.

can be identified with a circle. The connected component of the identity is $SO(2)$. We have $O(2) = SO(2) \amalg SO(2)P$ where $P$ is any $O(2)$ matrix of determinant $= -1$. So $\pi_0(O(2)) \cong \mathbb{Z}_2$. Similarly, $\pi_0(O(n)) \cong \mathbb{Z}_2$.

4. One can show that $\pi_0(Diff(T^2)) \cong GL(2,\mathbb{Z})$. Indeed there is a subgroup of $Diff(T^2)$ isomorphic to $GL(2,\mathbb{Z})$ of diffeomorphisms

$$\begin{pmatrix} \sigma^1 \\ \sigma^2 \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \sigma^1 \\ \sigma^2 \end{pmatrix} \tag{7.68}$$

which projects isomorphically to the quotient group.

5. If $G$ is a finite group then $\pi_0(G) \cong G$.

**Example 7.2.9**. In nonabelian gauge theory on a noncompact manifold the group of all gauge transformations is sometimes defined by the function group $\mathcal{G} = Map[X \to G]$ so that $g : X \to G$ goes to a constant "at infinity." The group of local gauge transformations $\mathcal{G}_0$ is the subgroup so that $g(x) \to 1$. We have $\mathcal{G}_0 \lhd \mathcal{G}$. The quotient group is then to be thought of as the group of "global gauge transformations." In fact, $\mathcal{G}/\mathcal{G}_0 \cong G$.

---

**Exercise** *Due Diligence*
a.) Check the details of the proof of Theorem 7.2.1 ! [63]
b.) Consider the *right cosets*. Show that $N\backslash G$ is a group.

---

---

**Exercise** *A Cancellation Theorem*
Suppose that we have a chain of normal subgroups

$$K \lhd N \lhd G \tag{7.69}$$

and moreover $K \lhd G$.
a.) Show that [64]

$$N/K \lhd G/K \tag{7.70}$$

[63] *Answer*: The main thing to check is that the product law defined by (7.43) is actually well defined. Namely, you must check that if $g_1N = g_1'N$ and $g_2N = g_2'N$ then $g_1g_2N = g_1'g_2'N$. To show this note that $g_1' = g_1n_1$ and $g_2' = g_2n_2$ for some $n_1, n_2 \in N$. Now note that $g_1'g_2' = g_1n_1g_2n_2 = g_1g_2(g_2^{-1}n_1g_2)n_2$. But, since $N$ is normal $(g_2^{-1}n_1g_2) \in N$ and hence $(g_2^{-1}n_1g_2)n_2 \in N$ and hence indeed $g_1g_2N = g_1'g_2'N$. Once we see that (7.43) is well-defined the remaining checks are straightforward. Essentially all the basic axioms are inherited from the group law for multiplying $g_1$ and $g_2$. Associativity should be obvious. The identity is $1_GN = N$ and the inverse of $gN$ is $g^{-1}N$. etc. ♠

[64] *Answer*: $(gK)(nK)(gK)^{-1} = (gng^{-1})K$ but $gng^{-1} = n'$ for some $n' \in N$ since $N \lhd G$, but then $n'K \in N/K$.

b.) Show that there is an isomorphism of groups [65]

$$G/N \cong (G/K)/(N/K) \tag{7.71}$$

---

---

**Exercise** *Even Permutations*

Example 7.2.2 has a nice generalization. Recall that a permutation is called *even* if it can be written as a product of an even number of transpositions.

a.) Show that the even permutations, $A_n$, form a normal subgoup of $S_n$.

b.) What is $S_n/A_n$?

---

---

**Exercise** *Subgroups Of Index Two*

a.) Suppose that $H \subset G$ is of index two: $[G : H] = 2$. Show that $H$ is normal in $G$. What is the group $G/H$ in this case? [66]

b.) Using (a) give another proof that $A_n \triangleleft S_n$ is a normal subgroup.

c.) As we will discuss later, the groups $A_n$ for $n \geq 5$ are simple groups. Accepting this for the moment give an infinite set of counterexamples to the converse of Lagrange's theorem. [67]

---

---

**Exercise**

Look at the 3 examples of homogeneous spaces $G/H$ in section 7.1. Decide which of the subgroups $H$ is normal and what the group $G/H$ would be.

---

[65] *Answer*: Define $\psi : (G/K)/(N/K) \to G/N$ by $\psi(gK \cdot (N/K)) \to gN$. The main thing we need to check is that this is well defined. But if $gK \cdot (N/K) = g'K \cdot (N/K)$ then $g'K = (gK)(nK) = (gn)K$ so $g' = gnk$ for some $n, K$ but $nk \in N$ so $g'N = gN$. Now compute the kernel of $\psi$: If $gN = N$ then $g \in N$ but then $(gK) \cdot (N/K) = N/K$.

[66] *Answer*: Suppose $G = H \amalg g_0 H$. Then take any $h \in H$. The element $g_0 h g_0^{-1}$ must be in $H$ or $g_0 H$. But if it were in $g_0 H$ then there would be an $h' \in H$ such that $g_0 h g_0^{-1} = g_0 h'$ but this would imply $g_0$ is in $H$, which is false. Therefore, for all $h \in H$, $g_0 h g_0^{-1} \in H$, and hence $H$ is a normal subgroup. Therefore $G/H \cong \mathbb{Z}_2$.

[67] *Answer*: Note that the order of $|A_n|$ is even and hence $\frac{1}{2}|A_n|$ is a divisor of $|A_n|$. However, a subgroup of order $|A_n|/2$ would have to be a normal subgroup, and hence does not exist, since $A_n$ is simple. More generally, a high-powered theorem, known as the Feit-Thompson theorem states that a finite simple non-abelian group has even order. Therefore if $G$ is a finite simple nonabelian group there is no subgroup of order $\frac{1}{2}|G|$, even though this is a divisor.

---

**Exercise**
Show that if the center $Z(G)$ is such that $G/Z(G)$ is cyclic then $G$ is Abelian. [68]

---

---

**Exercise** *Sylow subgroups of $A_4$*
Write down the 2-Sylow and 3-Sylow subgroups of $A_4$.

---

---

**Exercise** *Commutator Subgroups And Abelianization*

If $g_1, g_2$ are elements of a group $G$ then the *group commutator* is defined to be the element

$$[g_1, g_2] := g_1 g_2 g_1^{-1} g_2^{-1} \ . \tag{7.72}$$

If $G$ is any group the *commutator subgroup* usually denoted $[G,G]$ (sometimes denoted $G'$) is the smallest subgroup of $G$ containing all the group commutators.

a.) Show that $[G,G]$ is a normal subgroup of $G$. [69]

b.) Show that $G/[G,G]$ is abelian. This is called the *abelianization* of $G$. [70]

---

---

**Exercise** *A Less Than Perfect Group*

a.) Recall that a *simple* group is a group with no nontrivial normal subgroups. A *perfect* group is a group which is equal to its commutator subgroup. Show that a nonabelian simple group must be perfect.

b.) Show that $S_n$ is not a perfect group. What is the commutator subgroup? [71]

---

---

**Exercise** *Signed Permutations Again*

Recall our discussion of a natural matrix representation of $S_n$ and the group $W(B_n)$ of signed permutations from *** above.

[68] *Answer*: Every element of $G$ would be of the form $g_0^n z$ with $z \in Z$. But then it is easy to check: $g_0^n z g_0^m z' = g_0^m z' g_0^n z$ so $G$ is Abelian. So $G = Z(G)$, and in fact the cyclic subgoup must be trivial.

[69] *Answer*: Note that $g_0[g_1, g_2]g_0^{-1} = [g_1', g_2']$ where $g_i' = g_0 g_i g_0^{-1}$.

[70] *Answer*: Let $G' = [G,G]$ then $g_1 G' g_2 G' = (g_1 g_2)G' = g_2 g_1 (g_1^{-1} g_2^{-1} g_1 g_2)G' = g_2 g_1 G'$.

[71] The commutator subgroup is clearly a subgroup of $A_n$. In fact $A_n$ is generated by products of two transpositions and hence is generated by $(abc)$. But note that $(ab)(ac)(ab)(ac) = (abc)$. Therefore the commutator subgroup of $S_n$ is just $A_n$.

a.) Show that the subgroup of $W(B_n)$ of diagonal matrices is a normal subgroup isomorphic to $\mathbb{Z}_2^n$.

b.) Show that every signed permutation matrix can be written in the form $D \cdot \Pi$ where $D$ is a diagonal matrix of $\pm 1$'s and $\Pi$ is a permutation matrix.

c.) Conclude that the quotient of $W(B_n)$ by the normal subgroup of diagonal matrices is isomorphic to $S_n$.

d.) Show that every signed permutation can also be written as $\Pi' \cdot D'$. How is this decomposition related to writing it as $D \cdot \Pi$.

---

---

**Exercise** *Products Of Simple Groups*

Let $G_1$ and $G_2$ be simple groups.

a.) What are the normal subgroups of the Cartesian product $G_1 \times G_2$? [72]

b.) Suppose $G_i$, $i \in I$ is a set of simple groups. What are the subgroups of $\prod_{i \in I} G_i$?

---

♣This exercise should go to the first section when we introduce the matrix groups. ♣

---

**Exercise** *The Center Of* $U(N)$

Show that the center of $U(N)$ consists of the subgroup of matrices proportional to the unit matrix and is therefore isomorphic to $U(1)$. [73]

---

---

**Exercise** *Subgroups Which Are Not So Normal*

a.) Consider $O(n, \mathbb{R}) \subset GL(n, \mathbb{R})$. Is this a normal subgroup?

b.) Consider the subgroup of diagonal matrices in $SU(N)$. Is this a normal subgroup?

---

---

[72] *Answer*: Only $\{1\}$, $G_1 \times \{1_{G_2}\}$ , $\{1_{G_1}\} \times G_2$ and $G_1 \times G_2$.

[73] *Answer*: There are many proofs but one nice one is to use induction on $N$. First establish the result for $U(2)$ - here the matrix multiplication is easy and this can be done by hand. Now suppose that $\zeta \in U(N+1)$ is in the center. Decompose it as follows

$$\zeta = \begin{pmatrix} A & v \\ w & D \end{pmatrix}$$

where $A \in M_2(\mathbb{C})$, $v \in M_{2\times N-1}(\mathbb{C})$, $w \in M_{N-1\times 2}(\mathbb{C})$, and $D \in M_{N-1\times N-1}(\mathbb{C})$. Now insist that it commute with

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

with $u \in U(2)$ to show that $uAu^{-1} = A$, $uv = v$ and $wu^{-1} = w$ for all $u \in U(2)$. These equations imply $A$ is diagonal and $v, w = 0$.

**Exercise** *The Normalizer Subgroup*

If $H \subset G$ is a subgroup then we define the *normalizer of $H$ within $G$* to be the largest subgroup $N$ of $G$ such that $H$ is a normal subgroup of $N$. Note that $H$ is normal inside itself so such subgroups exist. If $N_1, N_2 \subset G$ are subgroups and $H$ is normal in both then they generate a subgroup in which $H$ is normal. So the maximal subgroup exists.

Define:

$$N_G(H) := \{g \in G | gHg^{-1} = H\} \tag{7.73}$$

a.) Show that (7.73) is a subgroup of $G$ and $H$ is a normal subgroup of $N_G(H)$.

b.) Show that $N_G(H)$ is the largest subgroup of $G$ which contains $H$ as a normal subgroup, so (7.73) is a formula for the normalizer subgroup of $H$ within $G$.

Note that there is no claim that $N_G(H)$ is a normal subgroup of $G$. In general, it is not.

---

---

**Exercise** *The Weyl Groups Of $SU(2)$*

a.) Let $D \subset SU(2)$ be the subgroup of diagonal matrices. Note that $D \cong U(1)$. Compute

$$N_{SU(2)}(D) \tag{7.74}$$

explicitly. [74]

b.) Compute the quotient group $N_{SU(2)}(D)/D$. [75]

c.) Show that conjugation by elements of the normalizer act by a permutation of the diagonal elements and the permutation only depends on the projection to the quotient. [76]

d.) Show that there is no subgroup of $N_{SU(2)}(D)$ which is isomorphic to $S_2 \cong \mathbb{Z}_2$ and whose conjugation action on $D$ induces the permutation action.

---

---

**Exercise** *Weyl Group Of $SU(N)$*

Every element of $SU(N)$ can be conjugated into the set $T$ Of diagonal matrices.

---

[74] *Answer*: The normalizer is the subgroup of $SU(2)$ that is the union of matrices of the form

$$\begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix}$$

or of the form

$$\begin{pmatrix} 0 & -z^{-1} \\ z & 0 \end{pmatrix}$$

where $z$ is a phase.

[75] *Answer* The quotient is isomorphic to $\mathbb{Z}_2$.

[76] The conjugation by the diagonal matrices in $N_{SU(2)}(D)$ on $D$ acts as the identity. The conjugation by the off-diagonal matrices in $N_{SU(2)}(D)$ on $D$ acts as permutation of the diagonal elements.

a.) Show that the normalizer $N(T)$ of $T$ within $SU(N)$ is larger than $T$ by considering $SU(N)$ matrices: (Here $i \neq j$):

$$U(i,j) := 1 - (e_{ii} + e_{jj}) + \mathrm{i}(e_{ij} + e_{ji}) \in SU(N) \tag{7.75}$$

The reader should carefully show that $U(i,j)$ is indeed in $SU(N)$.

b.) Show that the conjugation action by $U(i,j)$ on the subgroup of diagonal matrices permutes the $ii$ and $jj$ diagonal elements leaving all the others fixed.

c.) Show that the homomorphism $N(T) \to S_N$ must be surjective.

**Remark**: We remarked above that in a compact simple Lie group $G$ a maximal dimension torus subgroup will be unique - up to conjugation. Such a subgroup is known (with some abuse of language) "the" maximal torus. An example of a maximal torus for $SU(n)$ would be the diagonal subgroup. In general the Weyl group of $G$ is by definition

$$W(G) := N_G(T)/T \tag{7.76}$$

For example, in $SU(n)$ any maximal torus is conjugate to the subgroup $D \subset SU(n)$ of diagonal matrices. In this case, conjugation by $N_{SU(n)}(D)$ acts on $D$ by permutation of the diagonal elements and in fact

$$W(SU(n)) := N_{SU(n)}(D)/D \cong S_n \tag{7.77}$$

Note that the Weyl group is defined as a quotient of a subgroup of $G$. (Often this is abbreviated to "subquotient of $G$.") In general there is no subgroup of $G$ that is isomorphic to $W(G)$ and whose conjugation action induces the Weyl group action on $T$. (It is a common mistake to confuse $W(G)$ with a subgroup of $G$.)

---

---

**Exercise** *Representations Of $SU(N)$ That Are Not Representations Of $PSU(N)$*

Give an example of a representation of $SU(N)$ that is not a representation of the quotient $PSU(N)$. [77]

---

---

**Exercise** *Homomorphic Images And Normal Subgroups*

Suppose $N \triangleleft G$, and suppose that $\varphi : G \to \tilde{G}$ is a homomorphism to some other group $\tilde{G}$. Then $\varphi(N) \subset \tilde{G}$ is a subgroup. Is it a normal subgroup?

a.) If not, give a counterexample.

b.) Under what conditions is $\varphi(N)$ a normal subgroup? [78]

[77] The defining representation is not, because the center of $SU(N)$ acts nontrivially.

[78] *Answer*: In general it is not a normal subgroup. A simple counterexample is to take and inclusion homomorphism $\varphi : S_3 \to S_4$ and let $N = A_3$. However, if $\varphi$ is surjective then it is easy to see that it is a normal subgroup.

---

---

**Exercise** *The Connected Component Of The Identity Is A Subgroup*
Show that the path-connected component of the identity of a topological group is a subgroup. [79]

---

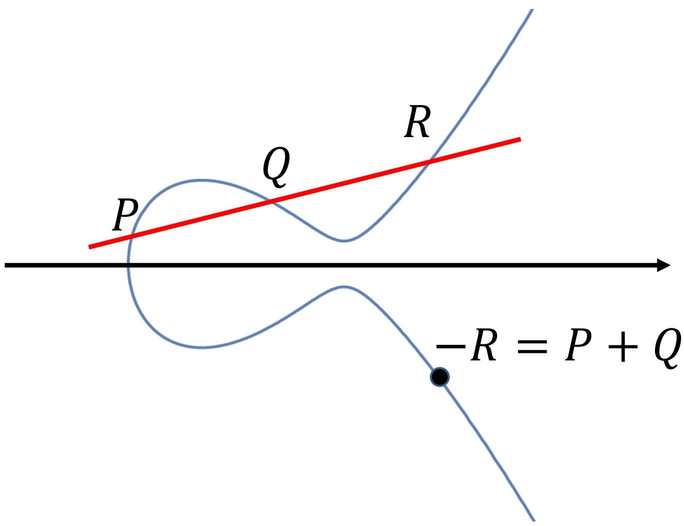



**Figure 3:** In a suitable range of real values of $f, g$ the real points on the elliptic curve have the above form. Then the elliptic curve group law is easily pictured as shown.

### 7.3.1 A Very Interesting Quotient Group: Elliptic Curves

Consider the Abelian group $\mathbb{C}$ of complex numbers with normal addition as the group operation. If $\tau$ is a complex number with nonzero imaginary part then $\mathbb{Z} + \tau\mathbb{Z}$ is the subgroup of complex numbers of the form $n_1 + \tau n_2$ where $n_1$ and $n_2$ are integers. Since $\mathbb{C}$ is Abelian we can form the Abelian group $\mathbb{C}/\mathbb{Z} + \tau\mathbb{Z}$. Note that $\Lambda := \mathbb{Z} + \tau\mathbb{Z}$ is a rank two lattice in the plane so that this quotient space can be thought of as a torus. As an Abelian group this group is isomorphic to $U(1) \times U(1)$. The explicit isomorphism is

$$(\sigma_1 + \tau\sigma_2) + \Lambda \mapsto (e^{2\pi i \sigma_1}, e^{2\pi i \sigma_2}) \tag{7.79}$$

Note that the identity element is $[0] = \Lambda$ and the inverse of $[z]$ is $[-z]$.

---

[79] *Answer*: Suppose there are continuous paths $\gamma_i : [0,1] \to G$ with $\gamma_i(0) = 1_G$ and $\gamma_i(1) = g_i$. Consider the path

$$\tilde{\gamma}(t) := \begin{cases} \gamma_1(2t) & 0 \le t \le 1/2 \\ g_1 \cdot \gamma_2(2t-1) & 1/2 \le t \le 1 \end{cases} \tag{7.78}$$

A remarkable fact is that this torus (minus one point) can be thought of as the space of solutions of the algebraic equation

$$y^2 = x^3 + fx + g \tag{7.80}$$

where $(x, y) \in \mathbb{C}^2$ and $f, g \in \mathbb{C}$. [80]

The mapping between $[z] \in \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ and $(x, y)$ and between $f, g$ and the complex number $\tau$ involves very interesting functions known as elliptic and modular functions. The solution set to (7.80) in $\mathbb{C}^2$ is known as an "elliptic curve." It is not difficult to describe the mapping. One introduces a holomorphic function of $z$ known as the *Weierstrass function*:

$$\wp(z|\tau) := \frac{1}{z^2} + \sum_{\omega \in \Lambda - \{0\}} \left( \frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right) \tag{7.81}$$

where $\omega = n_1 + n_2\tau \in \Lambda := \mathbb{Z} + \tau\mathbb{Z}$ and $z \notin \Lambda$ Note that for large values of $\omega$ the summand behaves like $\frac{2z}{\omega^3}$, so the series converges absolutely since $\int \frac{dxdy}{r^3}$ is convergent at $r \to \infty$. By general results in complex analysis the function is holomorphic for $z \in \mathbb{C} - \Lambda$. Note that it is also doubly-periodic:

$$\wp(z + m + m'\tau|\tau) = \wp(z|\tau) \tag{7.82}$$

for all $m, m' \in \mathbb{Z}$. So it descends to a function on the quotient to define a meromorphic function on a complex manifold.

For $[z] = 0$ the function has a second order pole. Put differently, the Weierstrass function has a double pole at every point $z \in \Lambda$. Indeed we can expand $\wp(z|\tau)$ around $z = 0$:

$$\begin{aligned} \wp(z|\tau) &= \frac{1}{z^2} + \sum_{k=1}^{\infty} (2k+1) G_{2k+2} z^{2k} \\ &= \frac{1}{z^2} + 3G_4 z^2 + 5G_6 z^4 + \cdots \end{aligned} \tag{7.83}$$

where

$$G_{2k+2} = \sum_{(n_1,n_2) \in \mathbb{Z}^2 - \{(0,0)\}} \frac{1}{(n_1 + n_2\tau)^{2k+2}} \tag{7.84}$$

are absolutely convergent and hence holomorphic functions of $\tau$ for $k \geq 1$ when $\mathrm{Im}\tau \neq 0$. They are famous functions known as *Eisenstein functions* and are basic examples of a fascinating set of functions known as *modular forms*. In order to produce equation (7.80) we will take $x = \wp(z|\tau)$ and define

$$\begin{aligned} y &:= \frac{\partial}{\partial z} \wp(z|\tau) \\ &= \frac{-2}{z^3} + 6G_4 z + 20G_6 z^3 + \cdots \end{aligned} \tag{7.85}$$

[80] We can restore the point at infinity using projective geometry. The equation $ZY^2 = X^3 + fXZ^2 + gZ^3$ makes sense for a point $[X : Y : Z] \in \mathbb{CP}^2$. Indeed note that the equivalence relation says that $[X : Y : Z] = [\lambda X : \lambda Y : \lambda Z]$ and the equation is homogeneous and of degree three. The equation (7.80) is the equation we get in the patch $Z \neq 0$ where we can fix the scaling degree of freedom by choosing $\lambda$ so that $Z = 1$. We then define $x, y$ by $[x : y : 1] = [X : Y : Z] = [X/Z : Y/Z : 1]$ which makes sense when $Z \neq 0$. The point at infinity has $Z = 0$. Therefore, by the equation $X = 0$, and since we have a point in $\mathbb{CP}^2$ we must have $Y \neq 0$, which can therefore be scaled to $y = 1$. So, the point at infinity is $[0 : 1 : 0] \in \mathbb{CP}^2$.

Now a small amount of algebra shows that we have the series expansion

$$y^2 - 4x^3 + 60G_4 x = -140G_6 + \mathcal{O}(z^2) \tag{7.86}$$

The expression on the left is double-periodic. But the pole at $z = 0$ has cancelled. So the pole at all values of $\omega \in \Lambda$ have cancelled! So the combination $y^2 - 4x^3 + 60G_4 x$ is entire. Moreover, because it is doubly-periodic all the values are taken on the closed rectangle given by a fundamental domain. So it is a bounded entire function. Therefore, by Liouville's theorem it must be constant! Thus all the higher terms in the series vanish! Thus we have the exact formula:

$$y^2 = 4x^3 - 60G_4 x - 140G_6 \tag{7.87}$$

which can be transformed to the expression (7.80) by rescaling

$$y = 2\gamma^3 \tilde{y} \qquad\qquad x = \gamma^2 \tilde{x} \tag{7.88}$$

for any nonzero $\gamma \in \mathbb{C}^*$.

The very simple Abelian group law

$$[z_1] + [z_2] = [z_1 + z_2] \tag{7.89}$$

expressed in terms of $(x, y)$ is rather nontrivial and closely related to some deep topics in number theory. We will describe it below, but some simple aspects are easily motivated by the isomorphism to $\mathbb{C}/\Lambda$. First of all, the identity element must correspond to the point at infinity, $(x = \infty, y = \infty)$ which can be given precise meaning in projective space as the point $[0 : 1 : 0]$. Second, because $\wp(z|\tau)$ is even in $z$ the inversion in the Abelian group must be $I(x, y) = (x, -y)$.

Now, if one considers $f, g$ to be real and studies the real solutions then the group law can be visualized as in Figure 3. (We are following the Wikipedia article here, which is quite clear.) One first defines the inverse $-P$ of a point $P$ on the curve with coordinates $P = (x, y)$ to be $-P := (x, -y)$. As we have seen this is compatible with $I([z]) = [-z]$. Then, for generic points $P$ and $Q$ we can define $P + Q$ by saying that

$$P + Q + R = 0 \tag{7.90}$$

if they are three collinear points on the elliptic curve. Since we know how to invert that means we can take the definition of $P + Q$ to be $-R$. Note that with these rules the statement $P + (-P) = 0$ shows that 0 must correspond to the point at infinity.

We can express the group law on $(x_P, y_P) + (x_Q, y_Q) = (x_R, -y_R)$ in explicit formulae as follows: We write the line between $P, Q$ as

$$y = sx + d \tag{7.91}$$

with

$$s = \frac{y_P - y_Q}{x_P - x_Q} \qquad d = y_P - x_P\left(\frac{y_P - y_Q}{x_P - x_Q}\right) = y_Q - x_Q\left(\frac{y_P - y_Q}{x_P - x_Q}\right) \tag{7.92}$$

Now the intersection of this line with the cubic equation has $x$ coordinates given by

$$(sx + d)^2 = x^3 + fx + g \tag{7.93}$$

and by simple rearrangement we can rewrite (7.93) as

$$x^3 - s^2x^2 + (f - 2sd)x + (g - d^2) = 0 \tag{7.94}$$

On the other hand, this equation must be of the form

$$(x - x_P)(x - x_Q)(x - x_R) = 0 \tag{7.95}$$

Expanding out (7.94) and equating the coefficient of $x^2$ we obtain

$$x_R = s^2 - x_P - x_Q \tag{7.96}$$

so we have $x_R$ explicitly as a function of $x_P, x_Q, y_P, y_Q$. Now the point $(x_R, y_R)$ must lie on the line $y = sx + d$ so we can also say that

$$y_R = y_P + s(x_R - x_P) \tag{7.97}$$

expressing $y_R$ and hence the coordinates of $R = (x_R, -y_R)$ as rational functions of $x_P, x_Q, y_P, y_Q$. It is not at all obvious that the above group law really satisfies the associativity constraint. The law also corresponds to a nontrivial identity on Weierstrass functions.

The above formulae hold for generic points. When points coincide or the line is tangent to the elliptic curve one must carefully degenerate the above expressions.

When $f, g$ are real but not in the range to give a figure like Figure 3 the algebraic equations above still define a group law. Indeed, these equations make sense over <u>any</u> field, such as $\mathbb{Q}$, or over finite fields, even though we do not have the corresponding torus $\mathbb{C}/\Lambda$ as an isomorphic model. These generalizations are of great importance in number theory and cryptography.

---

**Exercise** *Modular Transformations*

a.) Show that the transformation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : (z, \tau) \mapsto \left( \frac{z}{c\tau + d}, \frac{a\tau + b}{c\tau + d} \right) \tag{7.98}$$

defines a group action of $SL(2, \mathbb{Z})$ on pairs $\mathbb{C} \times \mathbb{H}$.

b.) Show that

$$\wp\left( \frac{z}{c\tau + d} \middle| \frac{a\tau + b}{c\tau + d} \right) = (c\tau + d)^2 \wp(z|\tau) \tag{7.99}$$

c.) Show that

$$G_{2k}(\frac{a\tau + b}{c\tau + d}) = (c\tau + d)^{2k} G_{2k}(\tau) \qquad k = 2, 3, 4, ... \tag{7.100}$$

d.) Using Fourier analysis prove that for $z \notin \mathbb{Z}$ we have

$$\sum_{n\in\mathbb{Z}} \frac{1}{(z+n)^2} = \frac{\pi^2}{\sin^2(\pi z)} \tag{7.101}$$

and conclude that for $\mathrm{Im} z > 0$

$$\sum_{n\in\mathbb{Z}} \frac{1}{(z+n)^{2k}} = \frac{(2\pi \mathrm{i})^{2k}}{(2k-1)!} \sum_{\ell=1}^{\infty} \ell^{2k-1} e^{2\pi \mathrm{i} \ell z} \tag{7.102}$$

From this derive that

$$G_{2k}(\tau) = 2\zeta(2k) + 2\frac{(2\pi \mathrm{i})^{2k}}{(2k-1)!} \sum_{n=1}^{\infty} \sigma_{2k-1}(n) q^n \tag{7.103}$$

where $\sigma_m(n) = \sum_{d \text{ divides } n} d^m$ and $q := e^{2\pi \mathrm{i}\tau}$. Furthermore show that

$$G_{2k}(\tau) = 2\zeta(2k) E_{2k}(\tau) \tag{7.104}$$

$$E_{2k}(\tau) = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} \frac{n^{2k-1} q^n}{1-q^n} \tag{7.105}$$

Here $B_{2k}$ are the Bernoulli numbers defined by the series expansion around $t = 0$:

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} \frac{B_n}{n!} t^n = 1 - \frac{1}{2} t + \sum_{k=1}^{\infty} \frac{B_{2k}}{(2k)!} t^{2k} \tag{7.106}$$

The functional equation for the Riemann zeta function leads to the identity:

$$B_{2k} = (-1)^{k+1} \frac{2(2k)!}{(2\pi)^{2k}} \zeta(2k) \qquad k \geq 1 \tag{7.107}$$

**Remark**: Holomorphic functions on the upper half-plane that transform like

$$f\left(\frac{z}{c\tau + d}, \frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^w f(\tau) \tag{7.108}$$

for an integer $w$ are called "modular forms of weight $w$." One can show that if one imposes a boundedness condition (essentially that there are no negative Fourier modes in the Fourier expansion) then modular forms only exist for even weights $> 2$. Defining $M_{2k}$ to be the vector space of such forms of weight $2k$ one key theorem in the subject says that

$$\oplus_{k>1} M_{2k} = \mathbb{C}[E_4, E_6] \tag{7.109}$$

where the RHS is the polynomial ring generated by $E_4$ and $E_6$. There are many many variations on the above ideas.

---

---

**Exercise** *Elliptic Integrals*

Show that if $z$ is in a small neighborhood of $z_0$ within a fundamnetal domain for the $\mathbb{Z}+\tau\mathbb{Z}$ action on $\mathbb{C}$ and $t, t_0$ are in a small neighborhood that does not "surround" any of the roots of $4x^3 - 60G_4 x - 140G_6$ then

$$z - z_0 = \int_{x_0}^{x} \frac{dt}{\sqrt{4t^3 - 60G_4(\tau)t - 140G_6(\tau)}} \tag{7.110}$$

where

$$x = \wp(z|\tau) \qquad x_0 = \wp(z_0|\tau) \tag{7.111}$$

Evidentally, if we consider the integral (7.110) and analytically continue in $t$ then it will be double valued and will depend on the choice of contour connecting $x_0$ to $x$.

**Remark**: In general an integral with a rational expression involving a square root of a cubic or quartic polynomial will be expressed in terms of elliptic functions. By contrast, integrals with square roots of quadratic functions can be expressed in terms of trigonometric functions:

$$\int^{x} \frac{dt}{\sqrt{at^2 + bt + c}} = \frac{1}{\sqrt{-a}} \cos^{-1}\left(-\frac{b + 2ax}{\sqrt{D}}\right) \tag{7.112}$$

where $D = b^2 - 4ac$ is the discriminant. Note that analytic continuation in $x$ will lead to a multiple valued function, in harmony with the multiple possible values taken by the inverse cosine. Elliptic integrals generalize this expression by replacing the quadratic polynomial $at^2 + bt + c$ by cubic and quartic polynomials. Elliptic functions show up in many exact results in physics. Perhaps the simplest example is the exact formula for the period of a pendulum of length $\ell$ with maximal angle $\theta_{mx}$ from the vertical:

$$t_2 - t_1 = \sqrt{\frac{\ell}{2g}} \int_{\theta_1}^{\theta_2} \frac{d\theta}{\sqrt{\cos\theta - \cos\theta_{mx}}} \tag{7.113}$$

The integral can be converted to (set $\sin u = \frac{\sin\theta/2}{\sin\theta_{mx}/2}$ and $k = \sin\theta_{mx}/2$):

$$t_2 - t_1 = \frac{1}{2}\sqrt{\frac{\ell}{g}} \int_{u_1}^{u_2} \frac{du}{\sqrt{1 - k^2 \sin^2 u}} \tag{7.114}$$

which in turn is an integral with a square root of a quartic polynomial, via $s = \sin u$

$$t_2 - t_1 = \frac{1}{2}\sqrt{\frac{\ell}{g}} \int_{s_1}^{s_2} \frac{ds}{\sqrt{(1 - s^2)(1 - k^2 s^2)}} \tag{7.115}$$

---

♣Explain more about how elliptic integrals give exact solution to the motion of a pendulum ♣

## 7.4 Quotient Groups And Short Exact Sequences

In mathematics one often encounters the notion of an *exact sequence*: Suppose we have three groups and two homomorphisms $f_1, f_2$

$$G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3 \tag{7.116}$$

| |
|---|
| We say the sequence is *exact at* $G_2$ if $\text{im} f_1 = \ker f_2$. |

This generalizes to sequences of several groups and homomorphisms

$$\cdots G_{i-1} \xrightarrow{f_{i-1}} G_i \xrightarrow{f_i} G_{i+1} \xrightarrow{f_{i+1}} \cdots \tag{7.117}$$

The sequence can be as long as you like. It is said to be *exact at* $G_i$ if $\text{im}(f_{i-1}) = \ker(f_i)$.

A *short exact sequence* is a sequence of the form

$$1 \xrightarrow{f_0} G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3 \xrightarrow{f_3} 1 \tag{7.118}$$

which is exact at $G_1$, $G_2$, and $G_3$. Here 1 refers to the trivial group with one element. (When we work with Abelian groups and work additively we will denote the trivial group by 0.) Note that the homomorphism $f_0$ is unique: $f_0(e) = 1_{G_1}$ where $e$ is the identity (and only) element of the trivial group 1. Henceforth we will not write $f_0$ explicitly. Similarly the homomorphism $f_3$ is also unique: It takes every element of $G_3$ to the identity(and only) element of the trivial group 1. Henceforth we will not write $f_3$ explicitly.

Thus, the statement that (7.118) is a short exact sequence is the statement that the following is true:

1. Exactness at $G_1$: The kernel of $f_1$ is the image of $f_0$, but that image is plainly $1_{G_1}$. It is an easy exercise to show that if a homomorphism has a trivial kernel then it is an injection of $G_1$ into $G_2$.

2. Exactness at $G_2$: $\text{im} f_1 = \ker f_2$.

3. Exactness at $G_3$: The kernel of $f_3$ is all of $G_3$. Exactness at $G_3$ means that this kernel is the image of the homomorphism $f_2$, and hence $f_2$ is a surjective homomorphism, i.e. $f_2$ maps $G_2$ onto all of $G_3$.

In particular, note that if $\mu : G \to G'$ is any group homomorphism then we automatically have a short exact sequence:

$$1 \to K \to G \xrightarrow{\mu} \text{im}(\mu) \to 1 \tag{7.119}$$

where $K$ is the kernel of $\mu$.

When we have a short exact sequence of groups there is an important relation between them, as we now explain.

**Theorem 7.4.1**: Let $K \subseteq G$ be the kernel of a homomorphism (4.33). Then $K$ is a normal subgroup of $G$.

**Proof**: $\mu(gkg^{-1}) = \mu(g)\mu(k)\mu(g^{-1}) = \mu(g)1_{G'}\mu(g)^{-1} = 1_{G'} \Rightarrow K$ is normal. ♠

---

**Exercise** *Is the image of a homomorphism a normal subgroup?*

If $\mu : G \to G'$ is a group homomorphism is $\mu(G)$ a normal subgroup of $G'$? Answer the question with a proof or a counterexample. [81]

---

It follows by Theorem 7.2.1, that $G/K$ has a group structure. Note that $\mu(G)$ is also naturally a group.

These two groups are closely related because

$$\mu(g) = \mu(g') \qquad \leftrightarrow \qquad gK = g'K \tag{7.120}$$

Thus we have

**Theorem 7.4.2**:

$$\boxed{\mu(G) \cong G/\mathrm{Ker}(\mu)} \tag{7.121}$$

*Proof*: We associate the coset $gK$ to the element $\mu(g)$ in $G'$.

$$\psi : gK \mapsto \mu(g) \tag{7.122}$$

Claim: $\psi$ is an isomorphism. You have to show three things:

1. $\psi$ is a well defined map:

$$gK = g'K \Rightarrow \exists k \in K, g' = gk \Rightarrow \mu(g') = \mu(gk) = \mu(g)\mu(k) = \mu(g) \tag{7.123}$$

2. $\psi$ is in fact a homomorphism of groups

$$\psi(g_1K \cdot g_2K) = \psi(g_1K) \cdot \psi(g_2K) \tag{7.124}$$

   where on the LHS we have the product in the group $G/K$ and on the RHS we have the product in $G'$. We leave this as an exercise for the reader.

3. $\psi$ is one-one, i.e. $\psi$ is onto and invertible. The surjectivity should be clear. To prove injectivity note that:

$$\mu(g') = \mu(g) \Rightarrow \mu(g^{-1}g') = 1 \Rightarrow \exists k \in K, g' = gk \Rightarrow g'K = gK \qquad \spadesuit \tag{7.125}$$

**Remarks**:

[81] *Answer*: Definitely not! Any subgroup $H \subset G$ is the image of the inclusion homomorphism. In general, subgroups are not normal subgroups.

1. If we have a short exact sequence

$$1 \to N \to G \to Q \to 1 \tag{7.126}$$

then it automatically follows that $N$ is isomorphic to a normal subgroup of $G$ (because the image of $N$ under an injective homomorphism is the kernel of a homomorphism $G \to Q$). By some abuse of notation we sometimes write $G/N$ instead of $G/f(N)$ where $f$ is the injective homomorphism. With this understood, $Q$ is isomorphic to $G/N$. For this reason we call $Q$ the *quotient group*. A frequently used terminology is that *"$G$ is an extension of $Q$ by $N$."* Some authors [82] will use the terminology that "$G$ is an extension of $N$ by $Q$." So it is best simply to speak of a group extension with kernel $N$ and quotient $Q$.

2. **VERY IMPORTANT**: In quantum mechanics physical states are actually represented by "rays" in Hilbert space, or better, by one-dimensional subspaces of Hilbert space, or, even better, by orthogonal projection operators of rank one. (This is for "pure states." More generally, "states" are described mathematically by density matrices.) When comparing symmetries of quantum systems with their classical counterparts, group extensions play an important role so we will discuss them rather thoroughly in §15 below.

Here are three nice examples of commonly encountered short exact sequences:

**Example 0**: Note that any direct product group $G_1 \times G_2$ fits into two short exact sequences:

$$1 \to G_1 \to G_1 \times G_2 \to G_2 \to 1 \tag{7.127}$$

$$1 \to G_2 \to G_1 \times G_2 \to G_1 \to 1 \tag{7.128}$$

The reader should write out the homomorphisms in these two cases. We will call such short exact sequences "trivial." Importantly, there exist nontrivial short exact sequences. We will see many examples below.

**Example 1**: Consider the group of fourth roots of unity, $\mu_4$ and the homomorphism $\pi : \mu_4 \to \mu_2$ given by $\pi(g) = g^2$. The kernel is $\{\pm 1\} = \mu_2$ and so we have:

$$1 \to \mathbb{Z}_2 \to \mathbb{Z}_4 \to \mathbb{Z}_2 \to 1 \tag{7.129}$$

As an exercise the reader should also describe this extension thinking of $\mathbb{Z}_4$ additively as $\mathbb{Z}/4\mathbb{Z}$. Note that $\mathbb{Z}_4$ is NOT isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

This example generalizes in a nice way as follows. Let $n$ be any integer and consider the roots of unity: $\mu_{n^2}$. There is a homomorphism $\varphi : \mu_{n^2} \to \mu_n$ given by $\varphi(z) = z^n$.

[82] notably, S. MacLane, one of the inventors of group cohomology,

Clearly, the kernel is $\mu_n$, the $n^{th}$ roots of unity. It is not hard to see the image is all of $\mu_n$ so

$$1 \to \mathbb{Z}_n \to \mathbb{Z}_{n^2} \to \mathbb{Z}_n \to 1 \tag{7.130}$$

**Example 2**: Consider the homomorphism

$$r_N : \mathbb{Z} \to \mathbb{Z}/N\mathbb{Z} \tag{7.131}$$

given by reduction modulo $N$. (Or, if you prefer to think multiplicatively, $r_N(n) = \omega^n$ where $\omega$ is a primitive $N^{th}$ root of 1.) The kernel is $K = N\mathbb{Z} \subset \mathbb{Z}$. As a group this kernel is isomorphic to $\mathbb{Z}$ and so we have

$$0 \to \mathbb{Z} \overset{m_N}{\to} \mathbb{Z} \overset{r_N}{\to} \mathbb{Z}/N\mathbb{Z} \to 0 \tag{7.132}$$

where $m_N(x) = Nx$ is simply multiplication by $N$.

**Example 3**: *Finite Heisenberg Groups*: Let $P, Q$ be $N \times N$ "clock" and "shift" matrices. To define these introduce an $N^{th}$ root of unity, say $\omega = \exp[2\pi i/N]$. Then

$$P_{i,j} = \delta_{i=j+1 \mathrm{mod} N} \tag{7.133}$$

$$Q_{i,j} = \delta_{i,j}\omega^j \tag{7.134}$$

Note that $P^N = Q^N = 1$ and no smaller power is equal to 1. Further note that [83]

$$QP = \omega PQ \tag{7.135}$$

For $N = 4$ the matrices look like

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad Q = \begin{pmatrix} \omega & 0 & 0 & 0 \\ 0 & \omega^2 & 0 & 0 \\ 0 & 0 & \omega^3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{7.136}$$

with $\omega = e^{2\pi i/4}$. The group of matrices generated by $P, Q$ and $\omega \mathbf{1}_{N\times N}$ is a finite subgroup of $GL(N, \mathbb{C})$.

It is obvious that $Q^N = 1$. Note that, in a standard ordered basis $\{e_i\}_{i=1}^N$ for $\mathbb{C}^N$ the operator corresponding to the matrix $P$ above takes

$$e_1 \overset{P}{\to} e_2 \overset{P}{\to} \cdots \overset{P}{\to} e_N \overset{P}{\to} e_1 \tag{7.137}$$

from which it is obvious that $P^N = 1$. Thanks to the relation $QP = \omega PQ$ the general group element can be written in the form

$$\omega^a P^b Q^c \tag{7.138}$$

[83] The fastest way to check that - and thereby to check that you have your conventions under control - is to compute $QPQ^{-1}$ because $(Q^{-1}PQ)_{ij} = Q_{ii}P_{ij}(Q_{jj})^{-1} = \omega P_{ij}$.

where $a, b, c$ are integers and the group element only depends on the residue class of $a, b, c$ modulo $N$. The group multiplication law is:

$$(\omega^{a_1} P^{b_1} Q^{c_1}) \cdot (\omega^{a_2} P^{b_2} Q^{c_2}) = \omega^{a_3} P^{b_3} Q^{c_3} \tag{7.139}$$

where

$$\begin{aligned} a_3 &= a_1 + a_2 + c_1 b_2 \\ b_3 &= b_1 + b_2 \\ c_3 &= c_1 + c_2 \end{aligned} \tag{7.140}$$

This group, as an abstract group is called a *finite Heisenberg group* and denoted $\mathrm{Heis}_N$. It is an extension

$$1 \to \mathbb{Z}_N \to \mathrm{Heis}_N \overset{\pi}{\to} \mathbb{Z}_N \times \mathbb{Z}_N \to 1 \tag{7.141}$$

where $\pi(\omega^a P^b Q^c) = (b \mathrm{mod} N, c \mathrm{mod} N)$ is a surjective homomorphism. The finite Heisenberg group has many many pretty applications to physics and we will return to this group several times below. See, for example section 11.17.2 below for a physical interpretation.

---

**Exercise** $A_n$
Use Theorem 7.1 to show that $A_n$ is a normal subgroup of $S_n$.

---

---

**Exercise** *Phases And Norms*
a.) Show that $z \mapsto |z|$ is a homomorphism $\mathbb{C}^* \to \mathbb{R}_{>0}$.
b.) Show that $\mathbb{C}^*/U(1) \cong \mathbb{R}_{>0}$.

---

---

**Exercise** *The Exponential Map And Short Exact Sequences*
Show that the exponential map $p(x) = e^{2\pi \mathrm{i} x}$ defines short exact sequences

$$0 \to \mathbb{Z} \to \mathbb{C} \to \mathbb{C}^* \to 1 \tag{7.142}$$

$$0 \to \mathbb{Z} \to \mathbb{R} \to U(1) \to 1 \tag{7.143}$$

These are very important in topology and in the theory of covering spaces.

---

---

**Exercise** *A Nontrivial Short Exact Sequence*

Show that (7.129) is a nontrivial short exact sequence by showing that $\mathbb{Z}_4$ is not isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. [84]

---

♣Should put in solutions for the next two exercises. ♣

---

**Exercise** *Induced Maps On Quotient Groups*

We will use the following result in §12.3: Suppose $\mu : G_1 \to G_2$ is a homomorphism and $H_2 \subset G_2$ is a subgroup. Recall that $\mu^{-1}(H_2) \subset G_1$ is a subgroup.

a.) If $H_1 \subset \mu^{-1}(H_2)$ is a subgroup show that there is an induced map $\bar{\mu} : G_1/H_1 \to G_2/H_2$. [85]

b.) Show that if $H_1$ and $H_2$ are normal subgroups then $\bar{\mu}$ is a homomorphism.

c.) In this case there is an exact sequence

$$1 \to \mu^{-1}(H_2)/H_1 \to G_1/H_1 \to G_2/H_2 \tag{7.144}$$

---

---

**Exercise** *Induced Maps On Quotients Of Abelian Groups*

Let $A, B$ be abelian groups and $A_1 \subset A$ and $B_1 \subset B$ subgroups, and suppose $\phi : A \to B$ is a homomorphism such that $\phi$ takes $A_1$ into $B_1$. That is $A_1 \subset \phi^{-1}(B_1)$.

It follows from the previous exercise that there is a homomorphism:

$$\bar{\phi} : A/A_1 \to B/B_1 \tag{7.145}$$

Show that if $\phi : A_1 \to B_1$ is *surjective* then

$$\ker\{\bar{\phi} : A/A_1 \to B/B_1\} \cong \frac{\ker\{\phi : A \to B\}}{\ker\{\phi : A_1 \to B_1\}} \tag{7.146}$$

---

---

**Exercise**

Let $n$ be a natural number and let

$$\psi : \mathbb{Z}/n\mathbb{Z} \to (\mathbb{Z}/n\mathbb{Z})^d \tag{7.147}$$

be given by the diagonal map $\psi(\omega) = (\omega, \cdots, \omega)$.

Find a set of generators and relations for $G/\psi(H)$.

---

[84] *Answer*: One answer is to note that every element in $\mathbb{Z}_2 \times \mathbb{Z}_2$ is order two, but $\mathbb{Z}_4$ has elements of order four. Under isomorphism, the order of a group element is preserved.

[85] *Answer*: Show that $\bar{\mu}(g_1 H_1) := \mu(g_1) H_2$ is a well-defined map.

**Exercise**

Let $G = \mathbb{Z} \times \mathbb{Z}_4$. Let $K$ be the subgroup generated by $(2, \omega^2)$ where we are writing $\mathbb{Z}_4$ as the multiplicative group of $4^{th}$ roots of 1. Note $(2, \omega^2)$ is of infinite order so that $K \cong \mathbb{Z}$. Show that $G/K \cong \mathbb{Z}_8$.

**Exercise** *The Finite Heisenberg Groups*

a.) Using the matrices of (7.133) and (7.134) show that the word

$$P^{n_1} Q^{m_1} P^{n_2} Q^{m_2} \cdots P^{n_k} Q^{m_k} \tag{7.148}$$

where $n_i, m_i \in \mathbb{Z}$ can be written as $\xi P^x Q^y$ where $x, y \in \mathbb{Z}$ and $\xi$ is an $N^{th}$ root of unity. Express $x, y, \xi$ in terms of $n_i, m_i$. [86]

b.) Find a presentation of $\mathrm{Heis}_N$ in terms of generators and relations. ♣Generators and relations have been moved to later. Move this exercise. ♣

c.) What is the order of $\mathrm{Heis}_N$ ? [87]

**Exercise** *Centrally symmetric shuffles*

Let us consider again the permutation group of the set $\{0, 1, \ldots, 2n-1\}$. Recall we let $W(B_n)$ denote the subgroup of $S_{2n}$ of centrally symmetric permutations which permutes the pairs $x + \bar{x} = 2n - 1$ amongst themselves.

Show that there is an exact sequence

$$1 \to \mathbb{Z}_2^n \to W(B_n) \to S_n \to 1 \tag{7.149}$$

and therefore $|W(B_n)| = 2^n n!$.

**Exercise** *The Weyl Group Of $SU(N)$*

Show that there is an exact sequence

$$1 \to T \to N(T) \to S_N \to 1 \tag{7.150}$$

[86] *Answer*: $x = \sum_i n_i$, $y = \sum_i m_i$, $\xi = \omega^c$ with $c = \sum_{i<j} m_i n_j$.

[87] *Answer* $N^3$

### 7.5 Conjugacy Classes In $S_n$

Above we discussed the cycle decomposition of elements of $S_n$. Now let us study how the cycles change under conjugation.

When showing that transpositions generate $S_n$ we noted the following fact:

*If $(i_1 i_2 \cdots i_k)$ is a cycle of length $k$ then, for any $g \in S_n$, $g(i_1 i_2 \cdots i_k)g^{-1}$ is also a cycle of length $k$.* It is the cycle where we replace $i_1, i_2, \ldots$ by their images under $g$. That is, if $g(i_a) = j_a$, $a = 1, \ldots, k$, then $g(i_1 i_2 \cdots i_k)g^{-1} = (j_1 j_2 \cdots j_k)$.

It therefore follows that:

*Any two cycles of length $k$ are conjugate.*

**Example** In $S_3$ there are two cycles of length 3 and they are indeed conjugate:

$$(12)(123)(12)^{-1} = (213) = (132) \tag{7.151}$$

Now recall that any element in $S_n$ can be written as a product of disjoint cycles. We can characterize a cycle decomposition by giving the number $\ell_j$ of cycles of length $j$, where $j$ ranges from 1 to $n$.

**Examples**:

1. The following two permutations in $S_{12}$ are conjugate:

$$(1,2)(3,4)(5,6)(7,8,9)(10,11,12) \tag{7.152}$$

$$(4,10)(7,8)(9,11)(1,12,6)(2,5,3) \tag{7.153}$$

   This has $\ell_1 = 0$, $\ell_2 = 3$, $\ell_3 = 2$, $\ell_j = 0$ for $j > 3$.

2. In $S_4$ there are 3 elements with cycle decomposition of type $(ab)(cd)$:

$$(12)(34), \qquad (13)(24), \qquad (14)(23) \tag{7.154}$$

   Note that these can be conjugated into each other by suitable transpositions. So this conjugacy class is determined by

$$\ell_1 = 0 \qquad \ell_2 = 2 \qquad \ell_3 = 0 \qquad \ell_4 = 0 \tag{7.155}$$

Associated to a cycle decomposition is a sequence of integers $\ell_j$. We can summarize this data with the notation:

$$(1)^{\ell_1}(2)^{\ell_2}\cdots(n)^{\ell_n} \tag{7.156}$$

Then, since we must account for all $n$ letters being permuted we must have:

$$\begin{aligned} n &= \underbrace{1+\cdots+1}_{\ell_1 \text{ times}} + \underbrace{2+\cdots+2}_{\ell_2 \text{ times}} + \cdots + \underbrace{n+\cdots}_{\ell_n \text{ times}} \\ &= 1\cdot\ell_1 + 2\cdot\ell_2 + \cdots n\cdot\ell_n \\ &= \sum_{j=1}^{n} j\cdot\ell_j \end{aligned} \tag{7.157}$$

But the cycles are disjoint, and any two cycles can be conjugated into each other:

*Therefore, the conjugacy classes in $S_n$ are labeled by specifying for each $j$ such that $1 \le j \le n$, a nonnegative integer, denoted $\ell_j$ such that*

$$\sum_{j=1}^{n} j\ell_j = n \tag{7.158}$$

*Here $\ell_j$ encodes the number of distinct cycles of length $j$ in the cycle decomposition of any element $\sigma$ in the conjugacy class $C(\sigma)$.*

**Definition** A decomposition of a positive integer $n$ into a sum of nonnegative integers is called a *partition of $n$.*

Therefore:

| The conjugacy classes of $S_n$ are in 1-1 correspondence with the partitions of $n$. |
|---|

**Definition** The number of distinct partitions of $n$ is called the partition function of $n$, and denoted $p(n)$. [88]

**Example** For $n = 4, 5$ $p(4) = 5$ and $p(5) = 7$ and the conjugacy classes of $S_4$ and $S_5$ are:

| Partition | Cycle decomposition | Typical $g$ | $\vert C(g)\vert$ | Order of $g$ |
|---|---|---|---|---|
| $4 = 1+1+1+1$ | $(1)^4$ | 1 | 1 | 1 |
| $4 = 1+1+2$ | $(1)^2(2)$ | $(ab)$ | $\binom{4}{2} = 6$ | 2 |
| $4 = 1+3$ | $(1)(3)$ | $(abc)$ | $2\cdot 4 = 8$ | 3 |
| $4 = 2+2$ | $(2)^2$ | $(ab)(cd)$ | $\frac{1}{2}\binom{4}{2} = 3$ | 2 |
| $4 = 4$ | $(4)$ | $(abcd)$ | 6 | 4 |

[88] This is a term in number theory. It is not to be confused with the "partition function" of a field theory!

| Partition | Cycle decomposition | $\|C(g)\|$ | Typical $g$ | Order of $g$ |
|---|---|---|---|---|
| $5 = 1+1+1+1+1$ | $(1)^5$ | 1 | 1 | 1 |
| $5 = 1+1+1+2$ | $(1)^3(2)$ | $\binom{5}{2} = 10$ | $(ab)$ | 2 |
| $5 = 1+1+3$ | $(1)^2(3)$ | $2 \cdot \binom{5}{3} = 20$ | $(abc)$ | 3 |
| $5 = 1+4$ | $(1)(4)$ | $6 \cdot \binom{5}{4} = 30$ | $(abcd)$ | 4 |
| $5 = 1+1+2+2$ | $(1)(2)^2$ | $5 \cdot \frac{1}{2}\binom{4}{2} = 15$ | $(ab)(cd)$ | 2 |
| $5 = 2+3$ | $(2)(3)$ | $2 \cdot \binom{5}{2} = 20$ | $(ab)(cde)$ | 6 |
| $5 = 5$ | $(5)$ | $4! = 24$ | $(abcde)$ | 5 |

♣Add table for $S_6$ . ♣

We will show below, twice, that the order of the conjugacy class of type (7.156) is

$$|C(g)| = \frac{n!}{\prod_{j=1}^{n} j^{\ell_j} \ell_j!} \,. \tag{7.159}$$

---

**Exercise** *Sign of the conjugacy class*

Let $\epsilon : S_n \to \{\pm 1\}$ be the sign homomorphism. Show that $\epsilon(g) = (-1)^{n+\sum_j \ell_j}$ if $g$ is in the conjugacy class (7.156).

---

### 7.5.1 Conjugacy Classes In $S_n$ And Harmonic Oscillators

There is a beautiful relation of conjugacy classes of the symmetric group with special collections of harmonic oscillators. We'll give a taste of how that happens here.

Let's review briefly some facts about the quantum mechanical harmonic oscillator: The classical harmonic oscillator is described by a phase space with coordinate $q \in \mathbb{R}$ (the displacement of the oscillator) and momentum $p \in \mathbb{R}$ with Hamiltonian

$$H = \frac{1}{2}(p^2 + \omega^2 q^2) \tag{7.160}$$

(we have scaled away the mass). In our considerations we will assume that $\omega > 0$ is a positive real number. The classical Poisson bracket $\{p, q\} = 1$ is quantized by postulating there are operators $\hat{p}, \hat{q}$ with

$$[\hat{p}, \hat{q}] = -\mathrm{i}\hbar \tag{7.161}$$

and there is a Hilbert space representing this operator algebra. One forms the linear combinations:

$$\begin{aligned} a &:= \frac{1}{\sqrt{2\hbar\omega}}(\omega\hat{q} + \mathrm{i}\hat{p}) \\ \bar{a} &:= \frac{1}{\sqrt{2\hbar\omega}}(\omega\hat{q} - \mathrm{i}\hat{p}) \end{aligned} \tag{7.162}$$

so that

$$[a, \bar{a}] = 1 \tag{7.163}$$

**Remarks**:

1. The operators $a, \bar{a}$ generate an algebra $\mathcal{A}$ (see Chapter 2 for the general notion of an *algebra*) known as a $*$ algebra. What this means is that there is an involutive $\mathbb{C}$-antilinear map $\mathcal{O} \to \mathcal{O}^*$ so that $(\mathcal{O}_1\mathcal{O}_2)^* = \mathcal{O}_2^*\mathcal{O}_1^*$. We define it by $\hat{q}^* = \hat{q}$ and $\hat{p}^* = \hat{p}$ and then extend by to linear combinations by demanding that $*$ is $\mathbb{C}$-antilinear. Then $\bar{a} = a^*$. In a $*$-representation of a $*$-algebra on a Hilbert space $*$ becomes Hermitian adjoint.

2. If we just consider the operator algebra without considering $H$ then there are a number of different ways to represent it. We could postulate there is a vector $|0\rangle$ with $a|0\rangle = 0$. Then the Hilbert space is spanned by $\bar{a}^n|0\rangle$ and $\bar{a} = a^\dagger$. But note that we could make a linear transformation to

$$\begin{aligned} b &= \alpha a + \beta \bar{a} \\ \bar{b} &= \gamma a + \delta \bar{a} \end{aligned} \tag{7.164}$$

As we saw in the exercise containing equation (3.60) et. seq., this will preserve the commutation relations: $[b, \bar{b}] = 1$, if $\alpha\delta - \beta\gamma = 1$. If we furthermore wish to preserve the $*$ structure so that $\bar{b} = b^*$ then $\delta = \alpha^*$ and $\gamma = \beta^*$. The group of such matrices:

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \tag{7.165}$$

with $|\alpha|^2 - |\beta|^2 = 1$ forms the group

$$SU(1,1) := \{A \in M_2(\mathbb{C}) | A^\dagger \eta A = \eta\} \tag{7.166}$$

We can define a "different;" representation of the Heisenberg algebra starting with $b|0\rangle_b = 0$. For a finite number of oscillators the representations will be equivalent by a *Bogoliubov transformation.* If $\mathcal{H}_a$ is the representation generated by a vector with $a|0\rangle_a = 0$ and $\mathcal{H}_b$ is the representation generated by a vector with $b|0\rangle_b = 0$ then there is a unitary map $\mathcal{U} : \mathcal{H}_b \to \mathcal{H}_a$ implementing (7.164). We will show in exercise (7.234) et. seq. below that in fact:

$$\mathcal{U}(|0\rangle_b) = \mathcal{N}(\cosh r)^{-1/2} e^{-\frac{1}{2}\Gamma(a^\dagger)^2}|0\rangle_a \tag{7.167}$$

where we parametrize

$$\begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} = \begin{pmatrix} \cosh r & -e^{-i\phi}\sinh r \\ -e^{i\phi}\sinh r & \cosh r \end{pmatrix} \tag{7.168}$$

and in terms of these parameters $\Gamma = e^{i\phi}\tanh(r)$. The parameter $r$ has an interpretation in terms of the expected number of "$b$ quanta" in the "$a$-vacuum":

$$ {}_a\langle 0|b^\dagger b|0\rangle_a = \sinh^2 r \tag{7.169} $$

Such Bogoliubov transformations play an important role in the theory of superfluids and superconductivity and in the theory of particle creation by time-varying electromagnetic and gravitational fields. Although the oscillator representations generated by $|0\rangle_a$ and $|0\rangle_b$ are equivalent, in the case of an infinite number of oscillators the representations can be inequivalent.

3. This choice of quantization based on $|0\rangle_a$ is preferred when we consider the oscillator Hamiltonian: The quantum Hamiltonian is $H = \omega(a^\dagger a + \frac{1}{2})$. The states $\bar{a}^n|0\rangle$ are eigenstates of $H$ with eigenvalue $\omega(n+\frac{1}{2})$. Assuming $|0\rangle$ has unit norm the normalized eigenstates $\frac{1}{\sqrt{n!}}\bar{a}^n|0\rangle$ form a complete ON basis for the Hilbert space.

In quantum statistical mechanics a very important quantity is the (physics) *partition function* defined to be

$$ \mathrm{Tr}_{\mathcal{H}_{\text{single h.o.}}} e^{-\beta H} = \frac{e^{-\frac{1}{2}\beta\omega}}{1-e^{-\beta\omega}} = \frac{1}{2\sinh\frac{1}{2}\beta\omega} \tag{7.170} $$

Here $\beta$ has the physical interpretation of $1/(kT)$ where $k$ is Boltzmann's constant and $T$ is the temperature above absolute zero. Using the (physics) partition function one derives thermodynamic quantities when the oscillator is connected to a heat bath.

Now, suppose we have a system which is described by an infinite collection of harmonic oscillators:

$$ [a_j, a_k] = 0 \qquad [a_j^\dagger, a_k^\dagger] = 0 \qquad [a_j, a_k^\dagger] = \delta_{j,k} \qquad j,k = 1,\dots \tag{7.171} $$

Suppose they have frequencies which are all a multiple of a basic harmonic which we'll denote $\omega_0$, so the frequencies associated with the oscillators $a_1, a_2, a_3, \dots$ are $\omega_0, 2\omega_0, 3\omega_0, \dots$. The motivation for choosing all frequencies to be multiples of a basic frequency comes from the theory of strings, as we will explain below.

If we write the standard sum of harmonic oscillator Hamiltonians we get, formally,

$$ H^{\text{formal}} = \sum_{j=1}^{\infty} j\omega_0(a_j^\dagger a_j + \frac{1}{2}) \tag{7.172} $$

This is formal, because of the infinite sum.

We represent the operator algebra by a "highest weight representation" starting with a vector (more precisely, a line - a one-dimensional vector space) annihilated by the $a_j$:

$$ a_j|0\rangle = 0 \qquad j = 1,2,3,\dots \tag{7.173} $$

and then consider the span of all states made by acting with $a_j^\dagger$ and then completing. The result is a Hilbert space

$$ \mathcal{H}_{\text{tot}} = \mathcal{H}_{a_1} \otimes \mathcal{H}_{a_2} \otimes \cdots \tag{7.174} $$

Now the formal Hamiltonian does not make sense as an operator on $\mathcal{H}_{\rm tot}$ because the groundstate energy is infinite. This is typical of the divergences of quantum field theory [89] An infinite number of degrees of freedom typically leads to divergences in physical quantities.

One way to make sense of the infinite ground state energy is via a procedure known as $\zeta$-function regularization, and a corresponding renormalization. The infinite ground state energy is

$$\sum_{j=1}^{\infty} \frac{j}{2}\omega = \frac{\omega_0}{2}\sum_{j=1}^{\infty}\frac{1}{j^{-1}} \tag{7.175}$$

We regularize the sum $\sum_j j$ by introducing the Riemann zeta function:

$$\zeta(s) := \sum_{j=1}^{\infty}\frac{1}{j^s} \tag{7.176}$$

This series converges absolutely to an analytic function of $s$ in the half-plane $Re(s) > 1$. One can show that it admits an analytic continuation in $s$ to the entire complex plane with a simple pole at $s = 1$. With this analytic continuation one finds $\zeta(-1) = -1/12$. [90] In this way we can define the renormalized ground state energy to be

$$\sum_{j=1}^{\infty} \frac{j}{2}\omega := \frac{\omega_0}{2}\zeta(-1) = -\frac{\omega_0}{24} \tag{7.177}$$

Equation (7.177) can be justified much more rigorously using the appearance of a ground state energy in a Euclidean path integral and $\zeta$-function definition of determinants, as described below. In any case, things work out very nicely if we take the Hamiltonian to be:

$$H = \sum_{j=1}^{\infty} j\omega_0 a_j^\dagger a_j - \frac{\omega_0}{24} \tag{7.178}$$

A natural basis for $\mathcal{H}$ given by states of the form:

$$(a_1^\dagger)^{\ell_1}(a_2^\dagger)^{\ell_2}\cdots(a_n^\dagger)^{\ell_n}|0\rangle \tag{7.179}$$

[89] The quantum field theory in question is that of a massless scalar field in a spacetime of 1+1 dimensions. See a few paragraphs below.

[90] One way to see that $\zeta(-1) = -1/12$ is to use the functional equation for the Riemann zeta function. It is straightforward to see that $2\pi^{-s/2}\Gamma(s/2)\zeta(s) = \int_0^\infty x^{s/2}(\vartheta({\rm i}x) - 1)dx$ where $\vartheta({\rm i}x) = \sum_{n\in\mathbb{Z}} e^{-\pi n^2 x}$ is the Riemann theta function. The integral on the RHS has an analytic continuation in $s$ and the functional equation $\Gamma(x+1) = x\Gamma(x)$ defines the analytic continuation of the $\Gamma$ function so this integral representation defines an analytic continuation of the Riemann zeta function. Now the theta function satisfies a nice modular property relating $\vartheta({\rm i}x)$ to $\vartheta({\rm i}/x)$. This can be proven by a straightforward application of the Poisson summation formula. See also section *** below for a physical derivation of the modular transformation law of the theta function. Defining $\xi(s) = \frac{1}{2}\pi^{-s/2}s(s-1)\Gamma(\frac{s}{2})\zeta(s)$ one proves the remarkable functional equation: $\xi(s) = \xi(1-s)$ by breaking the integral up into an interval $[0,1]$ and an interval $[1,\infty)$ and making the change of variable $x \to 1/x$. Now, we evaluate the functional equation at $s = 2$. Using the identity $\Gamma(x)\Gamma(1-x) = \pi/\sin\pi x$ one can evaluate $\Gamma(1/2) = \sqrt{\pi}$ and therefore $\Gamma(-1/2) = -2\sqrt{2}$. On the other hand $\zeta(2) = \pi^2/6$, as can be proven in many ways using elementary calculus.

This state has energy

$$\omega_0 \sum_{j=1}^{n} j\ell_j - \frac{\omega_0}{24} \tag{7.180}$$

It follows that the space of states with energy $n\omega_0$ above the ground state has dimension $p(n)$ and the above vectors are a natural basis in 1-1 correspondence with the conjugacy classes of $S_n$. This turns out to be significant in the boson-fermion correspondence in 1+1 dimensional quantum field theory: There is an equivalent formulation in terms of fermionic oscillators. It follows that the quantum statistical mechanical partition function of this infinite collection of oscillators is

$$Z^{osc}(\beta) = \mathrm{Tr}_{\mathcal{H}_{\mathrm{tot}}} e^{-\beta H} = e^{-\beta\omega_0/24}\left(1+\sum_{n=1}^{\infty} p(n) e^{-\beta n \omega_0}\right) := q^{-1/24}\left(1+\sum_{n=1}^{\infty} p(n) q^n\right) \tag{7.181}$$

where $q = e^{-\beta\omega_0}$.

On the other hand, since $\mathcal{H}_{\mathrm{tot}} = \mathcal{H}_{a_1} \otimes \mathcal{H}_{a_2} \otimes \cdots$ we have

$$\begin{aligned} Z^{osc}(\beta) &= \mathrm{Tr}_{\mathcal{H}_{\mathrm{tot}}} e^{-\beta H} \\ &= e^{-\beta\omega_0/24} \prod_{j=1}^{\infty} \mathrm{Tr}_{\mathcal{H}_{a_j}} e^{-\beta j \omega_0 a_j^\dagger a_j} \\ &= \frac{1}{q^{24} \prod_{j=1}^{\infty}(1-q^j)} \end{aligned} \tag{7.182}$$

We therefore have derived an infinite product representation of a generating function for the mathematical partition function $p(n)$ using the physical statmech partition function. The coincidence of names is, so far as I know, completely fortuitous.

The infinite product identity for the partition function can also be derived in a straightforward mathematical way as follows: Let $q$ be a complex number with $|q| < 1$. Notice that:

$$\begin{aligned} \frac{1}{\prod_{j=1}^{\infty}(1-q^j)} &= (1+q+q^2+\cdots)(1+q^2+q^4+\cdots)(1+q^3+q^6+\cdots)\cdots \\ &= \prod_{j=1}^{\infty}\left(1+\sum_{\ell_j=1}^{\infty}(q^j)^{\ell_j}\right) \\ &= 1+\sum_{n=1}^{\infty} p(n) q^n \end{aligned} \tag{7.183}$$

Expanding out (7.183) gives the first few values of $p(n)$:

$$\begin{aligned} &1+q+2q^2+3q^3+5q^4+7q^5+11q^6+15q^7+22q^8+30q^9+ \\ &+42q^{10}+56q^{11}+77q^{12}+101q^{13}+135q^{14}+\cdots \end{aligned} \tag{7.184}$$

and one can easily generate the first few 100 values using Maple or Mathematica or Sage or ...

The quantum statistical mechanical partition function of this collection of oscillators has a truly remarkable "modular transformation property" that relates $Z(\beta)$ to $Z(1/\beta)$. The precise statement is:

$$\beta^{-1/4}Z^{osc}(\beta) = \tilde{\beta}^{-1/4}Z^{osc}(\tilde{\beta}) \tag{7.185}$$

$$\beta\tilde{\beta} = \left(\frac{2\pi}{\omega_0}\right)^2 \tag{7.186}$$

Equations (7.185) and (7.186) should be viewed as a high-low temperature "duality." As an exercise the reader should show that the high temperature limit of $Z^{\rm osc}(\beta)$ diverges like:

$$Z^{\rm osc}(\beta) \quad \sim_{\beta\to 0} \quad \sqrt{\frac{\beta\omega_0}{2\pi}}\exp\left(\frac{\pi^2}{6\beta\omega_0}\right) \tag{7.187}$$

The property (7.185) and (7.186) is proven in textbooks on analytic number theory. In this context one often uses the variable

$$\tau = {\rm i}\frac{\beta\omega}{2\pi} \tag{7.188}$$

so that $q = e^{-\beta\omega} = e^{2\pi{\rm i}\tau}$. Analytic number theorists define the *Dedekind eta function* :

$$\eta(\tau) := \exp\left(\frac{2\pi{\rm i}\tau}{24}\right)\prod_{n=1}^{\infty}(1-q^n) \tag{7.189}$$

and prove the crucial identity:

$$\eta(-1/\tau) = (-{\rm i}\tau)^{1/2}\eta(\tau). \tag{7.190}$$

This equation is equivalent to the identity in equations (7.185) and (7.186). In fact, (7.189) admits an analytic continuation to the entire upper half plane ${\rm Im}(\tau) > 0$ as a holomorphic function. It follows that $Z^{\rm osc}(\beta)$ admits an analytic continuation to the $Re(\beta) > 0$. However, one must be careful about interpreting the analytic continuation in terms of physics. See below for the physical discussion.

An Interpretation From Physics

It turns out that (7.185) and (7.186), or equivalently, (7.190) has a beautiful interpretation from physics. Consider a circular string with coordinate $\sigma \sim \sigma + 2\pi$ and the displacement of the string $X(\sigma, t)$ is a real number depending on position and time. $X(\sigma, t)$ represents the displacement of the string from some equilibrium position. It therefore has units of length. The action is:

$$S = \frac{1}{4\pi\ell_s^2}\int dt \int_0^{2\pi} d\sigma \left((\partial_t X)^2 - \partial_\sigma X)^2\right) \tag{7.191}$$

where $\ell_s$ has units of length, or inverse mass, and we have temporarily set $\hbar = c = 1$ by choice of units. If one derives this action from the theory of elasticity then it becomes clear that $T = \ell_s^{-2}$ is the tension of the string. Note that we could absorb $\ell_s$ into the field $X$

to make a dimensionless field $\tilde{X}(t,\sigma) = X(t,\sigma)/\ell_s$. We will henceforth do this (and drop the tilde on $X$). Then the general solution of the classical equation of motion (the wave equation) $(\partial_t^2 - \partial_\sigma^2)X = 0$ is:

$$X(t,\sigma) = X_0 + pt + \mathrm{i}\frac{1}{\sqrt{2}}\sum_{n\neq 0}\left(\frac{\alpha_n}{n}e^{\mathrm{i}n(t+\sigma)} + \frac{\tilde{\alpha}_n}{n}e^{\mathrm{i}n(t-\sigma)}\right) \tag{7.192}$$

with complex numbers $\alpha_n = (\alpha_{-n})^*$ and $\tilde{\alpha}_n = (\tilde{\alpha}_{-n})^*$ and $X_0, p$ are real.

We can think of this as a $1+1$ dimensional field theory. Then spacetime is a cylinder $S^1 \times \mathbb{R}$ with Lorentzian metric. The $\alpha_n$ are the amplitudes of waves moving at the speed of light to the left, while the $\tilde{\alpha}_n$ are the amplitudes of waves moving at the speed of light to the right.

The Hamiltonian computed from the action is

$$H = \frac{1}{4\pi}\int_0^{2\pi} d\sigma\left((\partial_t X)^2 + \partial_\sigma X)^2\right) \tag{7.193}$$

and evaluating $H$ on the general solution of the equation of motion gives:

$$H = \frac{1}{2}p^2 + \frac{1}{2}\sum_{n\neq 0}(\alpha_{-n}\alpha_n + \tilde{\alpha}_{-n}\tilde{\alpha}_n) \tag{7.194}$$

When quantizing this system we find that $[X_0, p] = \mathrm{i}\hbar$ and

$$[\alpha_n, \alpha_m] = n\delta_{n+m,0} \qquad [\tilde{\alpha}_n, \tilde{\alpha}_m] = n\delta_{n+m,0} \qquad [\alpha_n, \tilde{\alpha}_m] = 0 \tag{7.195}$$

If we represent the Heisenberg algebras with vacua so that $\alpha_n|0\rangle = \tilde{\alpha}_n|0\rangle = 0$ for $n > 0$ then we get standard Harmonic oscillators by defining $a_n = \alpha_n/\sqrt{n}$ and $a_n^\dagger = \alpha_{-n}/\sqrt{n}$ for $n > 0$, and similarly for the right-moving modes $\tilde{\alpha}_n$. Therefore, the Hamiltonian becomes:

$$H = \frac{1}{2}\hat{p}^2 + \left[\sum_{n=1}^{\infty} n\left(a_n^\dagger a_n + \tilde{a}_n^\dagger \tilde{a}_n\right) - \frac{2}{24}\right] \tag{7.196}$$

We thus recognize $\omega_0 = 1$ as our basic frequency.

**Remark**: In equation (7.196) we have used the trick (7.177). This trick indeed gives the correct Casimir energy for a massless scalar field on an interval. The actions above are much the same except that now $\sigma \in [0,\pi]$ and we must impose boundary conditions at $\sigma = 0, \pi$. Making the simple choice $X(t,\sigma)|_{\sigma=0,\pi} = 0$ we find the general solution:

$$X(t,\sigma) = \sum_{n\neq 0}\left(\frac{\alpha_n}{n}e^{\mathrm{i}nt}\sin(n\sigma)\right) \tag{7.197}$$

so we only have one set of oscillators. If we make the interval of length $L$ then we scale $\sigma$ by $L$. Then $\partial_t X$ and $\partial_\sigma X$ scale by $1/L$ while the integration in the formula for the Hamiltonian goes from 0 to $\pi L$. The resulting the ground state energy - with $\zeta$-function definition is:

$$E_{\text{ground}} = -\frac{1}{24L} = -\frac{\hbar c}{24L} \tag{7.198}$$

where in the second equation we restored $\hbar$ and $c$ which had been set to 1. Of course, unless the system is coupled to gravity, the zero of energy is arbitrary. Here the zero of energy is defined by saying the massless scalar field on the real line has zero groundstate energy. Then the above formula for the Casimir energy is meaningful because we are comparing two energies. Put differently what is meaningful, and physical, and independent of the choice of regularization and renormalization of the naively infinite ground state energy is the Casimir force $-\frac{\partial E_{\text{ground}}}{\partial L}$. In analogous situations with the free electromagnetic field in $1+3$ dimensions the Casimir force has been experimentally measured. Beginning with [91] . The experiments have since been refined and agree with theory to within a percent.

Returning to the closed string, with both left-movers and right-movers we now study the quantum statmech partition function for this string:

$$Z(\beta) := Tre^{-\beta H} = \left(\frac{R}{\ell_s}\right)(2\pi\beta\omega)^{-1/2}\left(Z^{osc}((\beta)\right)^2 \tag{7.199}$$

There is one subtlety: It is infinite, because of the infinite volume of the target space. The zeromodes have a phase space density $dXdp$ and

$$\int dXdpe^{-\frac{1}{2}\beta p^2} \sim (\int dX)\beta^{-1/2} \tag{7.200}$$

We therefore make the range of $X$ periodic: We make the target space a periodic "box" - in this case, just a circle of radius $R$. It is possible to evaluate the finite $R$ corrections exactly and in the large $R$ limit the contribution of the zeromodes is

$$R\sqrt{2\pi/\beta} \tag{7.201}$$

as expected. [92] Recall the radius here is expressed in string units. If we restore $\ell_s$ we replace $R \to R/\ell_s$. Now, accounting for left- and right-moving oscillators we have:

$$Z(\beta) := Tre^{-\beta H} = R\sqrt{2\pi/\beta}(Z^{osc}((\beta))^2 \tag{7.203}$$

We have written this for $\beta$ real and positive. In the limit $R \to \infty$ it is the partition function per unit volume that has a finite limit.

Moreover, it is a standard and important result that for real $\beta$ the partition function $Tre^{-\beta H}$ can be written as a path integral with periodic Euclidean time of period $\beta$. The chain of logic is that the partition function is of the form

$$\text{Tr}e^{-\beta H} = \sum_{\psi_n}\langle\psi_n|e^{-\beta H}|\psi_n\rangle \tag{7.204}$$

[91] S.K. Lamoreaux, "Demonstration of the Casimir Force in the 0.6 to 6 $\mu$m Range," PRL 78 (1997)pp. 5-8; U. Mohideen and Anushree Roy, "Precision Measurement of the Casimir Force from 0.1 to 0.9 $\mu$m," PRL 81 (1998) 4549-4552 arXiv:physics/9805038.

[92] Technically, the exact partition function for finite radius $R$ target space is $\Theta_R/(\eta)^2$ where $\Theta_R$ is a Siegel-Narain theta function. The exact formula - in our present case - is

$$\Theta_R = \left(\sum_n e^{-\beta n^2/(2R^2)}\right)\left(\sum_w e^{-\beta w^2 R^2/2}\right) \tag{7.202}$$

The large $R$ asymptotics is evaluated with the help of the modular transformation law.

where $\psi_n$ is a basis of the space of states. But now $\langle\psi_1|e^{-\beta H}|\psi_2\rangle$ is an analytic continuation of the transition amplitude $\langle\psi_1|e^{-\mathrm{i}tH}|\psi_2\rangle$ to imaginary time. On the other hand $\langle\psi_1|e^{-\mathrm{i}tH}|\psi_2\rangle$ can be written as a path integral with initial and final conditions $\psi_2, \psi_1$, respectively. If we set $\psi_1 = \psi_2$ and sum over a complete basis then the domain of the path integral becomes that of all field configurations on the circle of Euclidean time. For more on this, see the classic book by Feynman and Hibbs.

If we apply the above principle to the case of our $1+1$ dimensional QFT, the quantum field $X$ at fixed time is a map from the circle to the real line so altogether we have a path integral on a torus $S^1 \times S^1$ with one factor for space and one factor for Euclidean time. Comparing with the Euclidean action of the string we see the metric on the torus is:

$$ds^2 = (d\sigma)^2 + \beta^2(d\sigma^2)^2 = (2\pi)^2\left[(d\sigma^1)^2 + \left(\frac{\beta}{2\pi}\right)^2(d\sigma^2)^2\right] \tag{7.205}$$

where we have chosen a dimensionless coordinate $\sigma^2 \sim \sigma^2+1$ in the Euclidean time direction and rescaled $\sigma = 2\pi\sigma^1$ so that $\sigma^1 \sim \sigma^1 + 1$.

Note that, if we exchange the Euclidean time circle for the spatial circle then we exchange $\sigma^1 \leftrightarrow \sigma^2$. [93] Up to an overall normalization of the metric this effectively changes $\beta$ to $\tilde{\beta}$ where

$$\beta\tilde{\beta} = (2\pi)^2 \tag{7.206}$$

which we recognize as the rule (7.186) for the modular transformation law! One can argue that the path integral is invariant under this change of coordinates and this will lead to a derivation of the modular transformation law from the physical path integral. But before doing that it is useful to generalize to complex $\beta$.

Now, in statistical physics it is natural to consider the analytic continuation of $Z(\beta)$ to the subset of the complex $\beta$-plane with positive real part. However, in this case, simply analytically continuing $\beta$ to be complex does not fit quite well with the interpretation in terms of the path integral on a torus. Put differently, if we simply took $\beta$ to be complex then the metric (7.205) would be complex. It turns out it is better to split the contributions of left-movers and right-movers in the following way:

We introduce separate Hamiltonians for the left- and right-moving degrees of freedom:

$$\begin{aligned} H_L &= \frac{1}{2}(H+P) \\ H_R &= \frac{1}{2}(H-P) \end{aligned} \tag{7.207}$$

where $P$ is the total momentum of the field $\sim \int_0^{2\pi}\partial_t X \partial_\sigma X$. Note that $H = H_L + H_R$.

[93] If we wish to preserve orientation we should exchange $\sigma^1 \leftrightarrow -\sigma^2$.

Explicitly, one finds: [94]

$$
\begin{aligned}
H_L &= \frac{1}{4}\ell_s\hat{p}^2 + \ell_s^{-1}\left[\sum_{n=1}^{\infty} n a_n^\dagger a_n - \frac{1}{24}\right] \\
H_R &= \frac{1}{4}\ell_s\hat{p}^2 + \ell_s^{-1}\left[\sum_{n=1}^{\infty} n \tilde{a}_n^\dagger \tilde{a}_n - \frac{1}{24}\right]
\end{aligned}
\tag{7.208}
$$

Now we can generalize the statmech partition function to be

$$Z(\beta_L, \beta_R) = \mathrm{Tr}_{\mathcal{H}} e^{-\beta_L H_L} e^{-\beta_R H_R} \tag{7.209}$$

where we can again continue to $Re(\beta_L) > 0$ and $Re(\beta_R) > 0$. If we choose

$$\beta_L = \beta = \beta_R^* \tag{7.210}$$

then the partition function can be written as:

$$Z(\beta) = \mathrm{Tr}_{\mathcal{H}} e^{-Re(\beta)H} e^{-\mathrm{i}Im(\beta)P} \tag{7.211}$$

One of the virtues of (7.211) is that it still has a nice interpretation in terms of a path integral on a torus: After we propagate in Euclidean time by $\mathrm{Re}(\beta\omega)$ we shift in the $\sigma$ coordinate by $\mathrm{Im}(\beta\omega)$ before gluing, because $P$ is the generator of translations in the $\sigma$ direction. The net result is that we can identify $Z(\beta)$ with the pathintegral on a torus with metric:

$$ds^2 = (2\pi)^2|d\sigma^1 + \tau d\sigma^2|^2 = (2\pi)^2|dz|^2 \tag{7.212}$$

where

$$\tau = \mathrm{i}\frac{\beta\omega_0}{2\pi} = \mathrm{i}\frac{\beta}{2\pi\ell_s} \tag{7.213}$$

We can identify the torus with our friend: $E_\tau := \mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})$ with a flat metric $ds^2 = |dz|^2$ where $\tau$ is a complex number in the upper half-plane.

Now we can evaluate the partition function in two ways. First, we can evaluate the trace directly, along the lines we did above and we obtain:

$$Z(\beta) = (R/\ell_s)\cdot(\mathrm{Im}\tau)^{-1/2}|Z^{osc}(\beta)|^2 \tag{7.214}$$

where again $(R/\ell_s)$ is the regularization of the volume divergence.

On the other hand we also have a representation as a Euclidean path integral:

$$Z^{\mathrm{path}}(\beta) = \int_{X:E_\tau\to\mathbb{R}} [dX(\sigma^1,\sigma^2)]e^{-S_E} \tag{7.215}$$

To obtain the action $S_E$ we first return to the original action and restore the $1+1$ dimensional Minkowskian metric:

$$S = -\frac{1}{4\pi\ell_s^2}\int_{S^1\times\mathbb{R}} d^2\sigma\sqrt{|\mathrm{det}\eta|}\eta^{\alpha\beta}\partial_\alpha X\partial_\beta X \tag{7.216}$$

[94] The splitting of the zeromode between left- and right-movers is a very subtle point we have elided here. To do this properly one needs to work out the quantization of a self-dual field.

we choose $\eta_{tt} = -1$. The generalization to an arbitrary metric $h_{\alpha\beta} d\sigma^\alpha d\sigma^\beta$ on the "world-sheet" $S^1 \times \mathbb{R}$ is clear:

$$S = -\frac{1}{4\pi\ell_s^2} \int_{S^1\times\mathbb{R}} d^2\sigma \sqrt{|\det h|} h^{\alpha\beta} \partial_\alpha X \partial_\beta X \tag{7.217}$$

This is the standard minimal coupling in general relativity. It is possible to "Wick rotate" the time coordinate to imaginary time and consider the path integral with $e^{-S_E}$ and Euclidean signature two-dimensional space time with:

$$S_E = -\frac{1}{4\pi\ell_s^2} \int_{S^1\times\mathbb{R}} d^2\sigma \sqrt{\det h} h^{\alpha\beta} \partial_\alpha X \partial_\beta X \tag{7.218}$$

where now $h_{\alpha\beta}$ has Euclidean signature. In (7.215) we use the metric (7.212). Now (7.215) is really just a Gaussian integral.

Recall that in finite dimensions we have

$$\int \prod_{i=1}^{n} dx^i e^{-x^i Q_{ij} x^j} = \frac{\pi^{n/2}}{\sqrt{\det Q}} \tag{7.219}$$

where $Re(Q) > 0$.

At least formally we can take this formula over to definite the infinite-dimensional Gaussian integrals that appear in QFT. Once again the constant mode where $X(\sigma^1, \sigma^2)$ is constant and the action is zero gives us some trouble. Once again, we regularize it by putting the field $X$ in a periodic box of length $R$. Once this is done we get:

$$Z(\beta) = \left(\frac{R}{\ell_s}\right) (\mathrm{Im}\tau)^{-1/2} \frac{1}{\sqrt{\det'(-\Delta_{E_\tau})}} \tag{7.220}$$

The determinant of the Laplacian $\Delta_{E_\tau}$ on scalars is then defined by $\zeta$-function regularization.

In general, if $\mathcal{O}$ is an operator with a discrete spectrum bounded below and growing sufficiently fast then we can define the zeta function of the operator

$$\zeta_{\mathcal{O}}(s) = \sum_\lambda D(\lambda) \lambda^{-s} \tag{7.221}$$

where $D(\lambda)$ is the degeneracy of the $\lambda$ eigenspace. For a good spectrum, like that described above this will converge at large $Re(s)$ and admit an analytic continuation, analytic in the neighborhood of zero. Then we define the determinant of the operator $\mathcal{O}$ to be

$$\det\mathcal{O} := \exp[-\frac{d}{ds}\zeta_{\mathcal{O}}|_{s=0}] \tag{7.222}$$

a formula which, one easily checks, is valid in finite dimensions.

In the case of the Laplacian on the torus with the above metric the spectrum of minus the Laplacian on scalars is $\{(\mathrm{Im}\tau)^{-1}|n_1 + n_2\tau|^2\}_{(n_1,n_2)\in\mathbb{Z}^2}$ and all the eigenspaces are singly degenerate. The $\zeta$-function can be evaluated explicitly (this is a nice exercise in complex analysis) and the resulting determinant is:

$$Z(\beta) = \left(\frac{R}{\ell_s}\right) (\mathrm{Im}\tau)^{-1/2} |\eta(\tau)|^{-2} \tag{7.223}$$

where

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty}(1-q^n), \tag{7.224}$$

In agreement with (7.214).

Now the path integral is invariant under "large" diffeomorphisms of the torus, that is, on diffeomorphisms which are not deformable to the identity. Such diffeomorphisms will act nontrivially on $\pi_1(T^2)$. If we take $z = x + \tau y$ with $x, y$ real and $x, y$ identified modulo one, then we can make the diffeomorphism that rotates by 90 degrees in the $x, y$ plane. Note that this exchanges $A$- and $B$-cycles: We are exchanging the spatial circle with the (Euclidean) time circle. The transformation also takes the torus to a torus with $\tau \to -1/\tau$ and the flat metric rescaled by a constant factor: If $ds^2 = |dz|^2$ with $z = x + \mathrm{i}y$ then the pull-back is $ds^2 = |\tau|^2 |dx' + \tau' dy'|^2$ with $x' = y$ and $y' = -x'$ and $\tau' = -1/\tau$.

What about the overall factor of $|\tau|^2$ in front of the metric? The massless scalar field above has a beautiful property known as *conformal invariance* that is, the action is invariant under *conformal transformations*

$$h_{\alpha\beta} \to \Omega^2 h_{\alpha\beta} \tag{7.225}$$

One has to be very careful about the quantum theory: In this case the partition function is not quite invariant but rather scales by an overall functional of $\Omega$ known as the Liouville action. But for a flat metric and constant $\Omega$, the overall scaling factor is just one.

We finally conclude that $(\mathrm{Im}\tau)^{-1/2}|\eta(\tau)|^{-2}$ is invariant under $\tau \to -1/\tau$ and since $\eta(\tau)$ is holomorphic one can deduce

$$\eta(-1/\tau) = \Phi(-\mathrm{i}\tau)^{1/2}\eta(\tau) \tag{7.226}$$

where $\Phi$ is a phase independent. Because it is a phase and because the other factors are (locally) holomorphic it must be a constant independent of $\tau$. Substituting $\tau = \mathrm{i}$ we see the phase $\Phi = 1$. In this way we have derived the very important result (7.190) above. Actually, we have found much more: If

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z}) \tag{7.227}$$

then our reasoning shows that

$$\eta(\gamma\cdot\tau) = \eta(\frac{a\tau+b}{c\tau+d}) = \Phi(\gamma)(-\mathrm{i}(c\tau+d))^{1/2}\eta(\tau) \tag{7.228}$$

where $\Phi(\gamma)$ is a phase. Because the modular group is generated by $\tau \to \tau+1$ and $\tau \to -1/\tau$ it follows that $\Phi(\gamma)$ is in fact a $24^{th}$ root of unity which depends on $\gamma$. There is a (subtle) explicit expression for this phase to be found in textbooks on analytic number theory.

As an interesting application of (7.190), when combined with the method of stationary phase, one can derive the Hardy-Ramanujan formula giving an asymptotic formula for large values of $n$:

$$p(n) \sim \frac{1}{\sqrt{2}} (\frac{1}{24})^{3/4} n^{-1} \mathrm{exp}\left(2\pi\sqrt{\frac{n}{6}}\right) \tag{7.229}$$

Note that this grows much more slowly than the order of the group, $n!$. So we conclude that some conjugacy classes must be very large! (See discussion in the next section on the class equation if this is not obvious.)

Analogs of equation (7.229) for a class of functions known as *modular forms* plays an important role in modern discussions of the entropy of supersymmetric (and extreme) black hole solutions of supergravity.

---

**Exercise** *Change Of Form Of Hamiltonian Under Bogoliubov Transformation*
Show that the standard harmonic oscillator Hamiltonian

$$H_b = \omega(b^\dagger b + \frac{1}{2}) \tag{7.230}$$

when $b, b^\dagger$ are related to $a, a^\dagger$ by a Bogoliubov transformation (7.164)

$$b = \cosh(r)a - e^{-\mathrm{i}\phi}\sinh(r)a^\dagger \tag{7.231}$$

becomes

$$H_b = \Omega a^\dagger a + \Delta(a^\dagger)^2 + \Delta^* a^2 + \Gamma \tag{7.232}$$

with

$$\begin{aligned} \Omega &= (\cosh(2r))\omega \\ \Delta &= -\frac{1}{2}e^{-\mathrm{i}\phi}\sinh(2r) \\ \Gamma &= \frac{1}{2}\cosh(2r)\omega \end{aligned} \tag{7.233}$$

**Remark**: In the Bogoliubov-deGennes effective Hamiltonian description of superconductivity $\Delta$ is the value of the Cooper pair condensate.

---

---

**Exercise** *Vacuum After Bogoliubov Transformation*
a.) If we represent $\hat{p}, \hat{q}$ on the Hilbert space $L^2(\mathbb{R})$ with

$$\begin{aligned} (\hat{q}\cdot\psi)(x) &= x\psi(x) \\ (\hat{p}\cdot\psi)(x) &= -\mathrm{i}\hbar\frac{\partial}{\partial x}\psi(x) \end{aligned} \tag{7.234}$$

then the groundstate with $a|0\rangle_a = 0$ corresponds - up to a phase - to a vector $\psi \in L^2(\mathbb{R})$ defined by the differential equation:

$$(\mathrm{i}\hat{p} + \omega\hat{q})\psi = 0 \qquad \Rightarrow \qquad \psi(x) = Ce^{-\frac{\omega x^2}{2\hbar}} \tag{7.235}$$

b.) Show that a normalized wavefunction annihilated by $b$ defined by

$$b = \cosh(r)a - e^{-\mathrm{i}\phi}\sinh(r)a^{\dagger} \tag{7.236}$$

is - up to a phase -

$$\langle x|0\rangle_b = \pi^{-1/4}\frac{1}{\sqrt{\cosh(r) + e^{\mathrm{i}\phi}\sinh(r)}}\exp(-\frac{1}{2}\Lambda x^2) \tag{7.237}$$

where [95]

$$\Lambda = \frac{\cosh(r) - e^{\mathrm{i}\phi}\sinh(r)}{\cosh(r) + e^{\mathrm{i}\phi}\sinh(r)} \tag{7.238}$$

♣Incomplete. Need to explain the relation to the squeezed state formula quoted above. ♣

---

---

**Exercise** *Deriving the Hardy-Ramanujan formula*

The function $Z^{osc}(\beta)$ has a nice analytic continuation into the right half complex plane where $\mathrm{Re}(\beta) > 0$. Note that $q^{1/24}Z^{osc}(\beta)$ is periodic under imaginary shifts $\beta \to \beta + \frac{2\pi\mathrm{i}}{\omega}$.

Write

$$p(n) = \int_{\beta_0}^{\beta_0+\frac{2\pi\mathrm{i}}{\omega}} d\beta e^{-n\beta\omega}q^{1/24}Z^{osc}(\beta) \tag{7.239}$$

and use the above transformation formula, together with the stationary phase method to derive (7.229). [96]

---

### 7.5.2 Conjugacy Classes In $S_n$ And Partitions

♣Probably this section should be moved to the representation theory of $S_n$ discussion. ♣

Above, we defined a *partition of* $n$ for a positive integer $n$ as a way of writing it as a sum of positive integers. There is another viewpoint based on a very closely related concept called, simply, a *partition.*

By definition, a *partition* is a sequence of nonnegative integers $\lambda := \{\lambda_1, \lambda_2, \lambda_3, \dots\}$ so that

a.) $\lambda_i$ are nonincreasing: $\lambda_i \geq \lambda_{i+1}$.

b.) The $\lambda_i$ eventually become zero.

The nonzero $\lambda_i$ are called the *parts of the partition.*

Given a partition, we define $|\lambda| := \sum_i \lambda_i$. If $n = |\lambda_i|$ then we get a partition of $n$ from:

$$n = \lambda_1 + \lambda_2 + \cdots + \lambda_k \tag{7.240}$$

We say we have a partition of $n$ with $k$ parts.

---

[95] *Hint*: Convert $b\psi = 0$ to a differential equation by substituting for $a, \bar{a}$ in terms of $\hat{p}$ and $\hat{q}$.

[96] *Answer*: Write $p(n) = \int_{-1/2+\mathrm{i}\beta}^{1/2+\mathrm{i}\beta} e^{-2\pi\mathrm{i}n\tau}q^{1/24}\eta(\tau)^{-1}d\tau$ where $\tau = x + \mathrm{i}\beta$ and the contour is along a horizontal line. One argues that, as $\beta \to 0^+$ the dominant terms in the integral come from the region near $x \cong 0$. (This is a rather subtle step to do correctly.) Then, using the modular transformation law one writes $\eta(\tau)^{-1} = (-\mathrm{i}\tau)^{1/2}\eta(-1/\tau)^{-1}$ and for $\mathrm{Im}(-1/\tau) \to \infty$ one approximates $\eta(-1/\tau)^{-1} \cong \exp[2\pi\mathrm{i}/(24\tau)]$. Now one applies the standard stationary phase technique. When this procedure is carried out more systematically one is led to the famous Rademacher expansion for coefficients of certain modular functions.

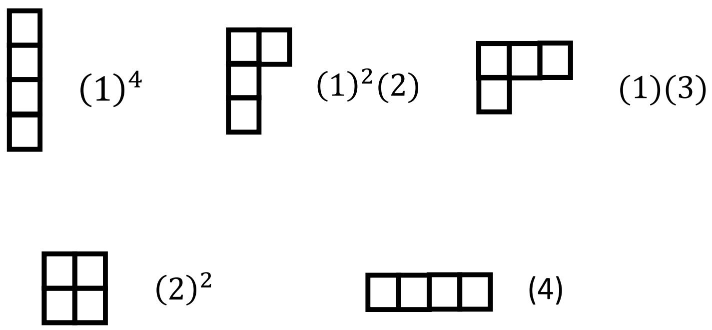



**Figure 4:** Young diagrams corresponding to the 5 different partitions of 4.

Partitions can be very effectively visualized by a diagram known as a *Young diagram* (a.k.a. a Young frame or a Ferrers diagram). This is a diagram with $\lambda_1$ boxes in the first row, $\lambda_2$ boxes in the second row and so forth. The boxes are arranged to make, roughly speaking, an upside-down $L$-shape. See Figure 4 for some examples. We will denote the diagram as $Y(\lambda)$. We will talk much more about these when discussing representations of the symmetric group and representations of $SU(n)$.

One way to associate a conjugacy class in $S_n$ with a partition is as follows. We let $m_i(\lambda)$ denote the number of rows in $Y(\lambda)$ with $i$ boxes. One can write a formula for it as a function of $\lambda$:

$$m_i(\lambda) := |\{j|\lambda_j = i\}| \tag{7.241}$$

The quantity $m_i(\lambda)$ is called the *multiplicity of $i$ in $\lambda$*. We can then assign a conjugacy class in $S_n$ with $n = |\lambda|$ :

$$C(\lambda) := (1)^{m_1}(2)^{m_2}\cdots \tag{7.242}$$

One of the beauties of the diagrammatic representation is that a certain kind of duality symmetry emerges which is not so obvious from the equations alone. It is worth noting that there is another associated partition and conjugacy class known as the conjugate partition. We let $\lambda'$ denote the partition conjugate to $\lambda$. It can be defined by saying that $Y(\lambda')$ is obtained from $Y(\lambda)$ by flipping on the main diagonal, i.e. we exchange rows and columns. So, the number of boxes in the $i^{th}$ row of $Y(\lambda')$, $\lambda'_i$ is exactly the number boxes in the $i^{th}$ column of $Y(\lambda)$. Because the diagram $Y(\lambda)$ has an inverted $L$-shape so does the conjugate diagram $Y(\lambda')$, in other words:

$$\lambda'_1 \geq \lambda'_2 \geq \cdots \tag{7.243}$$

so $\lambda' = \{\lambda'_1, \lambda'_2, \dots\}$ is another partition. Clearly the total number of boxes is unchanges so $|\lambda'| = n$. Note that $\lambda'' = \lambda$.

A little thought makes it clear that we have the relation:

$$\lambda'_i = |\{j|\lambda_j \geq i\}| \tag{7.244}$$

For example, $\lambda'_1$ is the number of boxes in the first column of $Y(\lambda)$. This is clearly the number of rows, and to have a nontrivial row we must have $\lambda_j \geq 1$. Now, to get a box in the second column we must have rows with $\lambda_j \geq 2$, and so on.

It follows from (7.244) that

$$m_i(\lambda) = \lambda'_i - \lambda'_{i+1} \tag{7.245}$$

So in terms of $\lambda'$ our conjugacy class (7.242) above becomes

$$C(\lambda) = (1)^{\lambda'_1 - \lambda'_2} (2)^{\lambda'_2 - \lambda'_3} \cdots (n)^{\lambda'_n} \tag{7.246}$$

But then it is clear that we could also assign a conjugacy class

$$C'(\lambda) = (1)^{\lambda_1 - \lambda_2} (2)^{\lambda_2 - \lambda_3} \cdots (n)^{\lambda_n} \tag{7.247}$$

It is perhaps less obvious that (**??**) defines a conjugacy class in $S_n$. One way to check this is to note the identity:

$$\lambda'_1 + \lambda'_2 + \cdots + \lambda'_n = (\lambda'_1 - \lambda'_2) + 2(\lambda'_2 - \lambda'_3) + 3(\lambda'_3 - \lambda'_4) + \cdots + (n-1)(\lambda'_{n-1} - \lambda'_n) + n\lambda'_n \tag{7.248}$$

assures us the total number of boxes is still $n$.

For any $n$ we can consider the set of all partitions. It is a finite set and we can consider it as a probability space giving equal weight to all partitions. When $n$ is large we can ask what the "typical" partition is in this measure space. That is, what are the "typical" conjugacy classes in $S_n$ when $n$ is large? This is an imprecise, and rather subtle question. To get some sense of an answer it is useful to consider the number $p_k(n)$ of partitions of $n$ into precisely $k$ parts (as in (7.240)). The generating function is (see exercise below)

$$\sum_{n=1}^{\infty} p_k(n) x^n = \prod_{j=1}^{k} \frac{x}{1 - x^j} \tag{7.249}$$

One natural guess, then, is that the "typical" partition has $k \cong \sqrt{n}$ with "most of the parts" on the order of $\sqrt{n}$. This naive picture can be considerably improved using the statistical theory of partitions. [97] Without going into a lot of complicated asymptotic formulae, the main upshot is that, for large $n$, as a function of $k$, $p_k(n)$ indeed is sharply peaked with a maximum around

$$\bar{k}(n) := \frac{\sqrt{6}}{2\pi} \sqrt{n} \log n \tag{7.250}$$

[97] It is a large subject. See P Erdös and J. Lehner, "The distribution of the number of summands in the partition of a positive integer," Duke Math. Journal **8**(1941)335-345 or M. Szalay and P. Turán, "On some problems of the statistical theory of partitions with application to characters of the symmetric group. I," Acta Math. Acad. Scient. Hungaricae, Vol. 29 (1977), pp. 361-379.

See Figure 5 for a numerical illustration. Moreover, and again speaking very roughly, the number of terms in the partition $\lambda_j$ with $\lambda_j \cong \frac{\sqrt{6}}{2\pi}\sqrt{n}$ is order $\sqrt{6n}/\pi$.

**Remarks**:

1. Recall that in our discussion of a string, or equivalently of a massless scalar field on the circle, there are $p(n)$ states in the energy eigenspace with energy $E = (n - \frac{1}{24})\omega$. Thus we can interpret the above result as a kind of equipartition theorem: The most likely state is the one where the energy is shared equally by the different oscillators.

2. A *Young tableau* is a Young diagram with $n$ boxes where the boxes have been filled in with integers drawn from $\{1, \dots, n\}$ so that no integer is repeated. Note that the symmetric group $S_n$ acts on Young tableau. For a given tableau $\mathcal{T}$ we can define two subgroups of the symmetric group: $R(\mathcal{T})$ are the permutations that only move numbers around within the rows and $C(\mathcal{T})$ are the permutations that only move numbers around within the columns. Using these subgroups one can construct projection operators that are used in constructing the irreducible representations of the symmetric group. To do this we note that if $V$ is any vector space then there is a canonical representation $\varphi$ of $S_n$ on $V^{\otimes n}$ by permuting factors. Given a Young tableau $\mathcal{T}$ we define
$$R := \sum_{r \in R(\mathcal{T})} \varphi(r) \tag{7.251}$$
$$C := \sum_{c \in C(\mathcal{T})} \epsilon(c)\varphi(c) \tag{7.252}$$
and $P = RC$ is proportional to a projection operator onto a subrepresentation of the symmetric group. Indeed note that, rather trivially: $\varphi(r)RC = RC$ for $r \in R(\mathcal{T})$ and $RC\varphi(c) = \epsilon(c)RC$. Somewhat less trivially, any operator of the form $\sum_{\sigma \in S_n} n(\sigma)\varphi(\sigma)$ that satisfies this property must be proportional to $RC$. Since $(RC)^2$ clearly satisfies the property, it must be proportional to $RC$. See section **** below for more details.

---

**Exercise** *Conjugate Partition*
Show that if
$$\lambda = \{5, 4, 3, 2, 2, 2, 1, 1\} \tag{7.253}$$
then
$$\lambda' = \{8, 6, 3, 2, 1\} \tag{7.254}$$

---

---

**Exercise** *Young tableau*

a.) For a Young tableau $\mathcal{T}$ show that

$$R(\mathcal{T}) \cong S_{\lambda_1} \times S_{\lambda_2} \times \cdots \tag{7.255}$$

$$C(\mathcal{T}) \cong S_{\lambda'_1} \times S_{\lambda'_2} \times \cdots \tag{7.256}$$

Note that the isomorphism class only depends on the partition $\lambda$ and not on the particular choice of tableau for that partition.

b.) Show that $R(\mathcal{T}) \cap C(\mathcal{T}) = \{1\}$.

---

---

**Exercise** *Generating Function For* $p_k(n)$

Prove equation (7.249). [98]

---

---

## 8. More About Group Actions And Orbits

♣NOTE BENE! THE MATERIAL IN THIS SECTION IS IDENTICAL TO SECTION 2 OF CHAPTER 3 ♣

In Section 5.1 above we introduced the notion of a group action on a set. In this section we develop this important idea a bit further.

### 8.1 Left And Right Group Actions

Let X be any set (possibly infinite). Recall the definition we gave in Section 5.1.

A *permutation* of X is a 1-1 and onto mapping $X \to X$. The set $S_X$ of all permutations forms a group under composition. A *transformation group* on X is a subgroup of $S_X$.

Equivalently, a *G-action on a set* $X$ is a map $\phi : G \times X \to X$ compatible with the group multiplication law as follows:

A *left-action* satisfies:

$$\phi(g_1, \phi(g_2, x)) = \phi(g_1 g_2, x) \tag{8.1}$$

A *right-action* satisfies

$$\phi(g_1, \phi(g_2, x)) = \phi(g_2 g_1, x) \tag{8.2}$$

In addition in both cases we require that

$$\phi(1_G, x) = x \tag{8.3}$$

[98] *Answer*: A partition of $n$ into exactly $k$-parts means that $\lambda_k \geq 1$. So now write

$$n - k = (\lambda_1 - \lambda_2) + 2(\lambda_2 - \lambda_3) + \cdots + (k-1)(\lambda_{k-1} - \lambda_k) + k(\lambda_k - 1)$$

This is a partition of $n - k$ as a sum of integers drawn from $\{1, \ldots, k\}$. Enumerating those is clearly given by $\prod_{j=1}^{k}(1 - x^j)^{-1}$.

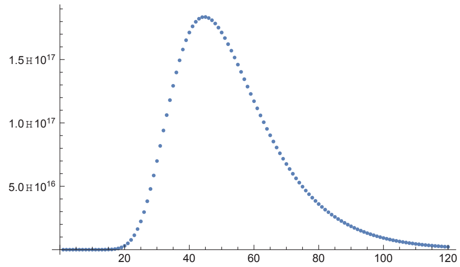



**Figure 5:** Showing the distribution of $p_k(n)$ as a function of $k$ for $n = 400$ and $1 \le k \le 120$. Note that the Erdös-Lehner mean value of $k$ is $\bar{k} = \frac{\sqrt{6}}{2\pi} 20 \log(20) \cong 46.7153$ is a very good approximation to where the distribution has its sharp peak. The actual maximum is at $k = 45$.

for all $x \in X$.

**Remarks**:

1. If $\phi$ is a left-action then it is natural to write $g \cdot x$ for $\phi(g,x)$. In that case we have

$$g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x. \tag{8.4}$$

   Similarly, if $\phi$ is a right-action then it is better to use the notation $\phi(g,x) = x \cdot g$ so that

$$(x \cdot g_2) \cdot g_1 = x \cdot (g_2 g_1). \tag{8.5}$$

2. If $\phi$ is a left-action then $\tilde{\phi}(g,x) := \phi(g^{-1},x)$ is a right-action, and vice versa. Thus there is no essential difference between a left- and right-action. However, in computations with nonabelian groups it is extremely important to be consistent and careful about which choice one makes. A common source of error is a confusion of a left-action with a right-action.

3. If $G$ is an Abelian group then any left-action is simultaneously a right-action.

4. A given set $X$ can admit more than one action by the same group $G$. If one is working simultaneously with several different $G$ actions on the same set then the notation $g \cdot x$ is ambiguous and one should write, for example, $\phi_g(x) = \phi(g,x)$ or speak of $\phi_g$, etc. A good example of a set $X$ with several natural $G$ actions is the case of $X = G$ itself. Then there are the actions of left-multiplication, right-multiplication, and conjugation. The action of $g$ on the group element $g'$ is:

$$\begin{aligned} L(g,g') &= gg' \\ \tilde{L}(g,g') &= g^{-1}g' \\ R(g,g') &= g'g \\ \tilde{R}(g,g') &= g'g^{-1} \\ C(g,g') &= g^{-1}g'g \\ \tilde{C}(g,g') &= gg'g^{-1} \end{aligned} \tag{8.6}$$

   where on the RHS of these equations we use group multiplication. The reader should work out which actions are left actions and which actions are right actions.

---

**Exercise** *A Funny Transformation*
Consider the map $\varphi : G \times G \to G$ defined by

$$\varphi(g,g') = g^{-1}g'g^{-1} \tag{8.7}$$

Is this a right-action or a left-action of $G$ on $X = G$ ? [99]

---

### 8.1.1 More About Induced Group Actions On Function Spaces

♣This subsection is now orphaned and should be moved to the beginning of this section on More about group actions. ♣

Let us return to the considerations of section 5.5.

Let $X$ be a $G$-set and let $Y$ be any set. There are natural left- and right- actions on the function space $\mathrm{Map}(X,Y)$. Given $\Psi \in \mathrm{Map}(X,Y)$ and $g \in G$ we need to produce a new function $\phi(g,\Psi) \in \mathrm{Map}(X,Y)$. The rules are as follows:

1. If $G$ is a left-action on $X$ then

$$\phi(g,\Psi)(x) := \Psi(g\cdot x) \qquad \text{right action on } \mathrm{Map}(X,Y) \tag{8.8}$$

2. If $G$ is a left-action on $X$ then

$$\phi(g,\Psi)(x) := \Psi(g^{-1}\cdot x) \qquad \text{left action on } \mathrm{Map}(X,Y) \tag{8.9}$$

3. If $G$ is a right-action on $X$ then

$$\phi(g,\Psi)(x) := \Psi(x\cdot g) \qquad \text{left action on } \mathrm{Map}(X,Y) \tag{8.10}$$

4. If $G$ is a right-action on $X$ then

$$\phi(g,\Psi)(x) := \Psi(x\cdot g^{-1}) \qquad \text{right action on } \mathrm{Map}(X,Y) \tag{8.11}$$

**Example**: Consider a spacetime $\mathcal{S}$. With suitable analytic restrictions the space of scalar fields on $\mathcal{S}$ is $\mathrm{Map}(\mathcal{S},\kappa)$, where $\kappa = \mathbb{R}$ or $\mathbb{C}$ for real or complex scalar fields. If a group $G$ acts on the spacetime, there is automatically an induced action on the space of scalar fields. To be even specific, suppose $X = \mathbb{M}^{1,d-1}$ is $d$-dimensional Minkowski space time, $G$ is the Poincaré group, and $Y = \mathbb{R}$. Given one scalar field $\Psi$ and a Poincaré transformation $g^{-1}\cdot x = \Lambda x + v$ we have $(g\cdot\Psi)(x) = \Psi(\Lambda x + v)$.

Similarly, suppose that $X$ is any set, but now $Y$ is a $G$-set. Then again there is a $G$-action on $\mathrm{Map}(X,Y)$:

$$(g\cdot\Psi)(x) := g\cdot\Psi(x) \qquad \text{or} \qquad \Psi(x)\cdot g \tag{8.12}$$

according to whether the $G$ action on $Y$ is a left- or a right-action, respectively. These are left- or right-actions, respectively.

We can now combine these two observations and get the general statement: We assume that both $X$ is a $G_1$-set and $Y$ is a $G_2$-set. We can assume, without loss of generality, that

[99] *Answer*: In general it satisfies neither the left-action condition nor the right-action condition, so it is neither. But if $G$ is Abelian it defines a left action and a right-action.

we have left-actions on both $X$ and $Y$. Then there is a natural $G_1 \times G_2$-action on $\mathrm{Map}(X,Y)$ defined by:

$$\phi((g_1, g_2), \Psi)(x) := g_2 \cdot (\Psi(g_1^{-1} \cdot x)) \tag{8.13}$$

note that if one writes instead $g_2 \cdot (\Psi(g_1 \cdot x))$ on the RHS then we do not have a well-defined $G_1 \times G_2$-action (if $G_1$ and $G_2$ are both nonabelian). In most applications $X$ and $Y$ both have a $G$ action for a single group and we write

$$\phi(g, \Psi)(x) := g \cdot (\Psi(g^{-1} \cdot x)) \tag{8.14}$$

This is a special case of the general action (8.13), with $G_1 = G_2 = G$ and specialized to the diagonal $\Delta \subset G \times G$.

**Example**: Again let $X = \mathbb{M}^{1,d-1}$ be a Minkowski space time. Take $G_1 = G_2$ and let $G = \Delta \subset G \times G$ be the diagonal subgroup, and take $G$ to be the Poincaré group. Now let $Y = V$ be a finite-dimensional representation of the Poincaré group. Let us denote the action of $g \in G$ on $V$ by $\rho(g)$. Then a field $\Psi \in \mathrm{Map}(X,Y)$ has an action of the Poincaré group defined by

$$g \cdot \Psi(x) := \rho(g)\Psi(g^{-1}x) \tag{8.15}$$

This is the standard way that fields with nonzero "spin" transform under the Poincaré group in field theory. As a very concrete related example, consider the transformation of electron wavefunctions in nonrelativistic quantum mechanics. The electron wavefunction is governed by a two-component function on $\mathbb{R}^3$:

$$\Psi(\vec{x}) = \begin{pmatrix} \psi_+(\vec{x}) \\ \psi_-(\vec{x}) \end{pmatrix} \tag{8.16}$$

Then, suppose $G = SU(2)$. We use, once again, the surjective homomorphism $\pi : G \to SO(3)$ defined by $\pi(u) = R$ where

$$u\vec{x} \cdot \vec{\sigma} u^{-1} = (R\vec{x}) \cdot \vec{\sigma} \tag{8.17}$$

Then the (double-cover) of the rotation group acts to define the transformed electron wavefunction $u \cdot \Psi$ by

$$(u \cdot \Psi)(\vec{x}) := u \begin{pmatrix} \psi_+(R^{-1}\vec{x}) \\ \psi_-(R^{-1}\vec{x}) \end{pmatrix} \tag{8.18}$$

In particular, $u = -1$ acts trivially on $\vec{x}$ but nontrivially on the wavefunction.

## 8.2 Some Definitions And Terminology Associated With Group Actions

There is some important terminology one should master when working with $G$-actions. First here are some terms used when describing a $G$-action on a set $X$:

**Definitions**:

1. A group action is *effective* or *faithful* if for any $g \neq 1$ there is some $x$ such that $g \cdot x \neq x$. Equivalently, the only $g \in G$ such that $\phi_g$ is the identity transformation is $g = 1_G$. A group action is *ineffective* if there is some $g \in G$ with $g \neq 1$ so that $g \cdot x = x$ for all $x \in X$. Note that the set of $g \in G$ that act ineffectively is a normal subgroup of $G$.

2. A group action is *transitive* if for any pair $x, y \in X$ there is some $g$ with $y = g \cdot x$.

3. A group action is *free* if for any $g \neq 1$ then for *every* $x$, we have $g \cdot x \neq x$.

In summary:

1. *Effective*: $\forall g \neq 1, \exists x$ s.t. $g \cdot x \neq x$.

2. *Ineffective*: $\exists g \neq 1$,s.t. $\forall x \; g \cdot x = x$.

3. *Transitive*: $\forall x, y \in X, \exists g$ s.t. $y = g \cdot x$.

4. *Free*: $\forall g \neq 1, \forall x, \; g \cdot x \neq x$

In addition there are some further important definitions:

1. Given a point $x \in X$ the set of group elements:

$$\mathrm{Stab}_G(x) := \{g \in G : g \cdot x = x\} \subset G \tag{8.19}$$

is called the *isotropy group at $x$*. It is also called the *stabilizer group* of $x$. It is often denoted $G^x$. The reader should show that $G^x$ is in fact a subgroup of $G$. Note that a group action is free iff for every $x \in X$ the stabilizer group $G^x$ is the trivial subgroup $\{1_G\}$. Another notation one will find is $\mathrm{Aut}(x)$, because in the category formed from a group action on $X$ this is the automorphism group of the object $x$.

2. A point $x \in X$ is a *fixed point* of the $G$-action if there exists some element $g \in G$ with $g \neq 1$ such that $g \cdot x = x$. So, a point $x \in X$ is a fixed point of $G$ iff $\mathrm{Stab}_G(x)$ is not the trivial group. Some caution is needed here because if an author says "$x$ is a fixed point of $G$" the author might mean that $\mathrm{Stab}_G(x) = G$. The definition we just gave above does not have that implication.

3. Given a group element $g \in G$ the *fixed point set* of $g$ is the set

$$\mathrm{Fix}_X(g) := \{x \in X : g \cdot x = x\} \subset X \tag{8.20}$$

The fixed point set of $g$ is often denoted by $X^g$. Note that if the group action is free then for every $g \neq 1$ the set $\mathrm{Fix}_X(g)$ is the empty set.

4. We repeat the definition from section **** above. The *orbit of $G$ through a point $x$* is the set of points $y \in X$ which can be reached by the action of $G$:

$$O_G(x) = \{y : \exists g \quad \text{such that} \quad y = g \cdot x\} \subset X \tag{8.21}$$

5. A *homogeneous space* is a set $X$ with a transitive group action. As we have mentioned a few times, the set of cosets of a subgroup $G/H$ are examples of homogeneous spaces.

**Remarks**:

♣First remark here is redundant with things said several times above. ♣

1. If we have a $G$-action on $X$ then we can define an equivalence relation on $X$ by defining $x \sim y$ if there is a $g \in G$ such that $y = g \cdot x$. (Check this is an equivalence relation!) The orbits of $G$ are then exactly the equivalence classes of under this equivalence relation. Therefore, $X$ is partitioned into a disjoint union of all the $G$-orbits.

2. The group action restricts to a transitive group action on any orbit.

3. If $x, y$ are in the same orbit then the isotropy groups $G^x$ and $G^y$ are conjugate subgroups in $G$. Therefore, to a given orbit, we can assign a definite *conjugacy class* of subgroups.

4. We can now explain the notation $Aut(b)$ used in equation (3.30). We consider the set of symmetric or antisymmetric bilinear forms

$$\mathcal{B}^\xi = \{b \in M_n(\kappa) : b^{tr} = \xi b\} \tag{8.22}$$

where $\xi = +$ for symmetric and $\xi = -$ for antisymmetric bilinear forms. The left $G = GL(n, \kappa)$ action on $\mathcal{B}^\xi$ is

$$\phi(g, b) := gbg^{tr} \ . \tag{8.23}$$

The group $Aut(b)$ is the stabilizer group of a given form. It is the group of "automorphisms" of that form.

5. *Stabilizer Groups And Quantum Information Theory.* Recall the definition of the Pauli group $P$ above. Note that if we have a chain of $N$ spin 1/2 particles, then the $N^{th}$ direct product

$$P^N = \underbrace{P \times \cdots \times P}_{N \text{ times}} \tag{8.24}$$

acts naturally on this chain of particles in the sense that it acts on the Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$ associated with a chain of $N$ Qbits. This group is useful in quantum information theory. For example if $S \subset P^N$ is a subgroup then we can study the subset of Hilbert space $\mathcal{H}^S = \{\psi | g\psi = \psi, \quad \forall g \in S\}$. This is the common fixed point set for the entire group. Since the group acts linearly it is a sub-Hilbert space: $\mathcal{H}^S \subset \mathcal{H} \cong (\mathbb{C}^2)^{\otimes N}$. For astutely chosen subgroups these are useful quantum code subspaces, known as *stabilizer codes.* Note that if the subgroup $S$ contains an operator of the form $z1^{\otimes n}$ with $z \neq 1$ then $\mathcal{H}^S = \{0\}$ so the space is rather trivial. For the Pauli group $P^N$ the only operators proportional to the identity are of the form $\pm i1^{\otimes N}$ and $-1^{\otimes N}$. Of course if $\pm i1^{\otimes N} \in S$ then its square $-1^{\otimes N} \in S$. Therefore if we want a nonzero subspace we should require that $-1^{\otimes N}$ is not in $S$.

In quantum information theory coded states are in code subspaces such as $\mathcal{H}^S$. If $\{E_i\}$ are a set of unitary gates considered as "error operators" then the error can be corrected - i.e. there is a quantum channel that undoes the error and restores the state to the code subspace if

$$PE_j^\dagger E_k P = \alpha_{jk} P \tag{8.25}$$

for an Hermitian matrix $\alpha_{jk}$ of scalars, where $P$ is the orthogonal projector $P : \mathcal{H} \to \mathcal{H}^S$. [100]

For stabilizer codes this condition can be expressed group theoretically: If the errors $E_j$ come from the Pauli group then $E_j^\dagger E_k$ should not be in $N(S)-S$ where $N(S)$ is the normalizer subgroup of $S$ in $P^N$. The reason is clear: If $E_j^\dagger E_k \in S$ then $PE_j^\dagger E_k P = P$. But if $E_j^\dagger E_k \in P^N - N(S)$ and $-1$ is not in $S$ then $E_j^\dagger E_k \in P^N - Z(S)$. But in the Pauli group every pair of group elements either commutes or anticommutes. So there is some element $s_0 \in S$ with $E_j^\dagger E_k s_0 = -s_0 E_j^\dagger E_k$. Since $-1$ is not in $S$, we have $s_0^2 = 1$ The projection operator is proportional to $\sum_{s\in S} s = \frac{1}{2}\sum_{s\in S}(s + ss_0)$. But now it is easy to show that

$$\left(\sum_{s\in S}(s+ss_0)\right) E_j^\dagger E_k \left(\sum_{s\in S}(s+ss_0)\right) = 0 \tag{8.29}$$

For more information on the above remarks see the textbook: Nielsen and Chuang, *Quantum Computation and Quantum Information*, chapter 10, especially Theorem 10.1 and section 10.5.

Point 3 above motivates the

**Definition** If $G$ acts on $X$ a *stratum* is a set of $G$-orbits such that the conjugacy class of the stabilizer groups is the same. The set of strata is sometimes denoted $X \parallel G$.

---

[100] To amplify on this: Since $\alpha_{jk}$ is unitarily diagonalizable the algebra of operators generated by $E_j$ is the same as an algebra generated by operators $F_j$ such that $PF_j^\dagger F_k P = d_k \delta_{j,k} P$. So WLOG we assume $\alpha_{j,k} = d_k\delta_{j,k}$ is diagonal. The error in sending a message is the quantum channel

$$\rho \mapsto \mathcal{E}(\rho) = \sum_k E_k^\dagger \rho E_k \tag{8.26}$$

Now note the polar decomposition: $E_j P = \sqrt{d_j} U_j P$ where $U_j$ is unitary. One then defines the Hermitian projection operators $P_j := U_j P U_j^\dagger$ and it follows from the fact that $\alpha_{j,k}$ is diagonal that $P_j P_k = \delta_{j,k} P_j$ are orthogonal projection operators. The recovery operation is the quantum channel

$$\rho \mapsto \mathcal{R}(\rho) = \sum_k U_k^\dagger P_k \rho P_k U_k \tag{8.27}$$

This is a recovery because, as is easily shown, if $P\rho P = \rho$, i.e. if the initial state is in the code subspace then

$$\mathcal{R}(\mathcal{E}(\rho)) = (\sum_k d_k)\rho \tag{8.28}$$

The overall normalization of the density matrix does not affect physical measurements.

**Exercise** *Group actions of $G$ on $G$*

Referring to equation (8.6). Which actions are left-actions and which actions are right-actions? [101]

---

---

**Exercise** *Effectivity Of Group Actions*

Suppose $X$ is a $G$-set.

Recall that a group action of $G$ on $X$ can be viewed as a homomorphism $\Phi : G \to S_X$.

a.) Show that the action is effective iff the homomorphism is injective.

b.) Show that the subset of group elements that act ineffectively is a normal subgroup $H \triangleleft G$.

c.) Show that there is an effective action of the group $G/H$ on $X$.

---

---

**Exercise**

Let $G$ act on a set $X$.

a.) Show that the stabilizer group at $x$, denoted $G^x$ above, is in fact, a subgroup of $G$.

b.) Show that the $G$ action is free iff the stabilizer group at every $x \in X$ is the trivial subgroup $\{1_G\}$.

c.) Suppose that $y = g \cdot x$. Show that $G^y$ and $G^x$ are conjugate subgroups in $G$. [102]

---

---

**Exercise** *Derangements*

A permutation in $S_n$ which acts on $\{1, \dots, n\}$ without fixed points is called a *derangement*. Show that the number of derangements in $S_n$ is given by

$$D_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!} \tag{8.30}$$

---

---

**Exercise** *Normalizers And Centralizers In The Pauli Group*

Let $S \subset P^N$ be a subgroup such that $-1$ is not in $S$. Then the normalizer is the same as the centralizer: $N(S) = Z(S)$. [103]

[101] *Answer*: $L, \tilde{R}, \tilde{C}$ are left-actions, while $\tilde{L}, R, C$ are right-actions.

[102] *Answer*: If $y = g_0 \cdot x$ and $g \cdot x = x$ then $(g_0 g g_0^{-1}) \cdot y = y$ so $g_0 G^x g_0^{-1} = G^{g_0 \cdot x} = G^y$.

[103] *Answer*: Every pair of elements of the Pauli group either commutes or anti-commutes. If $g \in N(S)$ and $s \in S$ then $gsg^{-1} = \pm s$. If $g$ is not in the centralizer then $gsg^{-1} = -s$ but then $gsg^{-1}s^{-1} = -1 \in S$, a contradiction.

**Exercise** *Simple Stabilizer Codes*

In quantum information theory the Pauli operators - thought of as quantum gates - are usually denoted:

$$X := \sigma^1 \qquad Y = \sigma^2 \qquad Z = \sigma^3 \tag{8.31}$$

a.) Consider the stabilizer code on a system of $N$ Qbits generated by $Z_i$, $i = 1, \dots, N$. Show that $\mathcal{H}^S$ is a one-dimensional line through $|0\rangle^{\otimes N}$.

b.) Consider the subgroup of $P^3$:

$$S = \{1, Z_1Z_2, Z_2Z_3, Z_1Z_3\} \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \tag{8.32}$$

Show that $\mathcal{H}^S$ is two-dimensional and generated by $|000\rangle$ and $|111\rangle$.

---

**Exercise** *Equivariant Maps As Fixed Points*

Suppose that $X$ and $Y$ are $G$-sets. As we have seen $Map(X,Y)$ is a $G$-set under the action (8.14). Show that the fixed points for this $G$-action on $Map(X,Y)$ are precisely the $G$-equivariant maps from $X$ to $Y$.

---

### 8.3 The Stabilizer-Orbit Theorem

There is a beautiful relation between orbits and isotropy groups: Note that for any $x \in X$, the orbit $\mathrm{Orb}_G(x)$ has a transitive left-$G$-action. On the other hand, the set of cosets $G/G^x$ also has a transitive left $G$-action. They are both $G$-sets, and in fact they are isomorphic as $G$-sets. This is the:

**Theorem** [Stabilizer-Orbit Theorem]: There is a natural isomorphism of $G$-sets

$$\psi : \mathrm{Orb}_G(x) \to G/G^x \tag{8.33}$$

Put differently: The points in the $G$-orbit of $x$ are in natural $1-1$ correspondence with the left cosets of $\mathrm{Stab}_G(x)$ in a way compatible with the left $G$-action on these sets.

*Proof*: Suppose $y$ is in a $G$-orbit of $x$. Then $\exists g$ such that $y = g \cdot x$. Define

$$\psi(y) := g \cdot G^x \ . \tag{8.34}$$

The first thing we need to do is check that $\psi$ is well-defined . This is easily checked:

$$y = g' \cdot x \qquad \rightarrow \qquad \exists h \in G^x \qquad g' = g \cdot h \qquad \rightarrow \qquad g'G^x = ghG^x = gG^x \tag{8.35}$$

Now let us check that $\psi$ is one-one. Given a coset $g \cdot G^x$ we may define

$$\psi^{-1}(gG^x) \equiv g \cdot x \tag{8.36}$$

Again, we must check that this is well-defined, and again this is easily checked. Since it inverts $\psi$, $\psi$ is 1-1. Moreover, the action is compatible with the $G$-action: For all $y \in \mathrm{Orb}_G(x)$ and $g \in G$:

$$\psi(g \cdot y) = g \cdot \psi(y) \qquad y \in \mathrm{Orb}_G(x) \tag{8.37}$$

so $\psi$ is a $G$-equivariant, (recall the definition (5.27)) and one-one. That is it is an isomorphism of $G$-sets ♠

Corollary: If $G$ acts transitively on a set $X$ then the isotropy groups $G^x$ for all the points $x \in X$ are conjugate subgroups of $G$, and any choice of a point $x \in X$ sets up a $1-1$ correspondence between points of $X$ and elements of the set of cosets $G/G^x$. Put differently: If $H \subset G$ is an isotropy subgroup $G^x$ for some $x \in X$ then we can identify $X$ with the set of left-cosets $G/H$.

**Remarks**:

1. Sets of the type $G/H$ are examples of *homogeneous spaces*. This theorem is the beginning of an important connection between the *algebraic* notions of subgroups and cosets to the *geometric* notions of orbits and fixed points. Below we will show that if $G, H$ are topological groups then, in some cases, $G/H$ are beautifully symmetric topological spaces, and if $G, H$ are Lie groups then, in some cases, $G/H$ are beautifully symmetric manifolds.

2. *Homogeneous Spaces And Spontaneous Symmetry Breaking* One way homogeneous spaces arise field theory is via the description of classical vacua of a scalar field theory with a global symmetry. Suppose $\phi$ is a (real) scalar field on $d$-dimensional Minkowski space, $\mathbb{M}^{1,d-1}$. So it is valued in some real vector space $V$ which is a carrier space of a finite dimensional representation of a global symmetry group $G$. Suppose that $U(\phi)$ is a $G$-invariant potential energy. This means that

$$U(\phi) = U(g \cdot \phi) \qquad \forall g \in G \tag{8.38}$$

Typically it is an invariant polynomial in $\phi$. Suppose we have a $G$-invariant metric, i.e. a symmetric bilinear form

$$b : V \otimes V \to \mathbb{R} \tag{8.39}$$

which is $G$-invariant:

$$b(g \cdot v_1, g \cdot v_2) = b(v_1, v_2) \qquad \forall g \in G \qquad \forall v_1, v_2 \in V \tag{8.40}$$

Our action will be

$$S[\phi] = \int_{\mathbb{M}^{1,d-1}} \left[\frac{1}{2}\eta^{\mu\nu} b(\partial_\mu\phi, \partial_\nu\phi) - U(\phi)\right] d^d x \tag{8.41}$$

where we have chosen a signature which is mostly minus. For a physically reasonable theory we should choose $b$ to be positive definite and $U : V \to \mathbb{R}$ to be bounded below.

If we choose a basis $\{e_s\}$ for $V$ then $\phi(x) = \sum_s e_s \phi^s(x)$ is a description in terms of a collection of $n$ scalar fields $\phi^s$, $s = 1, \dots, n$ where $n$ is the real dimension of $V$. Then

$$b_{st} = b(e_s, e_t) \tag{8.42}$$

and

$$S[\phi] = \int_{\mathbb{M}^{1,d-1}} \left[\frac{1}{2}\eta^{\mu\nu} b_{st}\partial_\mu\phi^s\partial_\nu\phi^t - U(\phi)\right] d^d x \tag{8.43}$$

Then conjugate momentum is

$$\Pi_s = b_{st}\partial_0\phi^t \tag{8.44}$$

and the Hamiltonian is

$$H = \int_{\mathbb{R}^{d-1}} \left[\frac{1}{2}b^{st}\Pi_s\Pi_t + \frac{1}{2}b_{st}\partial_i\phi^s\partial_i\phi^t + U(\phi)\right] d^d x \tag{8.45}$$

where $\Pi$ is the canonical conjugate momentum to $\phi$ and $b^{st}$ is the inverse of the matrix $b_{st}$. The energy will be bounded below if $b$ is positive semidefinite and $U$ is bounded below. For a good kinetic term we assume that $b$ is nondegenerate.

Then, in classical field theory we minimize the energy by setting $\Pi = 0$ and $\partial_i\phi = 0$ so the field $\phi(x)$ is constant on the spatial slice $\mathbb{R}^{d-1}$. Therefore, we can consider it to be a single vector $\phi \in V$. Moreover, the constant value of $\phi$ minimizes $U(\phi)$. Call the minimum value $U_0$. Of course if $\phi \in V$ minimizes $U$ then so does $g \cdot \phi$ for any $g \in G$, by $G$-invariance. The set of classical vacua $\mathcal{M}^{cv}$ is defined to be the set of field configurations minimizing the energy:

$$\mathcal{M}^{cv} := \{\phi(x) = \phi \in V | U(\phi) = U_0\} \tag{8.46}$$

Because $U$ is $G$-invariant $\mathcal{M}^{cv}$ is a $G$-space. It therefore decomposes as a union of $G$-orbits. In some important cases the $G$-action is transitive and we can identify the space of classical vacua as a homogeneous space.

Here is a simple example: Suppose $\phi \in \mathbb{R}^n$ is a real-valued vector and $G = SO(n)$ and the invariant metric $b$ is the standard Euclidean metric on $\mathbb{R}^n$. Then take

$$U(\phi) = \lambda(\phi \cdot \phi - v^2)^2, \tag{8.47}$$

where the coupling constant $\lambda > 0$ so that the energy is bounded below and $v$ is some nonzero real constant so that $v^2 > 0$. Then the set of classical vacua is

$$\mathcal{M} = \{\phi \in \mathbb{R}^n | \phi \cdot \phi = v^2\} \tag{8.48}$$

which is manifestly an $(n-1)$ dimensional sphere of radius $|v|$, and by the stabilizer-orbit theorem this can be identified with the homogeneous space $\mathcal{M} \cong S^{n-1} \cong SO(n)/SO(n-1)$. In this case the stabilizer of any classical vacuum $\phi \in \mathcal{M}$ is non-trivial and isomorphic to $SO(n-1)$. The symmetry group of the field configuration $\phi = 0$ is all of $SO(n)$, and the symmetry group of the action is $SO(n)$. But the choice of vacuum has reduced the "unbroken" symmetry. This is an example of *spontaneous symmetry breaking.*

More generally, when we choose a classical vacuum in $\mathcal{M}$ whose isotropy group $H$ is a proper subgroup of $G$ we say that there has been spontaneous symmetry breaking and that the vacuum has "broken" the symmetry from $G$ to $H$. The "spontaneous" refers to the fact that a physical system is built by excitations around some choice of the vacuum.

Note that, if $G$ and $H$ are Lie groups and $H$ is a proper subgroup then $G/H$ is a positive dimensional manifold. Consider field configurations of the form

$$\phi(x,t) = g(x,t) \cdot \phi \tag{8.49}$$

where $g(x,t)$ is a map from spacetime to the Lie group $G$. The in other words, we only consider the smaller space of maps to the manifold of classical vacua - the homogeneous space. The action, when restricted to these field configurations is, up to a constant that does not affect the dynamics,

$$S = \int_{\mathbb{R}^d} \left[ \frac{1}{2} \eta^{\mu\nu} b(g^{-1}\partial_\mu g \cdot \phi_0, g^{-1}\partial_\nu g \cdot \phi_0) \right] d^d x \tag{8.50}$$

Working out the dispersion relation for slowly varying functions $g(x,t)$ shows that there are massless modes of the field describing slow variation of the field along the orbits $G/H$ in the space of classical vacua. These the particles associated with these field modes are known as *Goldstone bosons.*

******************

DRAW MEXICAN HAT POTENTIAL

**********************

---

**Exercise** *Isomorphic Homogeneous spaces*

Suppose $G$ acts on $X$. Suppose that $x, y$ are in the same $G$-orbit. Construct an isomorphism of $G$-spaces [104]

$$\Psi : G/G^x \to G/G^y \tag{8.51}$$

[104] *Answer*: Choose $g_0 \in G$ such that $y = g_0 x$. Check that $\Psi : gG^x \mapsto (g_0 g g_0^{-1})G^y$ is well-defined and defines an isomorphism of $G$-spaces.

**Remark**: Note that there is no canonical isomorphism. This is why we cannot say that $G = G/H \times H$ as topological spaces. Rather $\pi : G \to G/H$ is a nontrivial $H$-bundle (in general). We will explicate this remark below.

---

**Exercise** *Orbits Of* $\mathbb{Z}_p$ *For Prime p*

Let $p$ be a prime and suppose the cycle group $\mathbb{Z}_p$ acts on a space $X$. Show that any orbit consists of either a single point, or of $p$ distinct points. [105]

---

---

**Exercise** *The Lemma that is not Burnside's*

Suppose a finite group $G$ acts on a finite set $X$ as a transformation group. A common notation for the set of points fixed by $g$ is $X^g$. Show that the number of distinct orbits is the averaged number of fixed points:

$$|\{orbits\}| = \frac{1}{|G|}\sum_g |X^g| \tag{8.52}$$

For the answer see. [106]

---

---

**Exercise** *Jordan's theorem*

Suppose $G$ is finite and acts transitively on a finite set $X$ with more than one point. Show that there is an element $g \in G$ with no fixed points on $X$. [107]

---

[105] *Answer*: By the stabilizer-orbit theorem the orbits are, as $G$-spaces, just $\mathbb{Z}_p/H$ where $H$ is a subgroup of $\mathbb{Z}_p$. But the only subgroups are the trivial group and the entire group.

[106] *Answer*: Write

$$\sum_{g\in G} |X^g| = |\{(x,g)|g\cdot x = x\}| = \sum_{x\in X} |G^x| \tag{8.53}$$

Now use the stabilizer-orbit theorem to write $|G^x| = |G|/|\mathcal{O}_G(x)|$. Now in the sum

$$\sum_{x\in X} \frac{1}{|\mathcal{O}_G(x)|} \tag{8.54}$$

the contribution of each distinct orbit is exactly 1.

[107] *Answer*: Apply the Burnside lemma. Since the action is transitive the LHS is 1. Break up the sum on the RHS into contributions from $g = 1$ and $g \neq 1$. The RHS becomes

$$RHS = \frac{1}{|G|}\left(|X| + \sum_{g\neq 1}|X^g|\right) \tag{8.55}$$

Suppose that every $g \in G$ has at least one fixed point. Then $X^g$ is nonempty for all $g$ and hence $RHS \geq \frac{1}{|G|}\left(|X| + (|G|-1)\right) = 1 + \frac{1}{|G|}(|X|-1)$. This is $1 + positive$ and cannot be equal to 1.

---

---

**Exercise** *Orders Of The Conjugacy Classes In* $S_n$
Prove equation (7.159) above. [108]

---

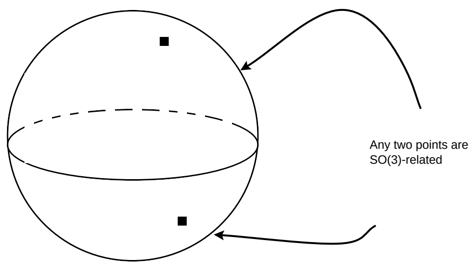



**Figure 6:** Transitive action of $SO(3, \mathbb{R})$ on the sphere.

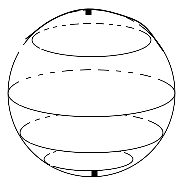


**Figure 7:** Orbits of $SO(2, \mathbb{R})$ on the two sphere.

### 8.4 Practice With Group Action Terminology

The concept of a $G$-action on a set is an extremely important concept, so let us consider a number of examples:

#### Examples

1. Let $G$ be any group and consider the group action defined by $\phi(g, x) = x$ for all $g \in G$. This is as ineffective as a group action can be: For every $x$, the istropy group is all of $G$, and for all $g \in G$, $\mathrm{Fix}(g) = X$. In particular, this situation will arise if $X$ consists of a single point. This example is not quite as stupid as might at first appear, once one takes the categorical viewpoint, for $pt//G$ is a very rich category indeed. See section 17 below on category theory.

[108] *Answer*: Consider the $G = S_n$ action on a particular conjugacy class by conjugation. This is clearly transitive. Now consider a particularly convenient permutation like:

$$(1)(2)\cdots(\ell_1)(\ell_1+1, \ell_1+2)\cdots(\ell_1+2\ell_2-1, \ell_1+2\ell_2)\cdots$$

and compute the stabilizer. The answer for the stabilizer is in equation (14.67) below. (You don't need to know the exact structure to compute the order.)

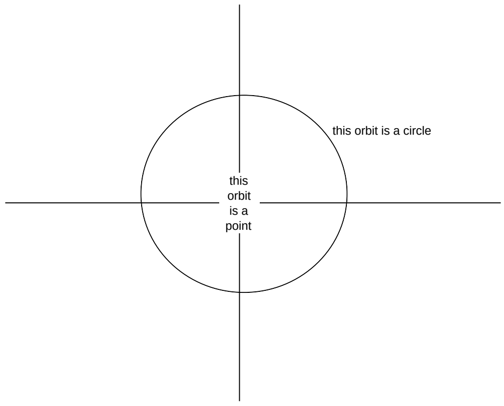



**Figure 8:** Notice not all orbits have the same dimensionality. There are two qualitatively different kinds of orbits of $SO(2,\mathbb{R})$.

2. Let $X = \{1, \cdots n\}$, so there is a left action of $S_X = S_n$ as we have discussed many times.

   a.) The action is effective: Every nontrivial permutation changes some number between 1 and $n$.

   b.) The action is transitive: For example, $i, j$ are mapped to each other by the permutation $(ij)$. (And many other permutations.)

   c.) The action is not free: The fixed point of any $j \in X$ is just the group of permutations $S_X^j$ that don't change $j$: $S_X^j \cong S_{n-1}$. Note that different $j$ have different stabilizer subgroups isomorphic to $S_{n-1}$, but they are all conjugate.

3. $GL(n,\mathbb{R})$ acts on $\mathbb{R}^n$ by matrix multiplication. If we act with a matrix on a column vector we get a left action. If we act on a row vector we get a right action. The action is:

   a.) Effective: If $g \neq 1$ some vector $\vec{x}$ is moved.

   b.) Not transitive: If $\vec{x} \neq 0$ it cannot be mapped to 0

   c.) Not free: $\vec{0}$ is a fixed point of the entire group.

   d.) There are two orbits.

   e.) The isotropy group of the vector $e_1$ is (under the left-action) the subgroup of matrices of the form

$$\mathrm{Stab}_{GL(n,\mathbb{R})}(e_1) = \{ \begin{pmatrix} 1 & v \\ 0 & B \end{pmatrix} | v \in Mat_{1\times(n-1)}(\mathbb{R}), \qquad B \in GL(n-1,\mathbb{R}) \} \tag{8.56}$$

The stabilizer group for all other nonzero vectors will be conjugate to this one. The stabilizer group of the origin is the entire group $GL(n,\mathbb{R})$.

4. If we restrict from $GL(n,\mathbb{R})$ to $SO(n,\mathbb{R})$ the picture changes completely. For simplicity consider the case $n=2$ acting on $\mathbb{R}^2$. The left- action is:

$$R(\phi) : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{8.57}$$

   The group action is effective. It is not free, and it is not transitive. There are now infinitely many orbits of $SO(2)$, and they are all distinguished by the invariant value of $x^2+y^2$ on the orbit. From the viewpoint of topology, there are two distinct "kinds" of orbits acting on $\mathbb{R}^2$. One has trivial isotropy group and one has isotropy group $SO(2)$. See Figure 8. These give two strata.

5. Orbits of $O(2)$ acting on $\mathbb{R}^2$. We have seen that $O(2)$ can be written as a disjoint union:

$$O(2) = SO(2) \amalg P \cdot SO(2) \tag{8.58}$$

   where $P$ is not canonical and can be taken to be reflection in any line through the origin. The orbits of $SO(2)$ and $O(2)$ are the same. We will find a very different picture when we consider the orbits of the Lorentz group.

6. Now consider a fixed $SO(2,\mathbb{R})$ subgroup of $SO(3,\mathbb{R})$, say, the subgroup defined by rotations around the $z$-axis, and consider the action of this group on a sphere in $\mathbb{R}^3$ of fixed radius. The action is <u>*not*</u> transitive. The $G$-orbits are shown in Figure 7. It is also not free: The north and south poles are fixed points.

7. Now consider the action of $SO(3,\mathbb{R})$ on a sphere of positive fixed radius in $\mathbb{R}^3$. (WLOG take it to be of radius one.) The action is then transitive on the sphere. Now the isotropy subgroup $\mathrm{Stab}_{SO(3)}(\hat{n}) \subset SO(3)$ of <u>any</u> unit vector $\hat{n} \in S^2$ is isomorphic to $SO(2)$:

$$\mathrm{Stab}_{SO(3)}(\hat{n}) \cong SO(2) \tag{8.59}$$

   But, for different choices of $\hat{n}$ we get different subgroups of $SO(3)$. For example, with usual conventions, if $\hat{n} = e_3$ is on the $x^3$-axis then the subgroup is the subgroup of matrices of the form

$$R_{12}(\phi) = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{8.60}$$

   but if $\hat{n} = e_1$ is on the $x^1$-axis the subgroup is the subgroup of matrices of the form

$$R_{23}(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{pmatrix} \tag{8.61}$$

   and if $\hat{n} = e_2$ is on the $x^2$ axis the isotropy subgroup is the group of matrices of the form

$$R_{13}(\phi) = \begin{pmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{pmatrix} \tag{8.62}$$

and so on. For any $\hat{n} \in S^2$ let $SO(2)_{\hat{n}} \subset SO(3)$ denote the subgroup, isomorphic to $SO(2)$, which stabilizes $\hat{n}$. According to the stabilizer-orbit theorem there is a natural one-one correspondence

$$S^2 \cong SO(3)/SO(2)_{\hat{n}} \tag{8.63}$$

Therefore, fixing any $\hat{n} \in S^2$ there is a map

$$\pi_{\hat{n}} : SO(3, \mathbb{R}) \to S^2 \tag{8.64}$$

Put simply, $\pi(R)$ rotates $\hat{n} \in S^2$ to $R \cdot \hat{n} \in S^2$:

$$\pi_{\hat{n}}(R) := R \cdot \hat{n} \in S^2 \tag{8.65}$$

Therefore, the inverse image of any other vector $\hat{k} \in S^2$:

$$\pi_{\hat{n}}^{-1}(\hat{k}) := \{R | R\hat{n} = \hat{k}\} \subset SO(3) \tag{8.66}$$

is the set of rotations which can be (noncanonically!) put in 1-1 correspondence with elements of $SO(2)$. That is because if $\hat{k} = R_1\hat{n}$ and $\hat{k} = R_2\hat{n}$ then $R_1^{-1}R_2\hat{n} = \hat{n}$ and therefore $R_2 = R_1R_0$ where $R_0 \in \mathrm{Stab}_{SO(3)}(\hat{n}) \cong SO(2)$.

So, for each point $\hat{k} \in S^2$ we can associate a copy of $SO(2)$ inside $SO(3)$, which is topologically a circle, and clearly every element of $SO(3)$ will be captured this way as $\hat{k}$ ranges over $S^2$. One might think that this means that, as manifolds, $SO(3)$ is diffeomorphic to $S^2 \times SO(2) = S^2 \times S^1$, but this turns out to be quite false. For example the homotopy groups of $SO(3)$ and $S^2 \times S^1$ are completely different.

Nevertheless, we can try to parametrize the general rotation by using this idea: We choose $\hat{n} = e_3$ to be the basepoint. Then the standard polar angles of a point on the sphere are defined by [109]

$$R_{12}(\phi)R_{23}(\theta)\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}\sin\theta\sin\phi\\ \sin\theta\cos\phi\\ \cos\theta\end{pmatrix} \tag{8.68}$$

But this does NOT mean every rotation matrix is of the form $R_{12}(\phi)R_{23}(\theta)$ ! It only gives us a parametrization of the cosets $SO(3)/SO(2)_{e_3}$. We now get the a parametrization of the general element of $SO(3)$ by including a factor on the right by a general element of the stabilizer group of $\hat{n} = e_3$:

$$R = R_{12}(\phi)R_{23}(\theta)R_{12}(\psi) \tag{8.69}$$

[109] One usually defines the polar angle $\phi$ so that $x = \sin\theta\cos\phi$. To get the standard parametrization of $(x, y, z)$ by polar angles replace

$$R_{12}(\phi) \to R_{12}(\frac{\pi}{2} - \phi) \ . \tag{8.67}$$

and if we want a generic element to have a unique set of "coordinates" $(\phi, \theta, \psi)$ we should choose the range:

$$\begin{aligned} \phi &\sim \phi + 2\pi \\ \psi &\sim \psi + 2\pi \\ 0 \leq &\theta \leq \pi \end{aligned} \tag{8.70}$$

These are the famous *Euler angles.* Although every $SO(3)$ matrix has an Euler angle presentation, and generic $SO(3)$ matrices have a unique such parametrization, there are some matrices for which the parametrization is not unique. For example, if we put $\theta = 0$ then the parametrization collapses to $R_{12}(\phi + \psi)$. If $(\phi, \theta, \psi)$ were truly good coordinates we could set $\theta = 0$ and get a two-dimensional subspace of $SO(3)$. But in fact we only get one-dimensional subspace because only the combination $\phi + \psi$ appears. This explains why the generic matrix has a unique Euler angle presentation, but the manifold $SO(3)$ is not the same as the manifold $S^2 \times S^1$.

8. $SU(2) \cong S^3$ *From Stabilizer-Orbit Theorem.* First we reprove the result - already seen in Example 2.7 above - that, as a manifold, $SU(2)$ can be identified with the unit three-dimensional sphere. We will give another proof using the Stabilizer-Orbit theorem. Consider the unit sphere in $\mathbb{R}^4$ as the space of unit vectors in a two-dimensional complex Hilbert space (the space of states of "one Qbit"):

$$S^3 \cong \{\vec{z} | \vec{z}^\dagger \vec{z} = 1\} \subset \mathbb{C}^2 \tag{8.71}$$

This is easily seen by writing

$$\vec{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{8.72}$$

and decomposing $z_1, z_2$ into their real and imaginary parts. Next, we note that $SU(2)$ has a transitive action on the unit sphere:

$$\phi_u : \vec{z} \mapsto u\vec{z} \tag{8.73}$$

The action is transitive because, given any unit vector we can find another orthogonal unit vector. But any two ON bases are related by some unitary transformation. By changing the phase of the second vector we can arrange that they are related by a special unitary transformation.

Therefore, we should invoke the stabilizer-orbit theorem and compute the stabilizer of, say

$$\vec{z}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad . \tag{8.74}$$

These are the upper triangular $SU(2)$ matrices with one on the diagonal: The stabilizer is trivial. So

$$SU(2) \cong S^3 \tag{8.75}$$

In particular, a general $SU(2)$ element must have the form

$$u = \begin{pmatrix} z_1 & s \\ z_2 & t \end{pmatrix} \tag{8.76}$$

with $|z_1|^2+|z_2|^2 = 1$ and $s, t$ to be determined. These can be determined by imposing the condition:

$$u^{-1} = u^\dagger \tag{8.77}$$

we solve for the other two matrix elements and recover the result from Example 2.7 that every $SU(2)$ element is of the form

$$u = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{8.78}$$

where

$$|\alpha|^2 + |\beta|^2 = 1 \ . \tag{8.79}$$

9. *$GL(2,\mathbb{C})$ And $SU(2)$ Act On $\mathbb{CP}^1$.* We continue to develop ideas from Section **** 5.3, example 4 **** Recall that $\mathbb{CP}^1$ can be identified with equivalence classes of points $(z_1, z_2) \in \mathbb{C}^2 - \{0\}$ with equivalence relation $(z_1', z_2') \sim (\lambda z_1, \lambda z_2)$. We denote equivalence classes by $[z_1 : z_2]$.

   Note that $\mathbb{CP}^1$ can also be thought of as the space of states of a single Qbit: $[z_1 : z_2]$ always has a representative with $|z_1|^2 + |z_2|^2 = 1$ and the representative is unique up to multiplication by a phase. We can use such a normalized representative to define a vector in a Hilbert space corresponding to a Qbit state:

$$\psi = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \tag{8.80}$$

   There is a well-defined action of $GL(2;\mathbb{C})$ on $\mathbb{CP}^1$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : [z_1 : z_2] \mapsto [az_1 + bz_2 : cz_1 + dz_2] \tag{8.81}$$

   (The reader should carefully check that this is a well-defined group action. Since the $GL(2,\mathbb{C})$ action on $\mathbb{C}^2 - \{0\}$ is transitive, the action on $\mathbb{CP}^1$ is transitive. Therefore choosing a point $p \in \mathbb{CP}^1$ we have an identification of $\mathbb{CP}^1$ as a homogeneous space:

$$GL(2,\mathbb{C})/B \cong \mathbb{CP}^1 \qquad B = \mathrm{Stab}_{GL(2,\mathbb{C})}(p) \ . \tag{8.82}$$

   For example, if we take $B$ to be the stabilizer of $[1 : 0]$ we compute

$$B = \{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} | a, d \in \mathbb{C}^* \qquad b \in \mathbb{C} \} \tag{8.83}$$

As we have just discussed, the $SU(2)$ action on normalized vectors is transitive. But thanks to the equivalence $[z_1 : z_2] = [\lambda z_1 : \lambda z_2]$ we can always find a representative so that $(z_1, z_2)$ is a unit vector in the Hilbert space $\mathbb{C}^2$. Therefore, the restriction of the $GL(2,\mathbb{C})$ action on $\mathbb{CP}^1$ to $SU(2)$ is still transitive. Now the stabilizer of $[1 : 0]$ is the subgroup of diagonal $SU(2)$ matrices and is isomorphic to $U(1)$. Therefore, there is also an identification

$$\mathbb{CP}^1 \cong SU(2)/U(1) \tag{8.84}$$

Thus $\mathbb{CP}^1$ as a homogeneous space can be presented both as a homogeneous space of compact Lie groups (hence it is compact) and as a homogeneous space of complex Lie groups. Hence $\mathbb{CP}^1$ is a compact complex manifold.

Moreover, as always with homogeneous spaces there is a natural map

$$\pi : SU(2) \to SU(2)/U(1) \cong \mathbb{CP}^1 \tag{8.85}$$

Since $S^3 \cong SU(2)$ we have a natural continuous map:

$$\pi : S^3 \to S^2 \tag{8.86}$$

whose fibers are copies of $S^1$. This is a famous map in mathematics and physics known as the *Hopf map* and has many beautiful properties. It appears in the physics of magnetic monopoles and in several other related contexts. It is very closely related to the map $\pi : SO(3) \to S^2$ defined above.

Note well that $S^3$ is not homeomorphic to $S^2 \times S^1$. This is easily seen by considering fundamental groups.

Another way of thinking about $\mathbb{CP}^1$ is that it is the space of lines through the origin in $\mathbb{C}^2$. This leads to the idea of Grassmannians described below.

---

**Exercise** $\mathbb{Z}_2$ *Actions On The Sphere*
Consider the action of $\mathbb{Z}_2$ on the sphere defined by (5.8):

$$\sigma \cdot (x^1, \dots, x^{n+1}) = (x^1, \dots, x^p, -x^{p+1}, \cdots, -x^{p+q}) \tag{8.87}$$

a.) For which values of $p, q$ is the action effective?
b.) For which values of $p, q$ is the action transitive?
c.) Compute the fixed point set of the nontrivial element $\sigma \in \mathbb{Z}_2$.
d.) For which values of $p, q$ is the action free?

---

---

**Exercise** $\mathbb{C}^*$ *Actions On* $\mathbb{CP}^{n-1}$

Consider the action of $G = \mathbb{C}^*$ on $\mathbb{CP}^{n-1}$ defined by (5.9).

a.) For which values of $(q_1, \ldots, q_n)$ is the action effective?

b.) For which values of $(q_1, \ldots, q_n)$ is the action transitive?

c.) What are the fixed points of the $\mathbb{C}^*$ action?

d.) What are the stabilizers at the fixed points of the $\mathbb{C}^*$ action?

---

---

**Exercise** $SL(2,\mathbb{R})$ *Action On The Upper Half-Plane*

a.) Show that (5.11) above defines a left-action of $SL(2,\mathbb{R})$ on the complex upper half-plane. [110]

b.) Is the action effective?

c.) Is the action transitive?

d.) Which group elements have fixed points?

e.) What is the isotropy group of $\tau = \mathrm{i}$ ? [111]

Conclude that

$$\mathcal{H} \cong SL(2,\mathbb{R})/SO(2) \tag{8.88}$$

---

---

**Exercise**

Using $GL(2,\mathbb{C})/B \cong \mathbb{CP}^1$ show that $GL(2,\mathbb{C})$ has a natural action on the Riemann sphere given by

$$z \mapsto \frac{az+b}{cz+d} \tag{8.89}$$

---

---

**Exercise**

Since there is a left-action of $G \times G$ on $X = G$ there is a left-action of the diagonal subgroup $\Delta \subset G \times G$ where $\Delta = \{(g,g)|g \in G\}$ is a subgroup isomorphic to $G$.

a.) Show that this action is given by $a \mapsto I(a)$, where $I(a)$ is the conjugation by $a$.

b.) Show that the orbits of $\Delta$ are the conjugacy classes of $G$.

c.) What is the stabilizer subgroup of an element $g_0 \in G$?

---

[110] *Hint*: Show that $\mathrm{Im}(g \cdot \tau) = \frac{\mathrm{Im}\tau}{|c\tau+d|^2}$.

[111] The isotropy group is the subgroup $SO(2,\mathbb{R}) \subset SL(2,\mathbb{R})$. To see this set $\frac{a\mathrm{i}+b}{c\mathrm{i}+d} = \mathrm{i}$ and conclude that $a = d$ and $b = -c$. Then since $ad - bc = 1$ we have $a^2 + b^2 = 1$ but this implies that the group element is in $SO(2,\mathbb{R})$.

---

**Exercise** *Spheres As Homogeneous Spaces*

a.) Show that there is a transitive action of $SO(n+1)$ on $S^n$, considered as a sphere of fixed radius in $\mathbb{R}^{n+1}$.

b.) Show that $S^n \cong SO(n+1)/SO(n)$.

c.) Give an inductive proof that $SO(n)$ is a connected manifold for $n \geq 2$.

---

### 8.4.1 More About The Relation Of $SU(2)$ And $SO(3)$

Let us now return to the basic homomorphism

$$R : SU(2) \to SO(3) \quad . \tag{8.90}$$

defined in section 5.4.1. Recall the key equation (5.39):

$$u\vec{x} \cdot \vec{\sigma} u^{-1} = (R(u)\vec{x}) \cdot \vec{\sigma} \tag{8.91}$$

Note that it is immediate from the definition that $R(-u) = R(u)$ and hence the kernel of $R$ must at least contain the subgroup of matrices proportional to the identity: $\{\pm 1\}$.

We will now prove that:

1. $\ker(R) = \{\pm \mathbf{1}_{2\times 2}\} = Z(SU(2))$.
2. Every proper rotation comes from some $u \in SU(2)$, i.e. the homomorphism $R$ is surjective

Thus we have the extremely important extension:

$$1 \to \mathbb{Z}_2 \overset{\iota}{\to} SU(2) \overset{R}{\to} SO(3) \to 1 \tag{8.92}$$

Thus, $SU(2)$ is a two-fold cover of $SO(3)$ and in fact

$$SO(3, \mathbb{R}) \cong SU(2)/\mathbb{Z}_2 \tag{8.93}$$

where the $\mathbb{Z}_2$ we quotient by is the center $\{\pm \mathbf{1}_{2\times 2}\}$. This is arguably the most important exact sequence in physics.

To prove the above two claims we will need to get to know $SU(2)$ a bit better.

There are many ways to parametrize $S^3$. One is to introduce a polar angle and stratify $S^3$ by two-dimensional spheres. Viewed this way, we can write the general $SU(2)$ element as

$$u = \cos\chi + \mathrm{i}\sin\chi \vec{n} \cdot \vec{\sigma} \tag{8.94}$$

where $0 \leq \chi \leq \pi$ and $\vec{n} \in \mathbb{R}^3$ with $\vec{n} \cdot \vec{n} = 1$, so $\vec{n} \in S^2$.

Now suppose that $u \in \ker(R)$. Then $u$ must commute with $\sigma^k$ for $k = 1, 2, 3$. But from (8.94) it is easy to check that $u$ commutes with $\sigma^k$ only if $\sin\chi = 0$ so $\cos\chi = \pm 1$. From this we conclude

$$\ker(R) = \{\pm \mathbf{1}_{2\times 2}\} = Z(SU(2)) \tag{8.95}$$

Now to prove surjectivity we return to the parametrization (8.94). We claim that if $u$ is parametrized as in that equation then

$$u\vec{x}\cdot\vec{\sigma}u^{-1} = \vec{y}\cdot\vec{\sigma} \tag{8.96}$$

where $\vec{y}$ is obtained from $\vec{x}$ by rotation by angle $-2\chi$ around the $\hat{n}$ axis. See the exercise below for the demonstration. Since every element of $SO(3)$ can be expressed as a rotation around some axis, this shows that $R$ is surjective. We have now completed the proof of equation (8.92).

**Remarks**:

1. Once again note that, although $SU(2)$ is a 2-fold cover of $SO(3)$ it is not true that $SU(2) = SO(3)\times\mathbb{Z}_2$, neither as manifolds (since both $SU(2)$ and $SO(3)$ are connected), nor as groups.

2. The short exact sequence (8.92) can be generalized in two ways. First, $SU(2)$ is isomorphic to a "spin group" $Spin(3)$. By definition, $Spin(3)$ is the group of matrices obtained by even products of the form

$$\hat{n}_1\cdot\sigma\hat{n}_2\cdot\sigma\cdots n_{2k-1}\cdot\sigma n_{2k}\cdot\sigma \tag{8.97}$$

Note that

$$\hat{n}_1\cdot\sigma\hat{n}_2 = \hat{n}_1\cdot\hat{n}_2 1_{2\times 2} + \mathrm{i}\hat{n}_1\times\hat{n}_2\cdot\sigma \tag{8.98}$$

is of the form (8.94) where $\chi$ is the angle between $\hat{n}_1$ and $\hat{n}_2$. The generalization to $Spin(n)$ involves considering the Clifford algebra in $n$-dimensions (See Chapter ***) and considering the group of even products of elements of the form $v\cdot\gamma$, where $v$ is a unit vector. Then one has the exact sequence

$$1\to\mathbb{Z}_2\to Spin(n)\to SO(n)\to 1 \tag{8.99}$$

A second generalization is

$$1\to Z(SU(n))\to SU(n)\to PSU(n)\to 1 \tag{8.100}$$

since $PSU(2)\cong SO(3)$.

We can put the above equation (8.92) to good use and show how the Euler angles give a nice parametrization of $SU(2)$.

Recall that the exponential of any matrix $A$ (or any endomorphism of a vector space) is defined by the usual series:

$$\exp A := 1 + A + \frac{1}{2!}A^2 + \cdots \tag{8.101}$$

Note that $(\vec{n}\cdot\vec{\sigma})^2 = 1_{2\times 2}$ and hence

$$u = \cos\chi + \mathrm{i}\sin\chi\vec{n}\cdot\vec{\sigma} = \exp(\mathrm{i}\chi\vec{n}\cdot\vec{\sigma}) \tag{8.102}$$

So, for any $\vec{n}\in S^2$ and $\chi\in\mathbb{R}$

$$R\left(\exp(\mathrm{i}\chi\vec{n}\cdot\vec{\sigma})\right) \tag{8.103}$$

is a rotation around the $\vec{n}$ axis by angle $-2\chi$.

As a nice check consider the restriction of the homomorphism $R$ to the subgroup of diagonal matrices $D$ of the form:

$$\begin{pmatrix} \xi & 0 \\ 0 & \xi^{-1} \end{pmatrix} \tag{8.104}$$

with $|\xi| = 1$. These act on $\vec{x}\cdot\vec{\sigma}$ by

$$\begin{pmatrix} \xi & 0 \\ 0 & \xi^{-1} \end{pmatrix}\begin{pmatrix} x_3 & x_1 - \mathrm{i}x_2 \\ x_1 + \mathrm{i}x_2 & -x_3 \end{pmatrix}\begin{pmatrix} \xi^{-1} & 0 \\ 0 & \xi \end{pmatrix} = \begin{pmatrix} x_3 & \xi^2(x_1 - \mathrm{i}x_2) \\ \xi^{-2}(x_1 + \mathrm{i}x_2) & -x_3 \end{pmatrix} \tag{8.105}$$

If we write

$$\xi = e^{-\mathrm{i}\phi/2} \tag{8.106}$$

for some angle $\phi$ then $\pi$ maps the diagonal matrix to $R\in SO(3)$ that is a rotation around the $x^3$ axis. It is a counterclockwise rotation by $\phi$ in the $(x^1, x^2)$ plane with the orientation $dx^1\wedge dx^2$. Moreover, it is useful to note that

$$e^{-\mathrm{i}\frac{\phi}{2}\sigma^3} = \begin{pmatrix} \xi & 0 \\ 0 & \xi^{-1} \end{pmatrix} \tag{8.107}$$

where the exponential of a matrix is defined by the usual series expansion. So, we conclude that

$$R(e^{-\mathrm{i}\frac{\phi}{2}\sigma^3}) = R_{12}(\phi) \tag{8.108}$$

**Nota Bene:** If we only know $\phi$ modulo $2\pi$ (as opposed to $\phi/2$ modulo $\pi$) then $e^{-\mathrm{i}\frac{\phi}{2}\sigma^3}$ but since $R(-u) = R(u)$ equation (8.108) is meaningful.

In a similar way the above result implies

$$R\left(e^{-\mathrm{i}\frac{\theta}{2}\sigma^1}\right) = R_{23}(\theta) \tag{8.109}$$

Note that in the Euler angle parametrization we could also have used rotation in the 13 plane. In this case we would use the $SO(2)$ subgroup of $SU(2)$ consisting of the real unitary matrices. We parametrize the group by

$$u(\theta/2) = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix} = e^{-\mathrm{i}\frac{\theta}{2}\sigma^2} \tag{8.110}$$

Of course, the eigenvalues are $e^{\pm\mathrm{i}\theta/2}$ and indeed

$$Se^{-\mathrm{i}\frac{\theta}{2}\sigma^2}S^{-1} = e^{-\mathrm{i}\frac{\theta}{2}\sigma^3} \tag{8.111}$$

where

$$S = \frac{1}{\sqrt{2}}(1 - \mathrm{i}\sigma^1) \in SU(2) \tag{8.112}$$

(all you have to check is $S\sigma^2 S^{-1} = \sigma^3$. ) We find that

$$\pi(e^{-\mathrm{i}\frac{\theta}{2}\sigma^2}) = R_{13}(\theta) := \begin{pmatrix} \cos(\theta/2) & 0 & \sin(\theta/2) \\ 0 & 1 & 0 \\ -\sin(\theta/2) & 0 & \cos(\theta/2) \end{pmatrix} \tag{8.113}$$

is rotation by $\theta$ around the $x^2$ axis.

We can parametrize all $SU(2)$ elements by

$$u = e^{\phi T^3} e^{\theta T^2} e^{\psi T^3} \tag{8.114}$$

where

$$T^i = -\frac{\mathrm{i}}{2}\sigma^i \qquad 1 \le i \le 3 \tag{8.115}$$

The range of Euler angles that covers $SO(3)$ once is $0 \le \theta \le \pi$ with $\phi$ and $\psi$ identified modulo $2\pi$. Because $SU(2)$ is a double cover we should extend the range of $\phi$ or $\psi$ by a factor of 2 if we want to cover the group $SU(2)$ once. For example, taking:

$$\begin{aligned} &0 \le \theta \le \pi \\ &\phi \sim \phi + 2\pi \\ &\psi \sim \psi + 4\pi \end{aligned} \tag{8.116}$$

Then for generic $SU(2)$ elements we will have a unique representation

$$u = e^{\phi T^3} e^{\theta T^2} e^{\psi T^3} = \exp[-\frac{\mathrm{i}}{2}\phi\sigma^3]\exp[-\frac{\mathrm{i}}{2}\theta\sigma^2]\exp[-\frac{\mathrm{i}}{2}\psi\sigma^3] = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \alpha \end{pmatrix} \tag{8.117}$$

with

$$\alpha = e^{-\mathrm{i}\frac{1}{2}(\phi+\psi)}\cos(\theta/2) \qquad \beta = -e^{-\mathrm{i}\frac{1}{2}(\phi-\psi)}\sin(\theta/2) \tag{8.118}$$

The Euler angle coordinates on $SU(2)$ break down at $\theta = 0, \pi$. At $\theta = 0$ the product only depends on $(\phi + \psi)$ even though we have a three-dimensional manifold Similarly at $\theta = \pi$ the product only depends on $(\phi - \psi)$.

**Remark**: As we will discuss later, a good parametrization near the identity would be

$$u = \exp[\theta^k T^k] \tag{8.119}$$

where we are exponentiating the general element of the Lie algebra $\mathfrak{su}(2)$

---

**Exercise** *Polar Angle Decomposition Of* $SU(2)$

a.) Prove that every element of $SU(2)$ can be written in the form of (8.94).

b.) Express $\alpha, \beta$ in terms of $\chi$ and $\hat{n}$.

c.) The coordinates $\chi, \hat{n}$ cover a product of an interval and a two-dimensional sphere. Prove that $[0,\pi]\times S^2$ is not topologically the same as $SU(2)$. Where does the $\chi, \hat{n}$ coordinate system go bad?

---

---

**Exercise** *Polar Angle Parametrization Of SU(2) And Rotations*

a.) Show that if $u$ is parametrized as in (8.94) then

$$u\vec{x}\cdot\vec{\sigma}u^{-1} = \vec{y}\cdot\vec{\sigma} \tag{8.120}$$

where $\vec{y}$ is obtained from $\vec{x}$ by rotation by angle $2\chi$ around the $\hat{n}$ axis. [112]

b.) Show that

$$e^{\mathrm{i}\psi n\cdot\sigma}x\cdot\sigma e^{-\mathrm{i}\psi n\cdot\sigma} = y\cdot\sigma \tag{8.122}$$

with

$$y = \cos(2\psi)(n\times(x\times n)) + \sin(2\psi)n\times x + (n\cdot x)n \tag{8.123}$$

---

---

**Exercise** *A Basis For The Lie Algebra* $\mathfrak{su}(2)$

a.) Show that every traceless anti-Hermitian $2\times 2$ matrix is a real linear combination of the three matrices $T^i$ defined in (8.115).

b.) Show that the matrix commutators satisfy

$$[T^i, T^j] = \epsilon^{ijk}T^k \tag{8.124}$$

with the convention $\epsilon^{123} = +1$.

---

---

**Exercise**

Show that in the Euler angle parametrization the shift

$$\psi \rightarrow \psi + 2\pi \tag{8.125}$$

takes $u \rightarrow -u$.

---

[112] *Answer*: To show this note that if $\vec{x}$ is parallel to $\hat{n}$ then $\vec{x}\cdot\vec{\sigma}$ and $\hat{n}\cdot\vec{\sigma}$ commute then $\vec{x} = \vec{y}$. Now suppose that $\vec{x}$ is perpendicular to $\hat{n}$. Then

$$\begin{aligned} u\vec{x}\cdot\vec{\sigma}u^{-1} &= u^2(\vec{x}\cdot\vec{\sigma}) \\ &= (\cos(2\chi) + \mathrm{i}\sin(2\chi)\hat{n}\cdot\vec{\sigma})(\vec{x}\cdot\vec{\sigma}) \\ &= \cos(2\chi)(\vec{x}\cdot\vec{\sigma}) - \sin(2\chi)(\hat{n}\wedge\vec{x})\cdot\sigma \end{aligned} \tag{8.121}$$

where in the last line we used that $\hat{n}\cdot\vec{x} = 0$.

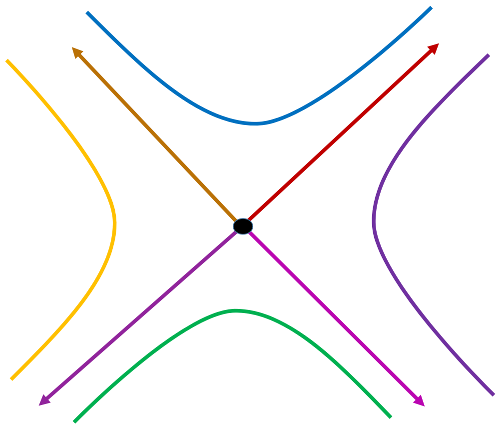


**Figure 9:** The distinct kinds of orbits of $SO_0(1,1,\mathbb{R})$ are shown in different colors. If we enlarge the group to include transformations that reverse the orientation of time and/or space then orbits of the larger group will be made out of these orbits by reflection in the space or time axis.

### 8.4.2 Extended Example: The Case Of $1+1$ Dimensions

Consider 1+1-dimensional Minkowski space with coordinates $x = (x^0, x^1)$ and metric given by

$$\eta := \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.126}$$

i.e. the quadratic form is $(x,x) = -(x^0)^2 + (x^1)^2$. The two-dimensional Lorentz group is defined by

$$O(1,1) = \{A | A^{tr} \eta A = \eta\} \tag{8.127}$$

This group acts on $\mathbb{M}^{1,1}$ preserving the Minkowski metric.

The connected component of the identity is the group of Lorentz boosts of rapidity $\theta$:

$$x^0 \to \cosh\theta\, x^0 + \sinh\theta\, x^1 \tag{8.128}$$

$$x^1 \to \sinh\theta\, x^0 + \cosh\theta\, x^1 \tag{8.129}$$

that is:

$$SO_0(1,1;\mathbb{R}) \equiv \{B(\theta) = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix} | -\infty < \theta < \infty\} \tag{8.130}$$

In the notation the $S$ indicates we look at the determinant one subgroup and the subscript 0 means we look at the connected component of 1. This is a group since

$$B(\theta_1)B(\theta_2) = B(\theta_1 + \theta_2) \tag{8.131}$$

so $SO_0(1,1) \cong \mathbb{R}$ as groups. Indeed, note that

$$B(\theta) = \exp\left[\theta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right] \tag{8.132}$$

It is often useful to define *light cone coordinates*: [113]

$$x^{\pm} := x^0 \pm x^1 \tag{8.133}$$

and the group action in these coordinates is simply:

$$x^{\pm} \to e^{\pm\theta} x^{\pm} \tag{8.134}$$

so it is obvious that $x^+ x^- = -(x,x)$ is invariant.

It follows that the orbits of the Lorentz group are, in general, hyperbolas. They are separated by different values of the Lorentz invariant $x^+x^- = \lambda$, but this is not a complete invariant, since the sign (or vanishing) of $x^+$ and of $x^-$ is also Lorentz invariant. For a real number $r$ define

$$\mathrm{sign}(r) := \begin{cases} +1 & r > 0 \\ 0 & r = 0 \\ -1 & r < 0 \end{cases} \tag{8.135}$$

Then $(\lambda, \mathrm{sign}(x^+), \mathrm{sign}(x^-))$ is a complete invariant of the orbits. That is, given this triple of data there is a unique orbit with these properties.

It is now easy to see what the different types of orbits there are. They are shown in Figure 9: They are:

♣Actually, the lightrays and hyperbolas have trivial stabilizer and hence are in the same strata. This is a problem with using strata. ♣

1. hyperbolas in the forward/backward lightcone and the left/right of the lightcone
2. 4 disjoint lightrays.
3. the origin: $x^+ = x^- = 0$.

It is now interesting to consider the orbits of the full Lorentz group $O(1,1)$ and its relation to the massless wave equations. But there are clearly elements of $O(1,1)$ not continuously connected to the identity such as:

$$P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad T = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.136}$$

In an exercise below you show that the Lorentz group $O(1,1)$ has four connected components

$$O(1,1) = \amalg_{(\sigma_1,\sigma_2)\in\mu_2\times\mu_2} O(1,1)_{(\sigma_1,\sigma_2)} \tag{8.137}$$

[113] Some authors will define these with a $1/2$ or $1/\sqrt{2}$. One should exercise care with this choice of convention.

where the connected component $O(1,1)_{(\sigma_1,\sigma_2)}$ consists of group elements of the form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{8.138}$$

with $sign(a) = \sigma_1$ and $sign(d) = \sigma_2$. We can write (noncanonically),

$$O(1,1) = SO_0(1,1) \amalg P \cdot SO_0(1,1) \amalg T \cdot SO_0(1,1) \amalg PT \cdot SO_0(1,1) \tag{8.139}$$

The $P$ and $T$ operations map various orbits of $SO_0(1,1)$ into each other: $P$ is a reflection in the time axis, i.e., a reflection of the spatial coordinate, while $T$ is a reflection in the space axis, i.e. a reflection of the time coordinate. Thus the orbits of the groups $SO(1,1)$, $SO_0(1,1) \amalg PT \cdot SO_0(1,1)$, and $O(1,1)$ all differ slightly from each other.

♣Should give more details here, or form an exercise. ♣

As an example of a physical manifestation of orbits let us consider the energy-momentum dispersion relation of a particle of mass $m$ with energy-momentum $(E,p) \in \mathbb{R}^{1,1}$.

1. Massive particles: $m^2 > 0$ have $(E,p)$ along an orbit in the upper quadrant:

$$\mathcal{O}^+(m) = \{(m\cosh\theta, m\sinh\theta) | \theta \in \mathbb{R}\} \tag{8.140}$$

2. Massless particles move at the speed of light. In 1+1 dimensions there is an interesting refinement of the massless orbits: Left-moving particles with positive energy have support on [114] $p_+ = \frac{1}{2}(E+p) = 0$ and $p_- = \frac{1}{2}(E-p) \neq 0$. Right-moving particles with positive energy have support on $p^- = 0$ and $p^+ \neq 0$. In $d+1$ dimensions with $d > 1$ the orbits of $SO_0(1,d)$ consisting of the forward and backward lightcones (minus the origin) are connected.

3. Tachyons have $E^2 - p^2 = m^2 < 0$ and have their support on the left or right quadrant. If we try to expand a solution to the wave-equation with $e^{\mathrm{i}(k_0 x^0 + k_1 x^1)}$ then $k_0^2 = \sqrt{k_1^2 + m^2}$ and so if the spatial momentum $k_1$ is sufficiently small then $k_0$ is pure imaginary and the wave grows exponentially, signaling and instability. This tells us our theory is out of control and some important new physical input is needed.

4. A massless "particle" of zero energy and momentum.

[114] Note the factors of two: We have

$$\begin{aligned} x^+ &= x^0 + x^1 \\ x^- &= x^0 - x^1 \end{aligned} \tag{8.141}$$

but

$$\begin{aligned} p_+ &= \frac{1}{2}(p_0 + p_1) \\ p_- &= \frac{1}{2}(p_0 - p_1) \end{aligned} \tag{8.142}$$

so that $x^0 p_0 + x^1 p_1 = x^+ p_+ + x^- p_-$. This is an example of the tricky factors of two one encounters when working with light-cone coordinates.

---

**Exercise** *Components Of The* $1+1$*-Dimensional Lorentz Group*

a.) Prove equation (8.137). [115]

b.) Show that the group of components of the Lorentz group is $O(1,1)/SO_0(1,1) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$

---

### 8.4.3 Orbits Of The Lorentz Group In $d > 2$ Dimensions As Homogeneous Spaces

We define $d$-dimensional Minkowski space $\mathbb{M}^{1,d-1}$ with $d > 2$ to be the vector space $\mathbb{R}^d$ with quadratic form

$$\eta = Diag\{-1, +1_{d-1}\} \tag{8.146}$$

and

$$O(1, d-1) = \{A | A^{tr} \eta A = \eta\} \tag{8.147}$$

There is a continuous surjective homomorphism

$$\varphi : O(1, d-1) \to \mu_2 \times \mu_2 \tag{8.148}$$

given by $\varphi(A) = (\kappa_1, \kappa_2)$. Here $\kappa_1 = \mathrm{sign}(\det A)$ while $\kappa_2$ is more subtle. It is $+1$ if $A$ preserves separately the two connected components of the light cones (minus the origin) and minus one if they are exchanged. Therefore, there is an exact sequence:

$$1 \to SO_0(1, d-1) \to O(1, d-1) \to \mu_2 \times \mu_2 \to 1 \tag{8.149}$$

We will show below, when analyzing $O(p,q)$ that as a manifold $O(1, d-1)$ has four connected components. So the kernel of $\varphi$ is the connected component of the identity. By

---

[115] *Answer*: The general matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{8.143}$$

is in $O(1,1)$ iff

$$\begin{aligned} a^2 - c^2 &= 1 \\ d^2 - b^2 &= 1 \\ ab &= cd \end{aligned} \tag{8.144}$$

The most general solution of the first two equations is

$$\begin{aligned} a &= \kappa_1 \cosh\theta \\ c &= \sinh\theta \\ d &= \kappa_2 \cosh\theta' \\ b &= \sinh\theta' \end{aligned} \tag{8.145}$$

where $\kappa_i \in \{\pm 1\}$ and $\theta, \theta' \in \mathbb{R}$. Now impose the third equation. The solutions split into two cases: If $\kappa_1/\kappa_2 = 1$ then $\theta = \theta'$. This gives two components. If $\kappa_1/\kappa_2 = -1$ then $\theta = -\theta'$, giving the other two components.

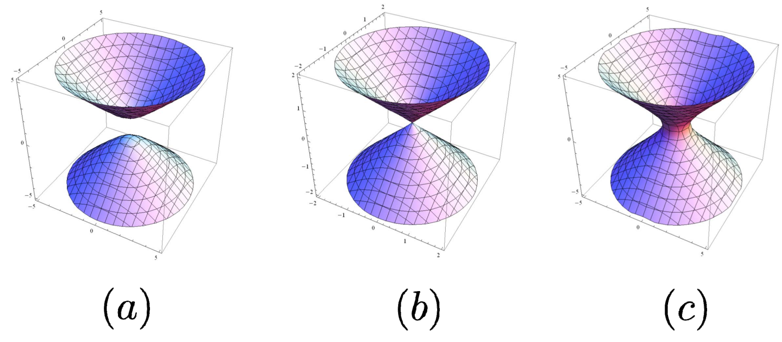


**Figure 10:** Illustrating orbits of the connected component of the identity in $O(1,3)$. In (a) the top and bottom hyperboloids are separate orbits, and if we include time-reversing transformations the orbits are unions of the two hyperboloids. In (b) there are three orbits shown with $x^0 > 0$ $x^0 < 0$ (the future and past, or forward and backward light cones), and the orbit consisting of the single point. In (c), once $x^2$ has been specified, there is just one orbit, for $d > 2$.

making (noncanonical) choices of $T$ and $P$ we can write:

$$O(1,d-1) = SO_0(1,d-1) \amalg P \cdot SO_0(1,d-1) \cdot T \cdot SO_0(1,d-1) \cdot PT \cdot SO_0(1,d-1) \quad (8.150)$$

where $T$ reverses time orientation and $P$ reverses space orientation.

The nature of the orbits of $SO_0(1,d-1)$ and $O(1,d-1)$ is slightly different from the $1+1$ dimensional case because of the zero-dimensional sphere $S^0$ is disconnected but the higher dimensional spheres are connected.

1. For $\lambda \in \mathbb{R}^*$ we can define the orbit of timelike vectors:

$$\mathcal{O}_{\text{timelike}}(\lambda) = \{x|(x^0)^2 - (\vec{x})^2 = \lambda^2 \qquad \& \qquad \text{sign}(x^0) = \text{sign}(\lambda)\} \quad (8.151)$$

   By the stabilizer-orbit theorem we can identify this with

$$\mathcal{O}_{\text{timelike}}(\lambda) \cong SO_0(1,d-1)/SO(d-1) \quad (8.152)$$

   by considering the isotropy group at $(x^0 = \lambda, \vec{x} = 0)$. See Figure 10(a).

2. For $\mu^2 > 0$ we can define

$$\mathcal{O}_{\text{spacelike}}(\mu^2) = \{x|(x^0)^2 - (\vec{x})^2 = -\mu^2\} \quad (8.153)$$

   By the stabilizer-orbit theorem we can identify this with

$$\mathcal{O}_{\text{spacelike}}(\mu^2) \cong SO_0(1,d-1)/SO_0(1,d-2) \quad (8.154)$$

by considering the isotropy group at $x = (x^0 = 0, x^1 = 0, \dots, x^{d-2} = 0, x^{d-1} = \mu)$. Unlike $1+1$ dimensions, the sign of $\mu$ does not distinguish different orbits for $d > 2$ because the sphere $S^{d-2}$ is connected. See Figure 10(c).

3.
$$\mathcal{O}^{\pm}_{\text{null}} = \{x | x^2 = 0 \qquad \& \qquad \text{sign}(x^0) = \pm 1\} \tag{8.155}$$

Vectors in this orbit are of the form $(x^0, |x^0|\hat{n})$ where $\hat{n} \in S^{d-2} \subset \mathbb{R}^{d-1}$ and the sign of $x^0$ is invariant under the action of the identity component of $O(1,3)$. (Show this!). We can think of $\hat{n} \in S^{d-2}$ as parametrizing the directions of light-rays. That is, the point where the light ray hits the celestial sphere. As mentioned above, for $d = 2$ the sphere $S^0$ has two disconnected components, leading to an $SO_0(1, d-1)$-invariant distinction between left- and right-movers. In one spatial dimension, a light ray either moves left or right, and this is a boost-invariant concept. In $d - 1 > 1$ spatial dimensions, we can rotate any direction of light ray into any other. See Figure 10(b). One can show that these orbits too are homogeneous spaces: [116]

$$\mathcal{O}^{\pm} \cong SO_0(1, d-1)/\mathcal{I} \tag{8.156}$$

4. The final orbit is of course $\{x = 0\}$.

**Remarks**

1. *Discrete Symmetries In Nature.* One can show, in a relativistic quantum theory that if a theory is invariant under infinitesimal Lorentz symmetries, then it is invariant under the connected component of the identity $SO_0(d-1, 1)$, because if the Lie algebra is represented by well-defined anti-unitary operators then so are the one-parameter subgroups generated by elements of the Lie algebra.

   However, it turns out that relativistic QFT's that have $SO_0(d-1, 1)$ symmetry can nevertheless fail to be invariant under the disconnected components of $O(d-1, 1)$.

   An example of a physical theory that IS invariant under all four components of $O(3, 1)$ is electromagnetism. Consider the classical equations in vacuum:

$$\begin{aligned} dF &= 0 \\ d * F &= 0 \end{aligned} \tag{8.157}$$

[116]The isotropy group of a light ray is $\mathcal{I} \cong ISO(d-2)$, where $ISO(d-2)$ is the Euclidean group on $\mathbb{R}^{d-2}$. The easiest way to show this is to use the Lie algebra of $so(1, d-1)$ and work with light-cone coordinates. Choosing a direction of the light ray along the $x^{d-1}$ axis and introducing light-cone coordinates $x^{\pm} := x^0 \pm x^{d-1}$, and transverse coordinates $x^i$, $i = 1, \dots, d-2$ if the lightray satisfies $x^- = 0$ then we have unbroken generators $M^{+i}$ and $M^{ij}$.

the equations are invariant under arbitrary isometries of the Lorentz metric. Written in terms of electric and magnetic fields in $3+1$ dimensions these equations become:

$$\begin{aligned} \nabla\cdot B &= 0 \\ \nabla\times E + \frac{1}{c}\frac{\partial B}{\partial t} &= 0 \\ \nabla\cdot E &= 0 \\ \nabla\times B - \frac{1}{c}\frac{\partial E}{\partial t} &= 0 \end{aligned} \tag{8.158}$$

invariance under $P$ means that if $(E(x,t),B(x,t))$ is a solution of the equations then so is $(-E(-x,t),B(-x,t))$. Invariance under $T$ means that $(E(x,-t),-B(x,-t))$ is a solution. This is enough to show that the classical theory is invariant in the absence of sources. It can be generalized to include sources. With some further work these arguments extend to the full quantum theory and in fact to QED (quantum electrodynamics). See, for example, S. Weinberg, *The Quantum Theory Of Fields, vol. 1* for a careful discussion of discrete symmetries in QFT.

♣Tricky point here about compatibility of general law of field transformations. The physical $F$ transforms as a twisted differential form under the time-reversing symmetries. ♣

Parity invariance means that, if you watch a video of a physical process, then you cannot tell whether you are looking at that process in a mirror, or not. For a long time it was assumed that parity is a symmetry not only of the electromagnetic and gravitational force but also of the nuclear forces, including the weak force responsible for nuclear beta decay where a neutron decays (possibly within a nucleus) according to the beta decay process:

$$n \to p + e^- + \bar{\nu}_e \ . \tag{8.159}$$

However, in 1956 T.D. Lee and C.N. Yang carefully reviewed the evidence for parity conservation in nature and pointed out that there was - at the time - no experimental test demonstrating that the weak interactions are parity invariant. Lee and Yang proposed some experimental tests of parity invariance of the weak nuclear force. Shortly thereafter C.S. Wu and her collaborators demonstrated experimentally in 1956 that parity is indeed violated by the weak nuclear force. For this remarkable achievement Lee and Yang [117] were awarded the 1957 Nobel prize in physics.

The basic idea of the Lee-Yang-Wu experiment is that the neutron has spin - which is invariant under parity - but in the beta decay process (8.159) electrons are preferentially emitted in one direction relative to the spin. (As it turns out, in our universe, they are preferentially emitted in the direction opposite to the spin vector of the neutron.) It was not practical to observe the process (8.159) directly with free neutrons so the Wu experiment looked at the decay of an unstable isotope of cobalt ${}^{60}_{27}\mathrm{Co}$ to an excited state of nickel ${}^{60}_{28}\mathrm{Ni}^*$, which subsequently decays electromagnetically to the stable isotope ${}^{60}_{28}\mathrm{Ni}$ by emitting two photons. The cobalt can be put in a definite spin state with a magnetic field (at low temperature) and then, since electromagnetism is parity invariant the photons will not be preferentially emitted along or against the

[117] but not C.S. Wu. Many physicists, including the author, consider this an outrage.

spin axis. One can then compare the photons with the electrons. It was found that the electrons are preferentially emitted in one direction.

Parity nonconservation leads to nontrivial role for chirality in the standard model and is a profound and fundamental aspect of nature.

Given the violation of parity by the weak nuclear force one naturally wonders about time reversal invariance. The question of time reversal invariance of a physical system can be phrased as follows: If you run a movie of a physical process backwards, is the resulting process physically possible. This is again true of classical electrodynamics and gravitation, as well as of QED. It is important here to distinguish between events that are extremely unlikely from those which are physically impossible. According to the laws of electromagnetism and gravity, the perfume in a room could collect itself into a small bottle. If you watched a movie of such an event you could say that - with high probability - the movie was run backwards. But, just based on the laws of QED and gravity, you could not say this with absolute and total certainty. In 1964 V. Fitch and J. Cronin discovered that in certain very rare processes in nature known as the decay of neutral Kaons, the decays actually do violate time-reversal invariance. That is a good thing, because, as noted by A. Sakharov, if there were no time-reversal invariance in the laws of physics it would be impossible to understand why there is matter/anti-matter asymmetry in the context of the big bang theory.

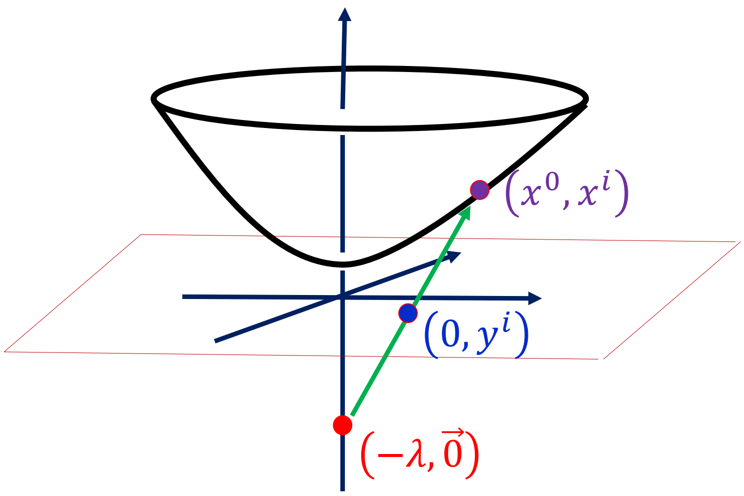



**Figure 11:** The picture shows $\mathcal{O}_{\text{timelike}}(\lambda)$ for $\lambda > 0$. Stereographic projection involves finding the point with $x^0 = 0$ on the line between a point on the orbit and the point $(-\lambda, \vec{0})$. This projects the orbit to the ball $\vec{y}^2 < \lambda^2$ in $d - 1$ dimensions .

### 8.4.4 Lorentzian And Euclidean (anti)-deSitter Spaces As Homogeneous Spaces

There are four important solutions of the Einstein equations that are related to homogeneous spaces involving Lorentz groups.

*Spheres a.k.a. Euclidean deSitter Space*

The sphere of radius $R$ in $\mathbb{R}^{n+1}$ admits a transitive action of $SO(n+1)$, and the stabilizer group of any point is isomorphic to $SO(n)$. So, as $G$-spaces we have an isomorphism

$$S^n \cong SO(n+1)/SO(n) \tag{8.160}$$

But more is true: $S^n$ inherits a metric from the Euclidean metric of the ambient Euclidean spaces

$$ds^2 = dX_{n+1}^2 + \sum_{i=1}^n dX_i^2 \tag{8.161}$$

The "pullback" to the solution set of

$$X_{n+1}^2 + \sum_{i=1}^n X_i^2 = R^2 \tag{8.162}$$

Gives the round metric on $S^n$ of radius $R$. It is a solution of the Einstein equations and can be considered to be a "Wick rotation of deSitter space."

*Hyperbolic Space a.k.a. Euclidean Anti-deSitter Space.*

The orbit $\mathcal{O}^+(R)$ above is the component of the solution space of the equation

$$-X_0^2 + \sum_{i=1}^n X_i^2 = -\lambda^2 \tag{8.163}$$

with $X_0 > 0$. It inherits a Euclidean signature metric from the ambient metric

♣Make notation $R, \lambda, \Lambda$ uniform. ♣

$$ds^2 = -dX_0^2 + \sum_{i=1}^n dX_i^2 \tag{8.164}$$

This gives one of many models of hyperbolic space, often known in physics as "Euclidean anti-deSitter space." The restriction of the metric to the orbit gives one model of the hyperbolic metric. We parametrize

$$\begin{aligned} x^0 &= \lambda \cosh\theta \\ x^i &= \lambda \sinh\theta n^i \end{aligned} \tag{8.165}$$

where $n^i$ is a vector on the unit sphere $S^{d-2} \subset \mathbb{R}^{d-1}$ and the metric becomes

$$ds^2 = d\theta^2 + \sinh^2(\theta) ds^2_{S^{d-2}} \tag{8.166}$$

where $ds^2_{S^{d-2}}$ is the metric on the unit $(d-2)$-dimensional sphere induced from the Euclidean metric in $\mathbb{R}^{d-1}$.

There are other very useful models for hyperbolic space. One proceeds by using stereographic projection. See Figure 11. A point $(x^0, x^i) \in \mathcal{O}_{\rm timelike}(\lambda)$ is projected to the point $(0, y^i)$ in $\mathbb{R}^{d-1}$ with

$$y^i = \lambda \frac{x^i}{\lambda + x^0} \tag{8.167}$$

Simple algebra shows that

$$\vec{y}^2 = \lambda^2 \frac{x^0 - \lambda}{x^0 + \lambda} = \lambda^2 \left(1 - \frac{2\lambda}{x^0 + \lambda}\right) < \lambda^2 \tag{8.168}$$

So the hyperbola projects to the interior of the $(d-1)$-dimensional ball (aka disk) of radius $\lambda$. The inverse transformation from the ball to the hyperbola is

$$\begin{aligned} x^0 &= \lambda \frac{1 + \vec{y}^2/\lambda^2}{1 - \vec{y}^2/\lambda^2} \\ x^i &= \frac{2}{1 - \vec{y}^2/\lambda^2} y^i \end{aligned} \tag{8.169}$$

We can now pull back the metric $-(dx^0)^2 + \sum_{i=1}^{d-1}(dx^i)^2$ to the hyperbola, which in this case amounts to substitution of the above formulae and taking derivatives to produce

$$ds^2 = 4 \frac{(d\vec{y})^2}{(1 - \vec{y}^2/\lambda^2)^2} \tag{8.170}$$

This is known as the Poincaré disk model for hyperbolic space. The "boundary sphere" $\vec{y}^2 = \lambda^2$ is at infinite distance.

There is a second method of projection from a point at infinity. Select one of the coordinates $x^i$ as special. We will take $x^{d-1}$, and denote the remaining "spatial" coordinates as $x^\alpha$, with $\alpha = 1, \dots, d-2$. (We are working with the case $d > 2$ here.) Set

♣Describe geometrical interpretation ♣

$$\begin{aligned} s^\alpha &= \lambda \frac{x^\alpha}{x^0 - x^{d-1}} \qquad & \alpha = 1, \dots, d-2 \\ z &= \lambda^2 \frac{1}{x^0 - x^{d-1}} > 0 \end{aligned} \tag{8.171}$$

The inverse transformation is:

$$\begin{aligned} x^\alpha &= \lambda \frac{s^\alpha}{z} \qquad & \alpha = 1, \dots, d-2 \\ x^0 - x^{d-1} &= \frac{\lambda^2}{z} \\ x^0 + x^{d-1} &= \frac{z^2 + \vec{s}^2}{z} \end{aligned} \tag{8.172}$$

Now pulling back the Lorentz metric gives the formula

$$ds^2 = \lambda^2 \frac{(dz)^2 + (ds^\alpha)^2}{z^2} \tag{8.173}$$

This is known as the Poincaré upper half plane model of hyperbolic space. The $\mathbb{R}^{d-2}$ plane at infinity $z = 0$ is mapped to the sphere $S^{d-2}$ at infinity when transforming from the plane to the disk.

*Lorentzian Anti-deSitter Space*

Consider the orbit of $O(2,n)$ on the hyperbola

$$-X_0^2 - X_{n+1}^2 + \sum_{i=1}^{n} X_i^2 = -\Lambda^2 \tag{8.174}$$

with $\Lambda$ real. Such an orbit will be a homogeneous space $O(2,n)/O(1,n)$ and the induced metric from

♣Careful here about components and covers. ♣

$$-dX_0^2 - dX_{n+1}^2 + \sum_{i=1}^{n} dX_i^2 \tag{8.175}$$

will be the Lorentz-signature AdS metric.

*Lorentzian deSitter Space*

Consider the orbit of $O(2,n)$ on the hyperbola

$$-X_0^2 + \sum_{i=1}^{n} X_i^2 = -\Lambda^2 \tag{8.176}$$

with $\Lambda$ real. Such an orbit will be a homogeneous space $O(1,n)/O(1,n-1)$ and the induced metric from

♣Careful here about components and covers. ♣

$$-dX_0^2 + \sum_{i=1}^{n} dX_i^2 \tag{8.177}$$

will be the Lorentz-signature deSitter metric.

*************************************

MORE DETAILS HERE. SEE GMP 2002, Ch. 3, Section 6

*************************************

### 8.4.5 Grassmannians

A very nice application of the Stabilizer-Orbit theorem is to the description of Grassmannians of a vector space as homogeneous spaces.

Let us recall some facts about complex projective space $\mathbb{CP}^n$. It can be described in several different ways.

♣Move this to the CPn section above? ♣

1. It is the space of orbits $(\mathbb{C}^{n+1} - \{0\})/\mathbb{C}^*$. Hence it is the space of equivalence classes $[z_1 : \cdots : z_{n+1}] = [\lambda z_1 : \cdots : \lambda z_{n+1}]$ with $(z_1, \dots, z_{n+1}) \neq 0$ and $\lambda \in \mathbb{C}^*$.

2. It is the moduli space of one-dimensional subspaces of $\mathbb{C}^{n+1}$.

3. It is the space of one-dimensional orthogonal projection operators

4. It is the space of pure states in a Hilbert space $\mathbb{C}^{n+1}$.

5. It is a homogeneous space $GL(n+1,\mathbb{C})/P$ and $SU(n+1)/(SU(n) \times U(1)$.

We will begin by generalizing the second point of view above.

Consider a finite dimensional vector space $V$, say of dimension $n$ and let $0 < k < n$ be an integer and define $Gr_k(V)$ to be the set of all $k$-dimensional linear subspaces of $V$. It is not hard to see that $GL(V)$ acts transitively on this space: If $W \subset V$ is a $k$-dimensional subspace and $T \in GL(V)$ then $T(W) = \{T(w)|w \in W\}$ is a $k$-dimensional subspace. It is an easy fact of linear algebra that for any two $k$-dimensional subspaces $W_1, W_2 \subset V$ there is a $T$ with $T(W_1) = W_2$, i.e. the action is transitive. Therefore, by the stabilizer orbit theorem there is an isomorphism of $G$-sets:

$$Gr_k(V) \cong GL(V)/Stab_{GL(V)}(W_0) \tag{8.178}$$

for any $k$-dimensional subspace $W_0 \subset V$.

To compute the stabilizer of some vector space $W_0$ choose an ordered basis $v_1, \dots, v_k$ for $W_0$ and any complementary basis for $V$ so that

$$v_1, \dots, v_k, u_1, \dots, u_{n-k} \tag{8.179}$$

is an ordered basis for $V$. Now, what is the subgroup of $T$ so that $T(W_0) = W_0$ ? By definition:

$$\begin{aligned} T(v_i) &= A_{ji}v_j + C_{\alpha i}u_\alpha \\ T(u_\alpha) &= B_{j\alpha}v_j + D_{\beta\alpha}u_\beta \end{aligned} \tag{8.180}$$

The condition $T(W_0) = W_0$ is then the condition that $C = 0$ for the matrix of $T$ relative to such a basis. So the stabilizer group is isomorphic to the subgroup $P$ of $GL(n, \kappa)$ with $C = 0$. That is:

$$P_{W_0} := \mathrm{Stab}_{GL(V)}(W_0) \cong P := \{g = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in GL(n, \kappa)\} \tag{8.181}$$

So we can identity

$$Gr_k(V) \cong GL(n, \kappa)/P \tag{8.182}$$

as $GL(V)$-spaces. The basepoint $P$ on the RHS corresponds to the subspace $W_0$.

In fact, the Grassmannian is a manifold and the representation in terms of homogeneous coordinates helps us to find local coordinates. The basic idea for finding local coordinates is that we try to find a unique representative $g_0$ in the coset $gP$. That is $g_0$ will satisfy some special conditions so that if $g_0$ and $g_0'$ both satisfy those conditions then

$$g_0 P = g_0' P \tag{8.183}$$

is sufficient to imply that $g_0 = g_0'$. One such condition is that $g_0$ should be of the form:

$$g_0 = \begin{pmatrix} 1_{k\times k} & 0 \\ \gamma_{(n-k)\times k} & 1_{(n-k)\times(n-k)} \end{pmatrix} \tag{8.184}$$

for some matrix $\gamma_{(n-k)\times k} \in Mat_{(n-k)\times k}(\kappa)$. Note that if we right multiply by an element of $P$ and ask that

$$\begin{pmatrix} 1 & 0 \\ \gamma & 1 \end{pmatrix} \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \gamma' & 1 \end{pmatrix} \tag{8.185}$$

Then $B = 0$ and $A = 1$ and $D = 1$ and hence $\gamma' = \gamma$.

**Remark**: When trying to represent cosets $gP$ by group elements $g$ we can view $g \to gp$ with $p \in P$ as a "gauge transformation." So, what we are doing here is trying to "fix the gauge" by choosing a condition such as (8.184) for the representative. There are many other ways of fixing a gauge freedom. We just made one choice above. In the quotient topology on $GL(V)/P$ the coordinates $\gamma$ are continuous. One could go further and produce an atlas and charts. Then they would be smooth coordinates.

We can try to impose the gauge freedom for cosets $gP$ where $gP$ is "not too far" from the identity $P$. "Not to far" means more precisely: that there is a representative $gP$ such that if $g$ in block diagonal form

$$g = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{8.186}$$

then $\alpha$ and $\delta - \gamma\alpha^{-1}\beta$ are invertible. These are precisely the conditions so that we can solve the equation

$$g \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \tilde{\gamma} & 1 \end{pmatrix} \tag{8.187}$$

for some $\tilde{\gamma} \in Mat_{(n-k)\times k}(\kappa)$.

Indeed, equation (8.187) implies that

$$\begin{aligned} \alpha A &= 1 \\ \gamma A &= \tilde{\gamma} \\ \alpha B + \beta D &= 0 \\ \gamma B + \delta D &= 1 \end{aligned} \tag{8.188}$$

These should be regarded as equations for $A, B, D, \tilde{\gamma}$ given $\alpha, \beta, \gamma, \delta$. If $g$ is "not too far" from the identity in the above sense then the solution is:

$$\begin{aligned} A &= \alpha^{-1} \\ B &= -\alpha^{-1}\beta(\delta - \gamma\alpha^{-1}\beta)^{-1} \\ D &= (\delta - \gamma\alpha^{-1}\beta)^{-1} \\ \tilde{\gamma} &= \gamma\alpha^{-1} \end{aligned} \tag{8.189}$$

The conclusion we can draw from this is that a neighborhood of $P$ in $GL(n,\kappa)/P$ consists of cosets $gP$ where any representative has $\alpha$ and $\delta - \gamma\alpha^{-1}\beta$ invertible. Coordinates in this neighborhood are then given by

$$\gamma \in Mat_{(n-k)\times k}(\kappa) \tag{8.190}$$

Under the isomorphism $Gr_k(V) \cong GL(n,\kappa)/P$ defined by a choice of $W_0$ we get a set of coordinates in the neighborhood of $W_0 \in Gr_k(V)$.

**Remark**: It is possible to push this idea further and produce explicit coordinate charts and transition functions. Let $\{e_1, \dots, e_n\}$ denote a basis for $V$ and let $J$ stand for a subset of $k$ integers between 1 and $n$ with $i_1 < \cdots < i_k$. Then the coordinate charts described above form an atlas if we consider just the charts around the subspaces $W_J$ spanned by $\{e_{i_1}, \dots, e_{i_k}\}$. The matrix groups $P_{W_J}$ will be related to each other by conjugation by permutation matrices.

♣Add material from Physics511-2014 ch. 2. Coordinate charts and transition functions. More material on Schubert cells, Plucker coordinates etc. ♣

*********************

MORE: GROUP ACTION ON RECTANGULAR MATRICES TO GET HOMOGENEOUS SPACE. COMPACT HOMOGENEOUS SPACE WITH ON BASES.

$\mathcal{S}_k(N) \cong U(N)/U(N-k)$: $k$-dimensional subspaces with ordered ON basis

$Gr_k(N) \cong U(N)/(U(N-k) \times U(k))$. In particular, it is compact and connected.

*********************

### 8.4.6 The Grassmannian Of Positive Definite Subspaces

There is an important generalization of the Grassmanian. We consider $\mathbb{R}^n$ with $n = p + q$ equipped with an indefinite metric of the form:

$$ds^2_{p,q} := (dx^1)^2 + \cdots (dx^p)^2 - (dy^1)^2 - \cdots (dy^q)^2 \tag{8.191}$$

The vector space together with this quadratic form is denoted $\mathbb{R}^{p,q}$. This metric is invariant under the generalization of the Lorentz group:

$$O(p,q;\mathbb{R}) := \{A \in GL(p+q;\mathbb{R}) | A^{tr} d_{p,q} A = d_{p,q}\} \tag{8.192}$$

where

$$d_{p,q} := \begin{pmatrix} 1_{p\times p} & 0 \\ 0 & -1_{q\times q} \end{pmatrix} \tag{8.193}$$

This group generalizes the rotation-reflection groups $O(n)$, the Lorentz groups $O(1,n)$ and the conformal groups $O(2,n)$.

We now let $\mathrm{Gr}^+_{p,q}$ denote the space of all $p$-dimensional subspaces $W \subset \mathbb{R}^{p,q}$ so that the restriction of the metric $ds^2_{p,q}$ is positive on $W$. This means that for any vector $\mathbf{w} \in W$ we have

$$\mathbf{w}^{\perp} d_{p,q} \mathbf{w} \geq 0 \tag{8.194}$$

with equality only for $\mathbf{w} = 0$.

**Example**: Consider $p = q = 1$. Then the positive definite subspaces are the spacelike lines through the origin. The space of such lines is naturally identified with the arc between the lightcones. Note that it is contractible.

♣EXPLAIN MORE. A FIGURE HERE WOULD BE NICE: LIGHTCONE AND AN ARC BETWEEN THE LIGHT CONES. ♣

Let us now consider the general case. One obvious positive subspace is:

$$\mathbb{R}^{p,0} = \{(\vec{x}, \vec{0}) \in \mathbb{R}^{p,q} | \vec{x} \in \mathbb{R}^p\} \tag{8.195}$$

Note that

$$\mathbb{R}^{p,q} \cong \mathbb{R}^{p,0} \oplus \mathbb{R}^{0,q} \tag{8.196}$$

is an orthogonal decomposition into a positive definite and negative definite subspace. In a similar way, if $W \subset \mathrm{Gr}^{+}_{p,q}$ then the orthogonal space defined by

$$W^{\perp} = \{\mathbf{u} \in \mathbb{R}^{p,q} | \mathbf{u}^{tr} d_{p,q} \mathbf{w} = 0 \qquad \forall \mathbf{w} \in W\} \tag{8.197}$$

is such that

$$\mathbb{R}^{p,q} = W \oplus W^{\perp} \tag{8.198}$$

and a consequence of the computation we are about to do is that $W^{\perp}$ is negative definite.

Now, we can give a very nice parametrization of points in $\mathrm{Gr}^{+}_{p,q}$ as follows. If $W \subset \mathbb{R}^{p,q}$ is a positive definite $p$-dimensional subspace then it must be the graph of a linear transformation: $C : \mathbb{R}^p \to \mathbb{R}^q$:

$$W = \{(x, Cx) | x \in \mathbb{R}^p\} \tag{8.199}$$

We can think of $C \in Mat_{q\times p}(\mathbb{R})$. The reason is that if $(x, y)$ and $(x, y')$ are two vectors in $W$ then $(0, y - y')$ would have to be in $W$, but then if $y - y' \neq 0$ the vector would have negative definite square norm. So $y = y'$ so there is a unique vector $y$ for each $x$, and since $W$ is a linear subspace it must be the graph of a linear function.

Now, since $W$ is positive definite it must be that if $x \neq 0$ then

$$\parallel x \parallel_p^2 - \parallel Cx \parallel_q^2 > 0 \qquad \Rightarrow \qquad \parallel x \parallel_p^2 > \parallel Cx \parallel_q^2 = (x, C^{tr} C x)_p \tag{8.200}$$

We call such a linear transformation *contractive*.

Note that $C^{tr}C$ is real and symmetric and hence diagonalizable, and has an ON basis of eigenvectors $v_i$ with eigenvalues $0 \leq \lambda_i < 1$. It follows that we can define an unambiguous square root $(1 - C^{tr}C)^{-1/2}$ and now we can check that

$$\mathbf{v}_i := \{\Big((1 - C^{tr}C)^{-1/2} v_i, C(1 - C^{tr}C)^{-1/2} v_i\Big) | i = 1, \dots, p\} \tag{8.201}$$

is an ON basis for $W$:

$$\mathbf{v}_i^{tr} d_{p,q} \mathbf{v}_j = \delta_{i,j} \ . \tag{8.202}$$

The nonzero eigenvalues of $C^{tr}C$ are the same as those of $CC^{tr}$. Therefore if we choose an ON basis $w_a$, $a = 1, \dots, q$ of eigenvectors of the symmetric $q \times q$ real symmetric matrix $CC^{tr}$ we can similarly form the vectors:

$$\mathbf{w}_a := \{\Big(C^{tr}(1 - CC^{tr})^{-1/2} w_a, (1 - CC^{tr})^{-1/2} w_a\Big) | a = 1, \dots, q\} \tag{8.203}$$

We can now compute:

$$\mathbf{v}_i^{tr} d_{p,q} \mathbf{w}_a = 0 \tag{8.204}$$

and hence $\mathbf{w}_a$ form a basis for $W^{\perp}$. Moreover, we can also compute

$$\mathbf{w}_a^{tr} d_{p,q} \mathbf{w}_b = -\delta_{a,b} \tag{8.205}$$

Therefore $W^{\perp}$ is a negative definite subspace, as asserted above.

As we have just seen, $(\mathbf{v}_1, \dots, \mathbf{v}_p, \mathbf{w}_1, \dots, \mathbf{w}_q)$ is an ON basis for $\mathbb{R}^{p,q}$ and hence the matrix that relates it to the standard ON basis of $\mathbb{R}^{p,0} \oplus \mathbb{R}^{0,q}$ must be in $O(p,q)$. That is

$$H_C := \begin{pmatrix} (1 - C^{tr}C)^{-1/2} & C^{tr}(1 - CC^{tr})^{-1/2} \\ C(1 - C^{tr}C)^{-1/2} & (1 - CC^{tr})^{-1/2} \end{pmatrix} \in O(p,q) \tag{8.206}$$

and moreover $H_C$ takes the standard decomposition (8.196) to (8.198). Thus the action of $O(p,q)$ is transitive and the Grassmannian of positive definite subspaces is a homogeneous space:

$$\mathrm{Gr}^+_{p,q} \cong O(p,q)/(O(p) \times O(q)) \tag{8.207}$$

Let $\mathcal{H}_{p,q}$ be the set of matrices $H_C$ for contractive linear maps $C : \mathbb{R}^p \to \mathbb{R}^q$. We can also identity $\mathcal{H}_{p,q}$ with the Grassmannian. On the other hand, the stabilizer of (8.196) within $O(p,q)$ is clearly $O(p) \times O(q)$. We conclude that

$$O(p,q) = \mathcal{H}_{p,q} \cdot (O(p) \times O(q)) \tag{8.208}$$

Since the space of contractive linear transformations is a contractible space we see that $O(p,q)$ contracts onto $O(p) \times O(q)$, which therefore has four components, so long as both $p > 0$ and $q > 0$.

**Example** For $O(1,1)$ $C$ is $v/c$ of the Lorentz boost. "Contractive" means that all speeds are less than the speed of light. More generally, for $O(n,1)$, $C : \mathbb{R} \to \mathbb{R}^n$ may be identified with the velocity vector in units of the speed of light: $\beta_i = v_i/c$. Note that $C^\dagger C = \vec{\beta}^2$, while $CC^\dagger = \beta\beta^{tr}$ is a rank one $n \times n$ matrix. Define a projection operator onto the "boost direction" by $\Pi = \hat{\beta}\hat{\beta}^{tr}$ where $\hat{\beta}^i$ is the unit vector in $\mathbb{R}^n$ parallel to $\beta^i$. This makes it easy to compute

$$\begin{aligned} (1 - CC^\dagger)^{-1/2} &= \exp[-\frac{1}{2}\log(1 - CC^\dagger)] \\ &= \exp[-\frac{1}{2}\log(1 - \parallel \beta \parallel^2)\Pi] \\ &= (1 - \Pi) + \frac{1}{\sqrt{1 - \parallel \beta \parallel^2}}\Pi \end{aligned} \tag{8.209}$$

Denoting the rapidity by $\theta$ we thus have

$$H_C = \begin{pmatrix} \cosh\theta & \sinh\theta\hat{\beta}^{tr} \\ \sinh\theta\hat{\beta} & (1 - \hat{\beta}\hat{\beta}^{tr}) + \cosh\theta\hat{\beta}\hat{\beta}^{tr} \end{pmatrix} \tag{8.210}$$

---

**Exercise** $H_C$ *As A Matrix Generalization Of A Boost*
Suppose that $p < q$ and that $C$ is of the form

$$C = Diag\{\mu_1, \dots, \mu_p\} \oplus 0_{p \times (q-p)} \tag{8.211}$$

where $0 \le |\mu_i| < 1$. Put $\mu_i = \tanh\theta_i$.

Show that $C$ is of the form

$$H_C = \Pi\Biggl( 1_{(q-p)\times(q-p)} \oplus \Biggl( \oplus_i \begin{pmatrix} \cosh\theta_i & \sinh\theta_i \\ \sinh\theta_i & \cosh\theta_i \end{pmatrix} \Biggr) \Biggr)\Pi^{-1} \tag{8.212}$$

where $\Pi$ is a permutation matrix.

---

---

**Exercise** $U(p,q)$

Define $\mathbb{C}^{p,q}$ to be the $(p+q)$-dimensional complex vector space with sesquilinear form $\langle z_1, z_2\rangle := z_1^\dagger d_{p,q} z_2$. Let $U(p,q)$ be the automorphism group of this form. In equations:

$$U(p,q) := \{A \in GL(p+q,\mathbb{C}) | A^\dagger d_{p,q} A = d_{p,q}\} \tag{8.213}$$

Give an analog of the above discussion for this and in particular show that the Grassmannian of positive definite $p$-dimensional subspaces is isomorphic to the homogeneous space

$$U(p,q)/(U(p)\times U(q)) \tag{8.214}$$

---

---

**Exercise** *Exponential Parametrization Of The Grassmannian*

Suppose that $\pi \in Mat_{k\times(n-k)}(\kappa)$, with $\kappa = \mathbb{R}, \mathbb{C}$. Show that:

$$\exp[\begin{pmatrix} 0 & \pi \\ \pi^\dagger & 0 \end{pmatrix}] = \begin{pmatrix} \cosh(\sqrt{\pi\pi^\dagger}) & \pi\frac{\sinh(\sqrt{\pi^\dagger\pi})}{\sqrt{\pi^\dagger\pi}} \\ \pi^\dagger\frac{\sinh(\sqrt{\pi\pi^\dagger})}{\sqrt{\pi\pi^\dagger}} & \cosh(\sqrt{\pi^\dagger\pi}) \end{pmatrix} \tag{8.215}$$

and that this is of the form $H_C$ for

$$C = \pi^\dagger \left( \frac{\tanh\sqrt{\pi\pi^\dagger}}{\sqrt{\pi\pi^\dagger}} \right) \tag{8.216}$$

---

## 8.5 Torsors And Principal Bundles

**Definition** A *torsor* or *principal homogeneous space* for a group $G$ is a $G$-set $X$ on which the action is transitive and free.

Note that since the action is free $\mathrm{Stab}_G(x) = \{1_G\}$ for every element of $X$. Therefore, by the stabilizer-orbit theorem, given a choice of $x \in X$ we can set up an isomorphism of $G$-spaces:

$$X \cong G/\{1_G\} \cong G \tag{8.217}$$

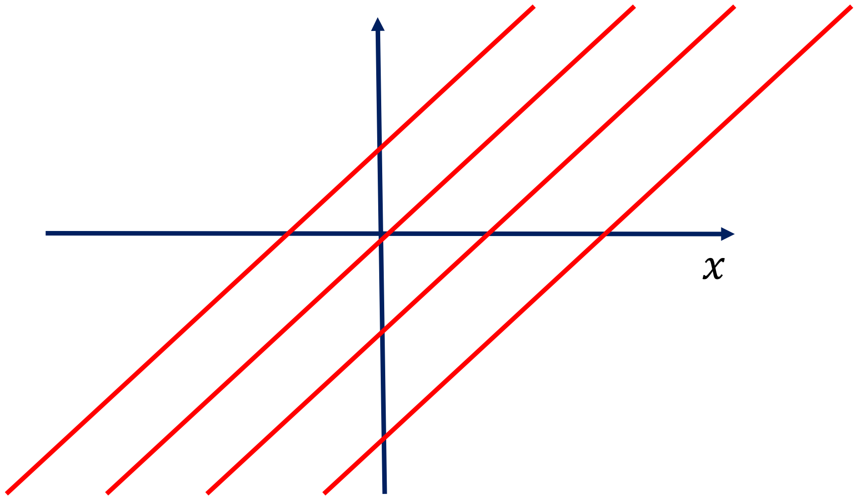



**Figure 12:** The torsors $\mathbb{Z} + x$, plotted in the $y$ direction are the points above a point $x$ on the horizontal axis. We construct the set by taking the line $y = x$ and then translating by integral shifts in the $y$-direction. The resulting set is also invariant under translation in the $x$-direction by integers. If we quotient by integer translation in the $x$-direction we can picture the quotient as an infinite spiral on a cylinder. The projection to the quotient of the $x$-axis becomes the projection $\mathbb{R} \to S^1$ given by the exponential map. As a generalization consider the family of lines $y = kx$ translated in the $x$ direction by integer shifts. This gives the principal $\mathbb{Z}$ bundles $P_k$ below for $k \neq 0$. If instead we consider the union of lines $y = n$ for $n \in \mathbb{Z}$ and quotient by integer translations in the $x$-direction we obtain the trivial $\mathbb{Z}$ bundle over $S^1$.

That is, we can set up a 1-1 correspondence between a torsor and elements of $G$, but in general there is no natural correspondence between $X$ and $G$ because to set up the 1-1 correspondence we needed to choose a point $x \in X$. A torsor has no distinguished element we can call the identity. Let us illustrate this idea with some examples:

1. Let $x \in \mathbb{R}$ be a real number. Consider the subset $X = \mathbb{Z} + x \subset \mathbb{R}$. This set is a torsor for $\mathbb{Z}$. But there is no natural zero in $X$. Indeed, let $x$ vary continuously, then any purported natural zero would vary continuously to any other number.

2. Imagine that the surface of the earth is flat and of infinite extent. Is this a copy of $\mathbb{R}^2$? Yes and no. We can identify it with $\mathbb{R}^2$, but not in any natural way: $\mathbb{R}^2$ is a vector space with a distinguished vector $\vec{0}$. Where should we put the origin? Rome? Beijing? Moscow? London? New York? Piscataway? Wuhan? Kiev? If the UN tried to assign an origin there would be endless disputes. However, there would never be any dispute about the vector in $\mathbb{R}^2$ needed to translate from New York to London. So, the difference of London minus New York is well-defined, but the sum is not. If London and New York represented vectors in a vector space then one could both add and subtract these vectors.

   The infinite flat earth is an example of two-dimensional affine Euclidean space $\mathbb{E}^2$. More formally: An *affine space* $\mathbb{E}^d$ *modeled on* $\mathbb{R}^d$ is a space of points with an action of $\mathbb{R}^d$ that translates the points so that nonzero vectors always move points and one can get from one point to any other by the action of a vector. But there is no natural choice of origin. In equations:

(a) If $v \in \mathbb{R}^d$ and $p \in \mathbb{E}^d$ then there is a point $p + v \in \mathbb{E}^d$ so that $(p + v) + v' = p + (v + v')$.

(b) If $p + v = p$ then $v = 0$.

(c) If $p, p' \in \mathbb{E}^d$ there is a (unique) vector $v \in \mathbb{R}^d$ so that $p' = p + v$. We can therefore say $p' - p = v$.

If we do choose an origin (this choice is arbitrary) then we can identify $\mathbb{E}^d \cong \mathbb{R}^d$. Indeed, it follows from the above statements that for any $p \in \mathbb{E}^d$, every $p' \in \mathbb{E}^d$ is of the form $p' = p + v$ for a unique $v \in \mathbb{R}^d$. So we map $\Psi_p : \mathbb{E}^d \to \mathbb{R}^d$ by taking $\Psi_p : p' \mapsto v$. In this language we can say that affine Euclidean space $\mathbb{E}^d$ is a principal homogeneous space for the Abelian group $\mathbb{R}^d$. [118]

3. Let $V$ be a finite-dimensional vector space over a field $\kappa$. The set $X = \mathcal{B}(V)$ of all ordered bases for $V$ is a $GL(n, \kappa)$ torsor: $g \cdot \{v_1, \dots, v_n\} := \{\tilde{v}_1, \dots, \tilde{v}_n\}$ where $\tilde{v}_i = g_{ji} v_j$. Any two such bases are related by some $g$, but, if we are just given an abstract vector space, there is no natural basis. In the case $\kappa = \mathbb{R}$ the torsor $\mathcal{B}(V)$ has two connected components. A choice of connected component is known as an *orientation of* $V$.

4. Of course, $G$ is has a transitive free left (or right) action by left- (or right-) multiplication on itself. So, a group $G$ is a $G$-torsor. When speaking of a torsor it is not necessarily true that there is no choice of origin or identity! Also note that $G$ does have a left (or right) action on itself by conjugation, but with this choice of group action $G$ is not a torsor.

5. For a nice post on torsors with other examples see https://math.ucr.edu/home/baez/torsors.html

---

[118] Affine Euclidean space still has a notion of distance: We do not need a choice of origin to speak of the distance between two points. If $p' - p = v$ then the distance is:

$$dist(p, p') := \parallel v \parallel \tag{8.218}$$

Now we can study the group of *isometries of* $\mathbb{E}^d$. This is the group of transformations $T : \mathbb{E}^d \to \mathbb{E}^d$ (not necessarily linear!) that preserves these distances:

$$dist(T(p), T(p')) = dist(p, p') \tag{8.219}$$

We denote it as $\mathrm{Euc}(d)$ and refer to it as the *Euclidean group.* We have defined $d$-dimensional Minkowski space $\mathbb{M}^{1,d-1}$ as a vector space, but we can equally well consider an affine Minkowskian space to be a torsor for $\mathbb{M}^{1,d-1}$ equipped with quadratic form on $p' - p = v$ given by

$$v \cdot v = v^\mu \eta_{\mu\nu} v^\nu \tag{8.220}$$

so the "distance squared" between two points is now $dist(p, p')^2 = v \cdot v$. Considered as an affine space we define the *Poincaré group* as the group of transformations $T : \mathbb{M}^{1,d-1} \to \mathbb{M}^{1,d-1}$ (not necessarily linear) preserving the quantity $v \cdot v$.

Now that we have some examples of torsors it is quite interesting to consider continuous families of torsors. We will illustrate this with several examples

1. Consider the exponential map
$$\pi : \mathbb{R} \to S^1 \tag{8.221}$$
$$\pi(x) := e^{2\pi \mathrm{i} x} \tag{8.222}$$
where we consider $S^1$ to be the set of complex numbers of modulus one. The *fiber over $z \in S^1$* is defined to be the pre-image:
$$\pi^{-1}(z) := \{x \in \mathbb{R} | e^{2\pi \mathrm{i} x} = z\} \tag{8.223}$$
For $x \in \pi^{-1}(z)$ we can write
$$x = \frac{1}{2\pi \mathrm{i}} \log z \tag{8.224}$$
but there are many branches of the log so this expression is ambiguous. A choice of a point in the fiber determines a choice of branch in the log. Suppose $z$ is given and let $x_0$ be some a solution to the equation
$$e^{2\pi \mathrm{i} x} = z \tag{8.225}$$
Then any other solution to the equation is of the form $x_0 + n$ for $n \in \mathbb{Z}$ and conversely every solution lies in the $\mathbb{Z}$-torsor $x_0 + \mathbb{Z}$. Now, in a way we will make precise below, the torsor varies continuously as $z$ varies continuously. Note that, locally, if we vary $z$ along a small continuous path in the circle then we can make a continuous path of real numbers $x_0$ in the fiber. So, locally, there is a neighborhood $\mathcal{U}$ of $x_0$ and in that neighborhood we can write a homeomorphism
$$\varphi : \pi^{-1}(\mathcal{U}) \to \mathcal{U} \times \mathbb{Z} \tag{8.226}$$
However, we cannot identify $\mathbb{R}$ globally with $S^1 \times \mathbb{Z}$. Note that $\mathbb{R}$ is connected while $S^1 \times \mathbb{Z}$ is not.

2. Here is another way of thinking of the previous example, together with a generalization. Let
$$\tilde{P} := \{(x, y) | y \in (\mathbb{Z} + x)\} \subset \mathbb{R}^2 \tag{8.227}$$
This subset of $\mathbb{R}^2$ is denoted in Figure 12. We have an obvious projection map $\tilde{\pi} : \tilde{P} \to \mathbb{R}$ with $\tilde{\pi}(x, y) = x$. The fiber of this map is the $\mathbb{Z}$-torsor $\mathbb{Z} + x$, by definition.

   Now note that $\tilde{P}$ is invariant under translation by integers in the $\mathbb{Z}$ direction
$$\phi(n, (x, y)) = (x + n, y) \tag{8.228}$$
If we quotient by this $\mathbb{Z}$-action we get a spiral in the cylinder with periodic coordinate $x$. We can identify the spiral with $\mathbb{R}$ (for example, by projection to the $y$-axis). We can identify $[x] \in \mathbb{R}/\mathbb{Z}$ with the circle. With these identifications $p(y) = [x]$ where $e^{2\pi \mathrm{i} y} = e^{2\pi \mathrm{i} x}$.

3. The previous picture can be generalized by considering the line $y = kx$ for $k \neq 0$ and all its translates under by integer shifts of $x$. The resulting space $\tilde{P}_k$ can be quotiented by the same action as (8.228) to give again a spiral $P_k$ in a cylinder with slope $k$. The fiber of $\pi : P_k \to S^1$ is the $\mathbb{Z}$-torsor $x + k\mathbb{Z}$ where the $\mathbb{Z}$ action is by translation by integer multiplets of $k$. Again the spiral can be identified with $\mathbb{R}$ by projection to $y$. Again the quotient of the $x$-axis by integer translation is $\mathbb{R}/\mathbb{Z} \cong S^1$. Now $p(y) = [x]$ with $e^{2\pi i y} = e^{2\pi i k x}$. For $k = 0$ the analog of $\tilde{P}_k$ is the union of translates in the $y$-direction by integer shifts of the $x$-axis. The quotient by integer shifts in the $x$-direction is the product $S^1 \times \mathbb{Z}$.

4. Let us return to our example $\pi_{\hat{n}} : SO(3) \to S^2$. Referring to our discussion above we saw that for each $\hat{k} \in S^2$ the fiber
$$\pi_{\hat{n}}^{-1}(\hat{k}) \tag{8.229}$$
is a principal homogeneous space for $SO(2)$: It is the set of rotations that take $\hat{n}$ to $\hat{k}$. But there is no canonical identification of this torsor with the group elements $SO(2)$ that varies continuously with $\hat{k}$ and includes $\hat{k} = \pm\hat{n}$.

5. Generalizing the previous example, let $H$ be a subgroup of any group $G$. Consider the projection from $G$ to the set of left $H$-cosets:
$$\pi : G \to G/H \tag{8.230}$$
defined by $\pi(g) := gH$. Note that the fibers of a coset $gH$ are:
$$\pi^{-1}(gH) = gH \tag{8.231}$$
Note: On the LHS of (8.231) $gH$ is best thought of as a <u>point</u> in the homogeneous space $G/H$. On the other hand, on the RHS, $gH$ is the fiber of the map $\pi$, more naturally thought of as a <u>subset</u> of $G$. The subset $gH \subset G$ can be put into 1-1 correspondence with $H$, but not in any natural way. While it is identified with $H$ as a set, it is not identified as a group because there is no natural element in $gH$ to identify with the unit in $H$.

The above examples are special cases of an extremely important idea in mathematics - that of a *principal fiber bundle*. In the language of this section, a principal $G$-bundle is a continuous family of $G$-torsors. Before giving the formal definition we look at one important special case: Let $X$ be any topological space. Consider the product space $P = X \times G$. This has two properties which will be key to the generalization:

1. There is a free right $G$-action: $(x, g) \cdot g_0 := (x, gg_0)$.

2. There is a surjective continuous map $\pi : P \to X$ given by projection on the first factor such that
$$\pi(p \cdot g_0) = \pi(p) \tag{8.232}$$
so that the $G$-action is transitive and free on the fibers.

In a principal $G$ bundle we generalize so that the above is the correct model locally but globally the situation can be twisted.

Here is the formal definition:

**Definition**: Let $G$ be a topological group. A *principal $G$-bundle* is a continuous surjective map of topological spaces $\pi : P \to X$, where $P$ is called the *total space* and $X$ is called the *base space* such that

1. There is a continuous and free right $G$ action on $P$ with
$$\pi(p \cdot g) = \pi(p) \tag{8.233}$$
so that the $G$ action is transitive on the fibers of $\pi$, namely, on the sets $\pi^{-1}(x)$ for $x \in X$. In other words, the fibers $\pi^{-1}(x)$ are $G$-torsors. [119]

2. $\pi : P \to X$ satisfies "local triviality"

The "local triviality" condition means, intuitively, that for any point in the base $x \in X$ there is a neighborhood that "looks like" the direct product example discussed above. This is the technical way of implementing the idea that the torsors "vary continuously."

Technically, we require that for all $x \in X$ there is a neighborhood $\mathcal{U}_x \subset X$ with a homeomorphism $\phi_{\mathcal{U}_x} : \pi^{-1}(\mathcal{U}_x) \to \mathcal{U}_x \times G$ such that

$$\begin{array}{ccc} \pi^{-1}(\mathcal{U}_x) & \xrightarrow{\phi_{\mathcal{U}_x}} & \mathcal{U}_x \times G \\ & \searrow^{\pi} & \downarrow{\scriptstyle \pi'} \\ & & \mathcal{U}_x \end{array} \tag{8.234}$$

commutes, where $\pi'$ is the canonical projection map onto the first factor. Moreover, we require that $\phi_{\mathcal{U}_x}$ is a $G$-equivariant map where we use the natural right $G$-action on the Cartesion product $\mathcal{U}_x \times G$: Namely, $g_0$ acts by $(y, g) \mapsto (y, gg_0)$ for $y \in \mathcal{U}_x$. Thus, the $G$-equivariance requirement states that if $p \in \pi^{-1}(\mathcal{U}_x)$ and $\phi_{\mathcal{U}_x}(p) = (y, g)$ then

$$\phi_{\mathcal{U}_x}(p \cdot g_0) = (y, g \cdot g_0) \ . \tag{8.235}$$

If we impose some mild topological conditions on $X$ then we can replace the set of "local trivializations" $\{(\mathcal{U}_x, \phi_x)\}_{x\in X}$ by an more manageable set given by an atlas $\{\mathcal{U}_\alpha\}$ of open covers of $X$. That is we have a collection $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$ where $\mathcal{U}_\alpha$ are open sets covering $X$ and
$$\phi_\alpha : \pi^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times G \tag{8.236}$$
are $G$-equivariant homeomorphisms and $\alpha$ ranges over some indexing set. If $X$ is compact it can be taken to be a finite indexing set.

[119]Therefore, if $X$ has more than one point the $G$-action is not transitive on all of $P$.

Given a collection of local trivializations $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$ it is interesting to consider how they compare on patch overlaps $\mathcal{U}_{\alpha\beta} := \mathcal{U}_\alpha \cap \mathcal{U}_\beta$. We have the diagram

$$\begin{array}{ccccc} \mathcal{U}_{\alpha\beta} \times G & \xrightarrow{\phi_\beta^{-1}} & \pi^{-1}(\mathcal{U}_{\alpha\beta}) & \xrightarrow{\phi_\alpha} & \mathcal{U}_{\alpha\beta} \times G \\ & {\scriptstyle\pi}\searrow & \downarrow{\scriptstyle\pi} & \swarrow{\scriptstyle\pi} & \\ & & \mathcal{U}_{\alpha\beta} & & \end{array} \tag{8.237}$$

where here $\phi_\beta^{-1}$ should, strictly speaking, be written as $(\phi_\beta|_{\pi^{-1}(\mathcal{U}_{\alpha\beta})})^{-1}$ and $\phi_\alpha$ should, strictly speaking, be written as $\phi_\alpha|_{\pi^{-1}(\mathcal{U}_{\alpha\beta})}$. We define the *transition function*

$$\phi_{\alpha\beta} : \mathcal{U}_{\alpha\beta} \times G \to \mathcal{U}_{\alpha\beta} \times G \tag{8.238}$$

to be the map

$$\phi_{\alpha\beta} := (\phi_\alpha|_{\pi^{-1}(\mathcal{U}_{\alpha\beta})}) \circ (\phi_\beta|_{\pi^{-1}(\mathcal{U}_{\alpha\beta})})^{-1} \tag{8.239}$$

Note that $\phi_{\alpha\beta}$ is a $G$-equivariant map. We will now use $G$-equivariance to cast $\phi_{\alpha\beta}$ in a useful form.

For any set $X$ consider a $G$-equivariant map

$$F : X \times G \to X \times G \tag{8.240}$$

such that

$$\begin{array}{ccc} X \times G & \xrightarrow{F} & X \times G \\ & {\scriptstyle\pi}\searrow & \downarrow{\scriptstyle\pi} \\ & & X \end{array} \tag{8.241}$$

that is, such that $\pi(F(x,g)) = \pi(x,g) := x$. Therefore $F(x,g) = (x, f(x,g))$ where $f : X \times G \to G$ is a map and the $G$-equivariance of $F$ means that $f(x, gg_0) = f(x,g)g_0$. But this implies that $f(x,g) = h(x)g$ where $h : X \to G$ is just some map. (Set $h(x) = f(x,1)$. In other words $F$ must be of the form $F(x,g) = (x, h(x)g)$. If $F$ is continuous, smooth, ... then $h$ will be continuous, smooth, ...

Now, applying the previous paragraph we see that $\phi_{\alpha\beta}$ must have the form:

$$\phi_{\alpha\beta} : (x,g) \mapsto (x, g_{\alpha\beta}(x)g) \qquad\qquad x \in \mathcal{U}_{\alpha\beta} \tag{8.242}$$

for some continuous function

$$g_{\alpha\beta} : \mathcal{U}_{\alpha\beta} \to G \tag{8.243}$$

The $G$-valued functions $g_{\alpha\beta}$ are called *clutching functions.*

It follows from the definition (8.239) of the transition functions that $g_{\alpha\beta}(x) = g_{\beta\alpha}(x)^{-1}$ and that on triple overlaps $\mathcal{U}_{\alpha\beta\gamma} := \mathcal{U}_\alpha \cap \mathcal{U}_\beta \cap \mathcal{U}_\gamma$

$$g_{\alpha\beta}(x) g_{\beta\gamma}(x) g_{\gamma\alpha}(x) = 1 \qquad\qquad x \in \mathcal{U}_{\alpha\beta\gamma} \tag{8.244}$$

a condition called the *cocycle condition.*

One virtue of using the data of an atlas $\{\mathcal{U}_\alpha\}$ and a set of continuous clutching maps $g_{\alpha\beta} : \mathcal{U}_{\alpha\beta} \to G$ satisfying $g_{\alpha\beta} = g_{\beta\alpha}^{-1}$ and the cocycle condition on triple overlaps is that one can construct a principal $G$ bundle by a gluing construction. One begins with the disjoint union $\amalg_\alpha(\mathcal{U}_\alpha \times G)$ and then takes $P = \amalg_\alpha(\mathcal{U}_\alpha \times G)/\sim$ where the equivalence relation identifies, for each $x \in \mathcal{U}_{\alpha\beta}$ a point $(x, g) \in \mathcal{U}_\alpha \times G$ with $(x, g_{\alpha\beta}(x)g) \in \mathcal{U}_\beta \times G$.

We now give a number of examples of principal bundles

**Examples**:

1. Let $X$ be any topological space and $G$ any topological group. Then $P = X \times G$ with $\pi : (x, g) \mapsto x$ is a principal bundle. A bundle of this form is known as the trivial bundle.

2. Let $H \subset G$ be a Lie subgroup of a Lie group $G$. Then $\pi : G \to G/H$ is a principal $H$ bundle over the homogeneous space $X = G/H$. (Note the free right $H$-action that commutes with $\pi$ is $g \mapsto g \cdot h$ so that $\pi(g) = gH = \pi(gh)$.

3. Returning to our discussion of the $SO(3)$ action on $S^2$, $\pi_{\hat{n}} : SO(3) \to S^2$ defines a principal $SO(2)$ bundle over $S^2$

4. Similarly, recalling the action of $SU(2)$ on $\mathbb{CP}^1$ and choosing some base point on $\mathbb{CP}^1$ we get a map $\pi : SU(2) \to \mathbb{CP}^1$ defining a principal $U(1)$ bundle over $\mathbb{CP}^1 \cong S^2$.

5. The previous example has a nice generalization. Consider the finite-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^{n+1}$ and consider the unit sphere in $\mathcal{H}$,

$$S(\mathcal{H}) = \{\psi \in \mathcal{H} | \langle \psi, \psi \rangle = 1\} \tag{8.245}$$

   This is the set of normalized vectors. In quantum mechanics with a finite dimensional Hilbert space one might try to represent physical states by such a unit vector $\psi$. Writing out the components of $\psi$ in real and imaginary parts shows that $S(\mathcal{H}) \cong S^{2n+1} \subset \mathbb{R}^{2n+2}$. Now, in quantum mechanics we are instructed to identify statevectors that differ by a phase $\psi \sim z\psi$ for $|z| = 1$. The quotient of $S(\mathcal{H})$ by this $U(1)$ action is just $\mathbb{CP}^n$. One can check the local triviality so that

$$\pi : S(\mathcal{H}) \to \mathbb{CP}^n \tag{8.246}$$

   is a principal $U(1)$ bundle.

6. *Principal G-bundles Over The Circle.* Let $G$ be a discrete group. Then $\mathbb{R} \times G$ is the trivial principal $G$- bundle over $\mathbb{R}$. We can make a more interesting bundle by considering the left $\mathbb{Z}$-action defined by choosing an element $g_0 \in G$ and defining for $n \in \mathbb{Z}$:

$$\phi^{(g_0)}(n, (x, g)) = (x + n, g_0^n g) \tag{8.247}$$

   so equivalence classes satisfy: $[(x, g)] = [(x + n, g_0^n g)]$. The quotient $P = (\mathbb{R} \times G)/\mathbb{Z}$ by this action is the total space of a principal $G$-bundle over $S^1$. The projection map is

$$\pi([(x, g)]) = e^{2\pi i x} \in S^1 \tag{8.248}$$

and the right $G$-action is

$$[(x,g)]\cdot g' = [(x,gg')] \tag{8.249}$$

Note well that even if $G$ is nonabelian this is well-defined because the equivalence relation defining $[(x,g)]$ is defined by a commuting left $G$-action.

Another way to describe $P_{g_0}$ is to consider the trivial bundle $[0,1]\times G$ and divide by an equivalence relation where $(0,g)$ is identified with $(1,g_0 g)$. We will say more about these bundles below.

♣Elaborate more on this. A figure would help. ♣

7. Our examples above of spirals in cylinders projecting to the circle are principal $G=\mathbb{Z}$-bundles of the form $P_{g_0}$ where $g_0 = k \in \mathbb{Z}$.

8. Another important special case of the general construction of the principal $G$ bundles $P_{g_0}$ over the circle concerns the case where $G=\mu_N$. Again, we view $S^1$ as the set of complex numbers of unit norm and let $g_0 = \exp[-2\pi \mathrm{i}/N]$. We can then define a one-one map:

$$\varphi : P_{g_0} \to S^1 \tag{8.250}$$

via

$$\varphi([(x,\zeta)]) = \zeta e^{2\pi \mathrm{i} x/N} \tag{8.251}$$

As $x$ ranges from 0 to 1 the complex number $e^{2\pi \mathrm{i}x/N}$ covers only an arc of angle $2\pi/N$ in the circle. But then the $N^{th}$ roots of unity $\zeta\in\mu_N$ fill out the rest of the circle, so the map $\varphi$ is in fact both surjective and injective. It can also be shown to be continuous. We now have a commutative diagram:

$$\begin{array}{ccc} P_{g_0} & \xrightarrow{\varphi} & S^1 \\ & {\scriptstyle \pi_1}\searrow \quad \swarrow{\scriptstyle \pi_2} & \\ & S^1 & \end{array} \tag{8.252}$$

where $\pi_2(w)=w^N$. So the fiber $\pi_2^{-1}(z)$ is the set of $N^{th}$ roots of $z$. Similarly, our example with $G=\mathbb{Z}$ and $k=1$ is the geometry behind choosing a branch of the logarithm.

A *map of $G$-torsors*, or, more properly, a *morphism of $G$-torsors* is a map that preserves the mathematical structure of being a $G$-torsor. So, if $\mathcal{X}_1$ and $\mathcal{X}_2$ are two $G$-torsors then a morphism of $G$-torsors is a map

$$\psi : \mathcal{X}_1 \to \mathcal{X}_2 \tag{8.253}$$

such that

$$\psi(y\cdot g) = \psi(y)\cdot g \tag{8.254}$$

for all $y\in\mathcal{X}_1$. In other words, $\psi$ is *equivariant.*

A *bundle map* between two principal $G$-bundles, or, more properly a *morphism of $G$-bundles* is a continuous map that preserves fibers and restricts on fibers to a be a morphism of $G$-torsors. In diagrams, if $\psi$ is a bundle map between two principal $G$-bundles $P_1$ and $P_2$ then $\psi : P_1 \to P_2$ is a map so that it preserves fibers, meaning that:

$$\begin{array}{ccc} P_1 & \xrightarrow{\psi} & P_2 \\ & {\scriptstyle \pi_1} \searrow \quad \swarrow {\scriptstyle \pi_2} & \\ & X & \end{array} \tag{8.255}$$

commutes. That is $\pi_2(\psi(p_1)) = \pi_1(p_1)$ for all $p_1 \in P$. Moreover, the map must be $G$ equivariant, or a morphism of torsors on the fibers, so that [120]

$$\psi(p_1 \cdot g) = \psi(p_1) \cdot g \tag{8.256}$$

If there are bundle maps $\psi_1 : P_1 \to P_2$ and $\psi_2 : P_2 \to P_1$ whose composition is the identity then the bundles are said to be *isomorphic bundles*. As an exercise below you prove that morphisms of principal bundles are always isomorphisms.

**Remarks**:

1. If a principal $G$-bundle $\pi : P \to X$ is isomorphic to the trivial bundle then it is said to be *trivializable*. Note that two different trivializations $\psi_1$ and $\psi_2$ of a trivializable bundle will differ by a continuous $G$-equivariant map $F : X \times G \to X \times G$. We have a diagram very similar to (8.237):

$$\begin{array}{ccccc} X \times G & \xrightarrow{\psi_1^{-1}} & P & \xrightarrow{\psi_2} & X \times G \\ & {\scriptstyle \pi} \searrow & \downarrow {\scriptstyle \pi} & \swarrow {\scriptstyle \pi} & \\ & & X & & \end{array} \tag{8.257}$$

   so that $\psi_1 = \psi_2 \circ F$. As we have seen above $F(x, g) = (x, h(x)g)$ is governed by a continuous map $h : X \to G$. It can very well happen that the continuous maps $Map(X, G)$ has nontrivial topology, so that there can be topologically inequivalent trivializations of a trivializable bundle. For example, we can consider a $U(1)$ bundle over the cylinder and if a "vortex" is located in the cylinder the trivialization of the principal $U(1)$ bundle over the two ends of the cylinder will differ by a trivialization differing by some nontrivial element of $\pi_1(U(1)) \cong \mathbb{Z}$. Similarly, considering $SU(2)$ bundles over $S^3 \times [0, 1]$ if an "instanton" is located in interior then the $SU(2)$ bundles on the boundary copies of $S^3$ will be trivializable but the trivializations will differ by a nontrivial element of $\pi_3(SU(2)) \cong \mathbb{Z}$. This is one reason why it is actually important to distinguish a trivial bundle from a trivializable bundle.

2. For other examples of how the distinction between trivializable and trivial can have physical implications in condensed matter physics see [121] .

[120] The reader familiar with the general theory of fiber bundles will note that (8.255) alone serves as the definition of a bundle map for a fiber bundle. But a principal bundle has more structure, and for a morphism of principal bundles we require the additional condition (8.256).

[121] See G. Moore, "A Comment On Berry Connections," e-Print arXiv:1706.01149 for some examples.

3. One can show that the principal $U(1)$ (8.246) is a nontrivial bundle. A consequence (from a theorem we prove below) is that one cannot continuously choose a phase of a statevector for the entire family of pure states on a Hilbert space.

Let us return to our example of $\pi : P_{g_0} \to S^1$. Denote the equivalence classes $[(x,g)]_{g_0} = [(x+n, g_0^n g)]_{g_0}$ where we added a subscript $g_0$ to emphasize the dependence on $g_0$. Note that the following diagram commutes for all $n \in \mathbb{Z}$: commutes:

$$\begin{array}{ccc} \mathbb{R}\times G & \xrightarrow{\psi_h} & \mathbb{R}\times G \\ \downarrow{\scriptstyle \phi_n^{(g_0)}} & & \downarrow{\scriptstyle \phi_n^{(hg_0h^{-1})}} \\ \mathbb{R}\times G & \xrightarrow{\psi_h} & \mathbb{R}\times G \end{array} \tag{8.258}$$

where $\psi_h : (x,g) \mapsto (x, hg)$ and $\phi^{(g_0)}$ is the $\mathbb{Z}$-action defined in equation (8.247) above. This implies that $\psi_h$ descends to a well-defined bundle map $P_{g_0} \to P_{hg_0h^{-1}}$:

$$\psi_h : [(x,g)]_{g_0} \mapsto [(x,hg)]_{hg_0h^{-1}} \tag{8.259}$$

Clearly $\psi_{h^{-1}}$ defines the inverse bundle map.

*Therefore, the isomorphism classes of principal $G$ bundles over the circle are labeled by conjugacy classes of elements of $G$. That is, $P_{g_0}$ and $P_{hg_0h^{-1}}$ are isomorphic principal $G$-bundles over $S^1$.*

Associated Bundles

Finally, let $\pi : P \to X$ be a principal $G$ bundle and let $Y$ be any $G$-space with a left $G$-action. Then we can define a $G$ action on $P \times Y$:

$$\phi_g(p,y) \mapsto (p\cdot g^{-1}, g\cdot y) \tag{8.260}$$

Notice this is a left $G$-action. The quotient space, usually denoted $P \times_G Y$ has a well-defined continuous map

$$\tilde{\pi} : P \times_G Y \to X \tag{8.261}$$

defined by $\tilde{\pi}([p,y]) = \pi(p)$. The fibers of $\tilde{\pi}$ can be identified with the space $Y$. One way to understand this is to use the local trivialization on the principal bundle to reduce to the case where $P = \mathcal{U} \times G$ is a trivial bundle. Then the equivalence class of $((x,g),y)$ always has a unique representative of the form $((x,1),y')$. Therefore we can identify

$$(\mathcal{U} \times G) \times_G Y \tag{8.262}$$

with $\mathcal{U} \times Y$. An answer to an exercise below has another viewpoint.

Note that the fibers of $\tilde{\pi}$ are not $G$-torsors but more general $G$-spaces. The map

$$\tilde{\pi} : P \times_G Y \to X \tag{8.263}$$

defines what is called *an associated bundle to the principal bundle $\pi : P \to X$ by the $G$-set $Y$*. Note that $P \times_G Y$ in general does not have a right action by $G$.

We can also describe the associated bundle in terms of gluing if we are given an atlas $\{\mathcal{U}_\alpha\}$ and a set of clutching functions $g_{\alpha\beta} : \mathcal{U}_{\alpha\beta} \to G$ for $P$. Then the associated bundle is obtained by gluing together sets of the form $\mathcal{U}_\alpha \times Y$ on overlaps $\mathcal{U}_{\alpha\beta}$ via:

$$\tilde{\phi}_{\alpha\beta} : (x, y) \to (x, g_{\alpha\beta}(x) \cdot y) \tag{8.264}$$

where $x \in \mathcal{U}_{\alpha\beta}$.

**Examples**

1. If the $G$-action on $Y$ is trivial then $P \times_G Y \cong X \times Y$ in a natural way.

2. Let $G = \mu_2$, and $\pi : P_{-1} \to S^1$ be the nontrivial principal $G$-bundle associated with taking the square root. Let $Y = [-1, 1]$ be the $G$-space where the nontrivial element of $\mu_2$ takes $y \to -y$. Then the associated bundle

$$(P_{-1} \times Y)/\mu_2 \tag{8.265}$$

is easily seen to be the Mobius band.

3. *Associated Vector Bundles* An important special case of the associated bundle construction is the case where $Y = V$ is a vector space which is the carrier space of a linear representation of $G$. Then we speak of "the vector bundle associated to $P$ via the representation $(T, V)$." If $T : G \to GL(V)$ is the homomorphism of the representation then the gluing maps for gluing $\mathcal{U}_\alpha \times V$ to $\mathcal{U}_\beta \times V$ over the intersection $\mathcal{U}_{\alpha\beta}$ are:

$$(x, v) \sim (x, T(g_{\alpha\beta}(x))(v)) \tag{8.266}$$

for $x \in \mathcal{U}_{\alpha\beta}$ and $v \in V$.

4. *The Magnetic Monopole Line Bundles.* Consider the principal $U(1)$ bundle $\pi : SU(2) \to SU(2)/U(1) \cong \mathbb{CP}^1 \cong S^2$, viewed as a principal $U(1)$ bundle over $S^2$. Consider $V_\pm$ the one-dimensional representations of $U(1)$ associated to the fundamental and its complex conjugate. The associated bundles

$$\mathcal{L}_\pm = P \times_{U(1)} V_\pm \tag{8.267}$$

are complex line bundles. For a very nice alternative description of these line bundles in terms of the eigenlines of the family of Hermitian operators $\vec{s} \cdot \vec{\sigma}$ with $\vec{s} \in \mathbb{R}^3 - \{0\}$ see the exercise below. These bundles play a huge role in physics. One frequently occurring instance stems from the following fact. Consider a continous family of Hermitian Hamiltonians $H_s$ parametrized by some manifold of control parameters $s \in \mathcal{S}$. [122] Suppose there is a level crossing at some point $s_*$. Generically the level crossing will involve two simple eigenlines. Then one can study the crossing behavior locally near $s_*$ focussing on the two-dimensional space spanned by these eigenlines.

[122] For more on this see LAUM.

One can show that the family of Hamiltonians on the two-level system is of the form $\vec{s}\cdot\vec{\sigma}$ with $\vec{s}\in\mathbb{R}^3$ with the level crossing happening at $\vec{s}=0$. Restricting to a sphere surrounding $\vec{s}=0$ the bundle of $\pm1$ eigenlines are isomorphic to the line bundles $\mathcal{L}_\pm$. See the exercise below.

Sections Of Bundles

If $\pi : E \to X$ is a principal bundle, or an associated bundle to a principal bundle (or any fiber bundle, if you know what that means) then we define a *section of* $\pi$ to be a map

$$s : X \to E \tag{8.268}$$

which is a right-inverse to $\pi$. That is

$$\pi \circ s(x) = x \qquad\qquad \forall x \in X \tag{8.269}$$

Note that $s(x)$ is always an element of $E$ which is in the fiber of $\pi : E \to X$ above $x$. One must be careful with this terminology because often in bundle theory when one speaks of a section it is implicitly assumed that one is discussing a continuous section. In this case $s : X \to E$ must be a continuous map.

Because of local triviality, continuous sections always exist locally. That is, near any $x \in X$ there will be continuous sections of the bundle $\pi^{-1}(\mathcal{U}) \to \mathcal{U}$ for some neighborhood $\mathcal{U}$ of $x$. It is less obvious what happens globally. In fact, we have

**Theorem** A principal $G$-bundle $\pi : P \to X$ is isomorphic to the trivial bundle $\pi : X\times G \to X$ iff there is a globally defined continuous section.

*Proof*: Suppose there is a globally defined section $s : X \to P$. Then let

$$\psi : X \times G \to P \tag{8.270}$$

be defined by $\psi(x, g) := s(x)g$. One checks this is a bundle morphism. Conversely, suppose there is a bundle morphism $\psi : P \to G\times X$. Then for each $x \in X$ there is a unique $s(x) \in P$ so that $\psi(s(x)) = (x, 1_G)$. So: A continuous section can be viewed as a continuous choice of basepoint to identify the $G$-torsor with the group $G$. ♠

If we apply the above theorem to the bundles $\pi : P_{g_0} \to S^1$ for a discrete group we find that they are trivializable only when $g_0 = 1_G$.

The situation is rather different for associated bundles. For example if $Y = V$ is a linear space then the space of sections is always nonempty and in fact is an infinite dimensional vector space. One way to see that it is nonempty is simply to take the $s(x) = [(p, 0)]$ where $p$ is any point in $\pi^{-1}(x)$. In general, if $g_{\alpha\beta} : \mathcal{U}_{\alpha\beta} \to G$ are the transition functions for $P$ and $Y = (T, V)$ is the representation space of $G$ with $T : G \to GL(V)$ then

$$(x, v_\alpha) \sim (x, T(g_{\alpha\beta}(x))v_\beta) \qquad\qquad x \in \mathcal{U}_{\alpha\beta} \tag{8.271}$$

defines the gluing rules for gluing $\mathcal{U}_\alpha \times V$ to $\mathcal{U}_\beta \times V$. (See the exercise below.) A section of a vector bundle can thus be identified with a collection of continuous maps $s_\alpha : \mathcal{U}_\alpha \to V$ such that for $x \in \mathcal{U}_{\alpha\beta}$ we have

$$s_\alpha(x) = T(g_{\alpha\beta}(x))s_\beta(x) \tag{8.272}$$

Using partitions of unity it is easy to construct infinitely many sections.
**Remarks**:

1. Going back to the line bundles $\mathcal{L}_\pm \to S^2$ discussed above, these bundles first entered physics in an important way in P.A.M. Dirac's magnificent 1931 paper that observed that the wave function of an electron in the presence of a magnetic monopole of magnetic charge $\pm 1$ must be a section of $\mathcal{L}_\pm$. They have played an important role ever since.

2. Going back to the bundles $S(\mathcal{H}) \to \mathbb{CP}^n$. Since the case $n = 1$ is nontrivial, all these bundles must be nontrivial. Now, a section of such a principal $U(1)$ bundle would be, physically, an identification of a normalized wavevector $\psi$ for each pure state $\rho$. The nontriviality of the principal $U(1)$ bundle means there is no continuous way to do this for all pure states.

Finally, we state a very nontrivial fact about how principal bundles can be classified. It uses two preliminary remarks.

The first remark is that if $f : X \to X'$ is a continuous map and $\pi' : P' \to X'$ is a principal $G$ bundle then we can form a principal $G$-bundle $\pi : P \to X$ by declaring $P$ to be the set of pairs $(p', x) \in P' \times X$ such that $f(x) = \pi'(p')$. For such a pair we define $\pi(p', x) := x$. It is not hard to see that the fiber of $\pi$ is naturally identified with the $G$-torsor $(\pi')^{-1}(f(x))$ and that the local triviality condition holds. The bundle $\pi : P \to X$ is called the *pullback of* $\pi' : P' \to X'$ *via* $f$ and is sometimes denoted $f^*(P')$.

As another preliminary remark, let $H_1, H_2$ be two commuting subgroups of $G$. Then there is a right $H_2$ action on $G/H_1$ and we have an $H_2$-bundle

$$\pi : G/H_1 \to G/(H_1 \times H_2) \tag{8.273}$$

Classification Of Isomorphism Classes Of Bundles. We commented above that the projections to homogeneous spaces $\pi : G \to G/H$ give examples of principal $H$-bundles. and if $H_1, H_2$ commute then we can generalize to give principal $H_2$ bundles $\pi : G/H_1 \to G/(H_1 \times H_2)$. There is a sense in which all principal $G$ bundles are related to such examples. We consider for simplicity the principal $U(k)$ bundles. One can generalize the Grassmannian of $k$-planes $\mathrm{Gr}_k(\mathbb{C}^N)$ to the Grassmannian of $k$-planes in a separable Hilbert space $\mathrm{Gr}_k(\mathcal{H})$. There is, in fact, a sense in which this is a limit of $\mathrm{Gr}_k(\mathbb{C}^N)$ for $N \to \infty$. Recall that we can model $\mathrm{Gr}_k(\mathbb{C}^N)$ as the homogeneous space $U(N)/(U(k) \times U(N-k)$ so we can talk about the principal $U(k)$ bundle $U(N)/U(N-k) \to \mathrm{Gr}_k(\mathbb{C}^N)$. The quotient $U(N)/U(N-k)$ can be interpreted as the space of $k$-dimensional subspaces of $\mathbb{C}^N$

equipped with an ordered ON basis. Therefore, there is an analogous space $E_k(\mathcal{H})$ of $k$-dimensional subspaces of $\mathcal{H}$ with an ordered ON basis, and a principal $U(k)$ bundle $\pi : E_k(\mathcal{H}) \to \mathrm{Gr}_k(\mathcal{H})$. One can show that isomorphism classes of $U(k)$ bundles over $X$ are in 1-1 correspondence with homotopy classes of maps $X \to \mathrm{Gr}_k(\mathcal{H})$.

---

**Exercise** *Bundle Maps For Trivial Bundles*

Show that the most general bundle map from the trivial bundle $\pi : X \times G \to X$ to itself is of the form:

$$\psi : (x, g) \mapsto (x, h(x)g) \tag{8.274}$$

for some continuous map $h : X \to G$.

---

---

**Exercise** *Morphisms Of Principal G-Bundles Are Isomorphisms*

a.) Show that any morphism of $G$-torsors is an isomorphism of $G$-torsors.

b.) Extend this to show that any morphism of principal $G$-bundles is an isomorphism of principal $G$-bundles.

---

---

**Exercise** *Fibers Of An Associated Bundle*

Show that the fibers of the associated bundle $\tilde{\pi} : P \times_G Y \to X$ described above are in 1-1 correspondence with the space $Y$. [123]

---

---

**Exercise** *The Monopole And Anti-Monopole Line Bundles*

a.) Show that if we consider the associated line bundle to the $U(1)$ principal bundle $\pi : SU(2) \to SU(2)/U(1) \cong S^2$ with equivalence relation

$$[(u, z)]_+ = [(u \begin{pmatrix} \zeta^{-1} & 0 \\ 0 & \zeta \end{pmatrix}, \zeta z)]_+ \tag{8.275}$$

[123] *Answer*: Let us define a map $f : \tilde{\pi}^{-1}(x) \to Y$. To do this we must <u>choose</u> an element $p_0 \in \pi^{-1}(x)$. Then if $[p, y] \in \tilde{\pi}^{-1}(x)$ it follows that $\pi(p) = x$ and hence $p = p_0 \cdot g_0$ for some $g_0$. Then we define $f[p, y] = g_0^{-1} \cdot y$. The reader needs to check that this map is well-defined. Since we can choose any $y \in Y$ the map is clearly surjective. Finally note that $[p_1, y_1] = [p_1, y_2]$ implies that $y_1 = y_2$ since the $G$-action on $P$ is free. Therefore the map is also injective. Note that the map does depend on a choice of $p_0$ for each $x$, so there is no canonical identification of the fibers with $Y$.

with $u \in SU(2)$ and $z \in \mathbb{C}$, and the $U(1)$ action is shown for $\zeta \in U(1)$. Parametrizing, as usual, $SU(2)$ elements as

$$u = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{8.276}$$

we have a well defined map from $SU(2) \times_{U(1)} \mathbb{C}$ to $\mathbb{C}^2$ given by

$$[(u, z)]_+ \mapsto \begin{pmatrix} \alpha z \\ -\bar{\beta} z \end{pmatrix} \tag{8.277}$$

Therefore, there is a well-defined map

$$\psi : SU(2) \times_{U(1)} \mathbb{C} \to S^2 \times \mathbb{C}^2 \tag{8.278}$$

given by

$$\psi^+([u, z]_+) := (\pi(u), \begin{pmatrix} \alpha z \\ -\bar{\beta} z \end{pmatrix}) \tag{8.279}$$

For a fixed value of $\pi(u)$ the image will be a <u>complex line</u> in $\mathbb{C}^2$. It is the complex line through

$$\begin{pmatrix} \alpha z \\ -\bar{\beta} z \end{pmatrix} \tag{8.280}$$

for any choice of $z \in \mathbb{C}^*$.

b.) Similarly, show that if we consider the associated line bundle with equivalence relation

$$[(u, z)]_- = [(u \begin{pmatrix} \zeta^{-1} & 0 \\ 0 & \zeta \end{pmatrix}, \zeta^{-1} z)]_- \tag{8.281}$$

we also get a map

$$\psi^-([u, z]_+) := (\pi(u), \begin{pmatrix} \beta z \\ \bar{\alpha} z \end{pmatrix}) \tag{8.282}$$

Again, the image of this map, for fixed $\pi(u)$ will be a complex line in $\mathbb{C}^2$.

c.) Show that if $\hat{x} \in S^2$ then $P(\hat{x}) := \hat{x} \cdot \vec{\sigma}$ has eigenvalues $\pm 1$.

d.)Now, using the usual spherical coordinates in $\mathbb{R}^3$:

$$x^1 = \sin\theta \cos\phi \qquad x^2 = \sin\theta \sin\phi \qquad x^3 = \cos\theta \tag{8.283}$$

show that

$$\hat{x} \cdot \vec{\sigma} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix} \tag{8.284}$$

and that the eigenline defined by

$$P(\hat{x})\psi = \psi \tag{8.285}$$

is the line in $\mathbb{C}^2$ proportional to

$$\begin{pmatrix} \cos\theta/2 \\ e^{i\phi} \sin\theta/2 \end{pmatrix} \tag{8.286}$$

e.) Similarly, show that the eigenline defined by

$$P(\hat{x})\psi = -\psi \tag{8.287}$$

is the line in $\mathbb{C}^2$ proportional to

$$\begin{pmatrix} -\sin\theta/2 \\ e^{i\phi}\cos\theta/2 \end{pmatrix} \tag{8.288}$$

f.) Using the parametrization of $SU(2)$ by Euler angles

$$u = \exp[-\frac{i}{2}\phi\sigma^3]\exp[-\frac{i}{2}\theta\sigma^2]\exp[-\frac{i}{2}\psi\sigma^3] \tag{8.289}$$

Thus, the lines (8.277) and (8.286) are the same and the lines (8.282) and (8.288) are the same. We can thereby view the fibers of the associated line bundles to $\pi : SU(2) \to S^2$ associated via the charge $\pm 1$ representations of $U(1)$ with the eigenlines of $P(\hat{x})$.

---

---

**Exercise** *Transition Functions For Associated Vector Bundles*

a.) Suppose that a principal $G$-bundle $\pi : P \to X$ has a chart with transition functions (8.242). Show that we can define natural local trivializations of an associated vector bundle associated to the representation $T : G \to GL(V)$ for a vector space $V$. Thus the trivializations are maps

$$\tilde{\phi}_\alpha : \tilde{\pi}^{-1}(\mathcal{U}_\alpha) \to \mathcal{U}_\alpha \times V \tag{8.290}$$

They are defined by declaring that if $\phi_\alpha(p) = (x, g)$ then

$$\tilde{\phi}_\alpha([(p, v)]) = (x, T(g)v) \tag{8.291}$$

b.) Show that on chart overlaps $x \in \mathcal{U}_{\alpha\beta}$ we have

$$\tilde{\phi}_\alpha \circ \tilde{\phi}_\beta^{-1}(x, v_\beta) = (x, T(g_{\alpha\beta}(x))v_\beta) := (x, v_\alpha) \tag{8.292}$$

c.) Conclude that a globally well-defined continuous section of the associated bundle is equivalent to a set of continuous maps $s_\alpha : \mathcal{U}_\alpha \to V$ such that for $x \in \mathcal{U}_{\alpha\beta}$

$$s_\alpha(x) = T(g_{\alpha\beta}(x)) \tag{8.293}$$

---

## 9. Centralizer Subgroups And Counting Conjugacy Classes

**Definition 9.1**: Let $g \in G$, the *centralizer subgroup* of $g$, (also known as the *normalizer subgroup* ), denoted, $Z(g)$, is defined to be:

$$Z(g) := \{h \in G | hg = gh\} = \{h \in G | hgh^{-1} = g\} \tag{9.1}$$

---

**Exercise** *Due Diligence*
a.) Check that $Z(g) \subset G$ is a subgroup.
b.) Show that $g^n \in Z(g)$ for any integer $n$.
c.) If $g_1 = g_0 g_2 g_0^{-1}$ show that $Z(g_1) = g_0 Z(g_2) g_0^{-1}$.
d.) Show that

$$Z(G) = \cap_{g \in G} Z(g) \tag{9.2}$$

---

---

**Exercise** *Is $Z(g)$ always an Abelian group?*
a.) Show that $Z(1) = G$. Answer the above question.
b.) Show that the centralizer of the transposition (12) in $S_n$ for $n \leq 3$ is isomorphic to $S_2$.
c.) Show that the centralizer of the transposition (12) in $S_n$ for $n \geq 4$ is isomorphic to $S_2 \times S_{n-2}$.

---

Recall that $C(g)$ denotes the conjugacy class of $g$. Using the Stabilizer-Orbit theorem we can establish a 1-1 correspondence between $C(g)$ and the cosets of $G/Z(g)$. As in the proof of that theorem we have a map $\psi : G/Z(g) \to C(g)$ by

$$\psi : g_i Z(g) \to g_i g g_i^{-1} \in C(g) \tag{9.3}$$

It is 1-1 and onto.

Since conjugacy is an equivalence relation $G$ decomposes as a disjoint union of the orbits, which in this case are the conjugacy classes. When $G$ is a finite group this decomposition leads to some useful theorems based on simple counting ideas. When $|G|$ is finite we can usefully write:

$$|G| = \sum_{conj.\ classes} |C(g)| \tag{9.4}$$

The sum is over distinct conjugacy classes. What is $g$ in this formula? For each class we may choose any representative element from that class.

Now, if $G$ is finite, then by the above 1-1 correspondence we may write:

$$|C(g)| = \frac{|G|}{|Z(g)|} \tag{9.5}$$

which allows us to write the above decomposition of $|G|$ in a useful form sometimes called the *class equation*:

$$|G| = \sum_{conj.\ classes} \frac{|G|}{|Z(g)|} \tag{9.6}$$

Again, we sum over a complete set of distinct non-conjugate elements $g$. Which $g$ we choose from each conjugacy class does not matter since if $g_1 = hg_2h^{-1}$ then $Z(g_1) = hZ(g_2)h^{-1}$ are conjugate groups, and hence have the same order. So, for each distinct conjugacy class we just choose any element we like.

### 9.1 $0+1$-Dimensional Gauge Theory

Recall the definition above of a morphism of bundles, and an isomorphism of bundles. An automorphism of a principal $G$-bundle $\pi : P \to X$ is an isomorphism of the principal bundle with itself. Since the composition of automorphisms is an automorphism, and automorphisms are invertible, and the identity is an automorphism the set of automorphisms form a group under composition, called the *group of automorphisms of the bundle* $\pi : P \to X$. Recall the bundles $P_{g_0} \to S^1$ determined by the group element $g_0$. As we showed above, the invertible bundle maps are of the form $\psi_h : P_{g_0} \to P_{hg_0h^{-1}}$. Therefore, the isomorphism classes of $G$-bundles over the circle are in 1-1 correspondence with the conjugacy classes in $G$, and the group of automorphisms of $P_{g_0}$ is precisely the centralizer group $Z(g_0)$.

A *gauge theory* is a physical theory where physical quantities are defined by summing over principal $G$-bundles with connection. We haven't defined the term "connection," but for principal $G$-bundles with discrete group there is a unique connection so we can discuss those here. [124] In the case of $0+1$ dimensional gauge theory a basic quantity of interest is the "partition function" on a one-dimensional manifold. The only closed connected one-dimensional manifold is the circle. So we have

$$Z(S^1) = \sum F(P_g) \tag{9.7}$$

where the sum is over isomorphism classes of principal $G$ bundles over the circle. and $F$ is a function on the set of such bundles, or equivalently, a function on all the principal $G$-bundles that only depends on the isomorphism class. We call this a *gauge invariant Boltzman factor.* The standard physical Boltzman factors involve curvature and holonomy. In this setting there is no curvature, so $F$ should be proportional to a character of $g$ in some representation.

In gauge theory we must also divide by the "volume" of the group of automorphisms of the bundle so

$$Z(S^1) = \sum_{cc} \frac{\chi_\rho(g)}{|Z(g)|} \tag{9.8}$$

for some character $\chi_\rho$. We will see later from the orthogonality relations for characters that the sum is zero unless $\rho$ contains some copies of the trivial representation, so we might as well take $\chi_\rho(g) = 1$.

Now, with the choice $F(g) = 1$ we can use the stabilizer-orbit theorem to rewrite our partition function as:

$$Z(S^1) = \sum_{cc} \frac{1}{|Z(g)|} = \frac{1}{|G|} \sum_{g \in G} 1 \tag{9.9}$$

[124] See section 18 below for a discussion of gauge theory that uses minimal prerequisites and is sufficient to understand this remark.

We can interpret the second sum as a sum over all the $G$-bundles $P_g$ weighted by 1 and divided by the full "volume" $|G|$ of the gauge group. So

$$Z(S^1) = \frac{1}{|G|} \sum_{P_g} 1 \tag{9.10}$$

Of course, by the counting formula above we see that $Z(S^1) = 1$. The Hilbert space of this theory is one-dimensional. These results might seem to be unexciting, but some of the conceptual ideas we have just used are very powerful and do produce exciting results in richer contexts.

## 9.2 Three Mathematical Applications Of The Counting Principle

In this section let $p$ be a prime number.

Application 1:

**Theorem**: If $|G| = p^n$ then the center is nontrivial, i.e., $Z(G) \neq \{1\}$.

*Proof*: Observe that an element $g$ is central *if and only if* $C(g) = \{g\}$ has order 1. Now let us use the class equation. We can usefully split up the sum over conjugacy classes as a sum over the center and the rest:

$$|G| = |Z(G)| + \sum_i |\mathcal{C}_i| \tag{9.11}$$

where the sum over $i$ is a sum over the distinct conjugacy classes more than one element. As we noted above, by the stabilizer orbit theorem

$$|\mathcal{C}_i| = \frac{|G|}{|Z(g_i)|} \tag{9.12}$$

where $g_i$ is any element of the conjugacy class $\mathcal{C}_i$. But, for these conjugacy classes $|Z(g_i)| < |G|$ and by Lagrange's theorem, and the assumption that $p$ is prime, $|Z(g_i)| = p^{n-n_i}$ for some $n_i < n$. Therefore, the second term on the RHS of (9.11) is divisible by $p$ and hence $p||Z(G)|$. ♠

Application 2: *Cauchy's theorem*:

In a similar style, we can prove the very useful:

**Theorem**: If $p$ divides $|G|$ then there is an element $g \in G$, $g \neq 1$ with order $p$.

*Proof 1*: This is a nice application of the stabilizer-orbit theorem. Consider the set

♣Proof 1 is not really an application of the class equation, rather it is an application of stabilizer-orbit. ♣

$$X = \{(g_1, \dots, g_p) | g_1 \cdots g_p = 1\} \subset G^p \tag{9.13}$$

Note that the cyclic group $\mathbb{Z}_p$ acts on this set with the standard generator acting by

$$\omega \cdot (g_1, \dots, g_p) = (g_p, g_1, g_2, \dots, g_{p-1}) \tag{9.14}$$

A fixed point of the $\mathbb{Z}_p$-action corresponds to an element of the form $(g, \dots, g)$ such that $g^p = 1$. If $g \neq 1$ then this corresponds to an element of order $p$. Now, by the stabilizer-orbit theorem, the orbits of any $\mathbb{Z}_p$ action (on any set) have cardinality either 1 or $p$. Let $N_1$ be the number of orbits of length one and let $N_p$ be the number of orbits of length $p$. Note that the order of $X$ is just $|G|^{p-1}$ since one can always solve for $g_p$ in terms of $g_1, \dots, g_{p-1}$. Then, by the counting principle we have:

$$|G|^{p-1} = N_1 + pN_p \tag{9.15}$$

It follows that $p$ divides $N_1$. Also $N_1 > 0$ because $(1, ..., 1)$ is a fixed point of the $\mathbb{Z}_p$ action. Therefore $N_1 = kp > 1$ and hence there are other fixed points, i.e. there are group elements of order $p$. In fact, there must be at least $(p-1)$ of them. ♠

*Proof 2*: We can also prove Cauchy's theorem using induction on the order of $G$, dividing the proof into two cases: First we consider the case where $G$ is Abelian and then the case where it is nonabelian.

Case 1: $G$ is Abelian:

If $|G| = p$ then $G$ is cyclic and the statement is obvious: Any generator has order $p$. More generally, note that if $G$ is a cyclic group $\mathbb{Z}/N\mathbb{Z}$ with $N > p$ and $p$ divides $N$ then $\overline{N/p} \in \mathbb{Z}/N\mathbb{Z}$ has order $p$. This establishes the result for cyclic groups.

Now suppose our Abelian group has order $|G| > p$. Choose an element $g_0 \neq 1$ and suppose that $g_0$ does not have order $p$. Let $H = \langle g_0 \rangle$. If $H = G$ then $G$ would be cyclic but then as we just saw, it would have an element of order $p$. So now assume $H$ is a proper subgroup of $G$. If $p$ divides $|H|$ then $H$ (and hence $G$) has an element of order $p$ by the inductive hypothesis. If $p$ does not divide $|H|$ then we consider the group $G/H$. But this has order strictly less than $|G|$ and $p$ divides the order of $G/H$. So there is an element $aH$ of order $p$ meaning $a^p = g_0^x$ for some $x$. If $g_0^x = 1$ we are done. If not then there is some smallest positive integer $y$ so that $g_0^{xy} = 1$ but then $a^y$ has order $p$. We have now proved Cauchy's theorem for abelian groups.

Case 2: $G$ is non-Abelian: By the class equation we can write

$$|G| = |Z(G)| + \sum{}' \frac{|G|}{|Z(g_i)|} \tag{9.16}$$

If $p$ divides the order of the centralizer $Z(G)$ then we can apply our previous result about Cauchy's theorem for Abelian groups. If $p$ does not divide $Z(G)$ then there must be some $g_i$ so that $p$ does not divide $\frac{|G|}{|Z(g_i)|}$ but this means $p$ divides $|Z(g_i)|$, but now by the inductive hypothesis $Z(g_i)$, and hence $G$ has an element of order $p$. This completes the proof. ♠

Application 3: *Sylow's theorem*:

Finally, as a third application we give a simple proof of Sylow's first theorem: If $p$ is prime and $p^k$ divides $|G|$ then $G$ has a subgroup of order $p^k$.

♣Proof 1 is not really an application of the class equation, rather it is an application of stabilizer-orbit. ♣

*Proof 1*: The first proof is again an application of the stabilizer-orbit theorem. [125] Suppose $|G| = p^{k+r}u$ with $gcd(u,p) = 1$ and $r \geq 0$ and $k > 0$. We will show that $G$ has a subgroup of order $p^k$. Consider the power set $\mathcal{P}(G)$, namely the set of all subsets of $G$, and consider the subset of $\mathcal{P}(G)$ of all subsets (not subgroups!) of $G$ of cardinality $p^k$. Call this set of subsets $\mathcal{P}(G,p^k)$. The cardinality of $\mathcal{P}(G,p^k)$ is clearly:

$$|\mathcal{P}(G,p^k)| = \binom{p^{k+r}u}{p^k} = p^r u \prod_{j=1}^{p^k-1} \frac{p^{k+r}u/j - 1}{p^k/j - 1} \tag{9.17}$$

In the product we have a ratio of rational numbers of the form $p^{k+r}u/j-1$ (the denominator is a special case of this form). Any rational number $r$ can be expressed as a product of prime powers $r = \prod_{\tilde{p} \text{ prime}} \tilde{p}^{v_{\tilde{p}}(r)}$ where the $v_{\tilde{p}}(r) \in \mathbb{Z}$ is known as the *valuation of* $r$ *at* $\tilde{p}$ and the product runs over all primes $\tilde{p}$. Now, given a specific prime $p$, note that if $a, b$ are relatively prime to $p$ then

$$\frac{p^k a}{b} - 1 = \frac{p^k a - b}{b} \tag{9.18}$$

and hence for such rational numbers $r$ the integer $v_p(r) = 0$ for the prime $p$. It follows that $p^r$ divides $\mathcal{P}(G,p^k)$ and that it is the maximal power which does so.

Now note that $G$ acts on $\mathcal{P}(G,p^k)$ via:

$$\phi_g : S \mapsto g \cdot S := \{gh | h \in S\} \tag{9.19}$$

where we are denoting an element of $\mathcal{P}(G,p^k)$ by $S$. Consider the stabilizer subgroup $G^S$ of any $S \in \mathcal{P}(G,p^k)$. Note that if $h \in S$ then every element $g \cdot h \in S$ for $g \in G^S$. (Why? Because $g \cdot S = S$ if $g \in G^S$.) But this means that $G^S \cdot h$ is a subset of $S$. Since the left-$G$ action is free

$$|G^S| = |G^S \cdot \alpha| \leq |S| = p^k \tag{9.20}$$

We now aim to show that some stabilizer group $G^S$ has order exactly $p^k$. This will be our subgroup predicted by Sylow's theorem. Suppose, on the contrary that no stabilizer group has order $p^k$. Then every stabilizer group satisfies $|G^S| < p^k$, and therefore it is divisible at most by $p^{k-1}$. Now, by the stabilizer-orbit theorem

$$|G| = |G^S| \cdot |\mathcal{O}(S)| \tag{9.21}$$

where $\mathcal{O}(S)$ is the $G$-orbit through $S$. Now $p^{k+r}$ divides $|G|$ and if $G^S$ is divisible by at most $p^{k-1}$ then $p^s$ divides $|\mathcal{O}(S)|$ for $s > r$. But now

$$|\mathcal{P}(G,p^k)| = \sum_{\text{distinct orbits}} |\mathcal{O}(S_i)| \tag{9.22}$$

If all the orbits on the RHS were divisible by $p^s$ with $s > r$ then $|\mathcal{O}(S_i)|$ would be divisible by $p^s$ with $s > r$. But this is not true. Therefore, some orbit is divisible by $p^r$ and no higher power. Therefore some $|G^S|$ is divisible by $p^k$, therefore $|G^S| = p^k$. Since $G^S$ is a stabilizer group it is a subgroup of $G$. ♠

[125] We are following the nice article on Wikipedia here.

*Proof 2*: The more conventional proof is similar to that of Cauchy's theorem. We work by induction on $|G|$, and divide the proof into two cases:

Case 1: $p$ divides the order of $Z(G)$.: By Cauchy's theorem $Z(G)$ has an element of order $p$ and hence a subgroup $N \subset Z(G)$ of order $p$. $N$ is clearly a normal subgroup of $G$ (being a subgroup of the center of $G$) so $G/N$ is a group. It is clearly of order $p^{k-1}m$. So, by the inductive hypothesis there is a subgroup $\bar{H} \subset G/N$ of order $p^{k-1}$. Now let $H = \{g \in G | gN \in \bar{H}\}$. It is not hard to show that $H$ is a a subgroup of $G$ containing $N$ and in fact $H/N = \bar{H}$. Therefore $|H| = p^k$, so $H$ is a $p$-Sylow subgroup of $G$.

Case 2: $p$ does not divide the order of $Z(G)$.: In this case, by the class equation $p$ must not divide $|C(g)| = |G|/|Z(g)|$ for some nontrivial conjugacy class $C(g)$. But that means that for such an element $g$ we must have that $p^k$ divides $|Z(g)| < |G|$. So $Z(g)$ has a $p$-Sylow subgroup which can serve as a $p$-Sylow subgroup of $G$. ♠

♣Again, there is a nice proof using the orbit-stabilizer theorem. See Wikipedia article. Give this in the section on Orbit-Stabilizer below? ♣

---

**Exercise**
If $p^k$ divides $|G|$ with $k > 1$ does it follow that there is an element of order $p^k$? [126]

---

---

**Exercise** *Groups Whose Order Is A Square Of A Prime Number*
If $|G| = p^2$ where $p$ is a prime then show that
1. $G$ is abelian
2. $G \cong \mathbb{Z}_p \times \mathbb{Z}_p$ or $\mathbb{Z}_{p^2}$.

---

---

**Exercise**
Write out the class equation for the groups $S_4$ and $S_5$.

---

---

**Exercise**
Find the centralizer $Z(g) \subset S_n$ of $g = (12 \dots n)$ in $S_n$.

---

---

**Exercise**

[126] *Answer*: NO! $\mathbb{Z}_p^k$ is a counterexample: It has order $p^k$ and every element has order $p$.

Prove that if $|G| = 15$ then $G = \mathbb{Z}/15\mathbb{Z}$.

---

---

**Exercise** *Groups whose order is a product of two primes*

Suppose that $G$ has order $pq$ where $p$ and $q$ are distinct primes. We assume WLOG that $p < q$. We now also assume that and $p$ does not divide $q - 1$.

a.) Show that $G$ is isomorphic to $\mathbb{Z}_{pq}$.

Warning!! This is hard. [127]

b.) Why is it important to say that $p$ does not divide $q - 1$? [128]

c.) Show that this result implies that if a nonabelian group has odd order then the order must be $\geq 21$. (And in fact, there does exist a nonabelian group of order 21.)

---

## 10. Generators And Relations

The presentation (6.28) of the symmetric group is an example of presenting a group by *generators and relations.*

**Definition 10.1** A subset $\mathcal{S} \subset G$ is a *generating set* for a group if every element $g \in G$ can be written as a "word" or product of elements of $\mathcal{S}$. That is any element $g \in G$ can

[127] *Answer.* By Cauchy's theorem we know there is an element $a$ of order $p$ and an element $b$ of order $q$. We can easily reduce to the case the center of $G$ is trivial. In general the subgroup $Z(G)$ must have order $pq, p, q$, or 1. If $Z(G)$ has order $pq$ then $a$ and $b$ commute and $G \cong \mathbb{Z}_p \times \mathbb{Z}_q \cong \mathbb{Z}_{pq}$. If $|Z(G)|$ has order $p$ or $q$ then $G/Z(G)$ must be cyclic of order $q$ or $p$, respectively. Hence by an easy exercise above $G$ is cyclic. This leaves us with the hard case where $Z(G) = \{1\}$ is the trivial subgroup. Let us consider the conjugacy classes of the powers of $a$, $C(a)$, $C(a^2)$,.... Since $Z(a)$ has order at least $p$ and its order must divide $pq$ and it can't be the whole group (since $Z(G) = \{1\}$) it must be that $Z(a) = \{1, a, \dots, a^{p-1}\}$ and hence $C(a)$ has order $q$. Indeed, for <u>any</u> element $g \in G$ that is not the identity it must be that $Z(g)$ has order $p$ or $q$ and $C(g)$ has order $q$ or $p$. Now note that $Z(a) \supset Z(a^2) \supset \cdots$. So, as long as $a^x$ is not one, it must be that $Z(a^x) = Z(a)$ and $C(a^x)$ has order $q$. Now we claim that the different conjugacy classes $C(a)$, $C(a^2)$,..., $C(a^{p-1})$ are all <u>distinct</u>. The statement that these are distinct can be reduced to the statement that it is not possible to have $bab^{-1} = a^x$ for any $x$, so now we verify this latter statement. If it were the case that $bab^{-1} = x$ then since the general element of the conjugacy class is $b^j ab^{-j}$ the conjugacy class would have to be $\{a, a^x, a^{2x}, \dots, a^{(q-1)x}\}$. But that set must be the set $C(a) = \{a, a^2, \cdots, a^{p-1}\}$ of $p$ elements. Since $q > p$ it must be that $b^{j_1} ab^{-j_1} = b^{j_2} ab^{-j_2}$ where $1 \leq j_1, j_2 \leq (q-1)$ and $j_1 \neq j_2$. So we have to have $b^j ab^{-j} = a$ for some $1 \leq j \leq (q-1)$. But then $b^j \neq 1$. But then such an element $b^j$ would be in $Z(a)$. This is impossible. So we can never have $bab^{-1} = a^x$ and hence $C(a)$, $C(a^2)$,..., $C(a^{p-1})$ are all <u>distinct</u>. Now the class equation says that

$$pq = 1 + (p-1)q + X$$

where $X$ accounts for all the other conjugacy classes. As we have remarked these must have order $p$ or $q$ and hence $X = rp + sq$ for nonnegative integers $r, s$. But now

$$q - 1 = rp + sq$$

But this is impossible: If $s \geq 1$ the RHS is too large. So $s = 0$ but then $p$ would have to divide $q - 1$.

[128] *Answer*: Consider $p = 2$ and $q = 3$ and note that $S_3$ is not isomorphic to $\mathbb{Z}_6$.

be written in the form

$$g = s_{i_1} \cdots s_{i_r} \tag{10.1}$$

where, for each $1 \leq k \leq r$ we have $s_{i_k} \in \mathcal{S}$.

Every group $G$ has at least one generating set, namely $G$ itself, but this is rarely useful. A group is said to be *finitely generated* if there exists a generating set $\mathcal{S}$ that is finite. That is, there is a finite list of elements $\{s_1, \ldots s_n\}$ so that all elements of the group can be obtained by taking products – "words" – in the "letters" drawn from $\mathcal{S}$. For example, we have shown above that the symmetric group is finitely generated by the transpositions. Typical Lie groups such as $SU(n)$ or $SO(n, \kappa)$ (over $\kappa = \mathbb{R}, \mathbb{C}$) are not finitely generated.

A group is said to be presented by *generators and relations* if there is a generating set $\mathcal{S}$ and a set of words $R_\alpha$ in those generators - known as the *relations* - so that, in the group $R_\alpha = 1$ and, moreover, all the identities of the form "word in the generators = 1 " are consequences of the relations $R_\alpha = 1$. That is, equalities between different words in the generators can always be proven by repeated use of the identities $R_\alpha = 1$. [129]

♣Rearrange order. Free groups. Then Nielsen Schrier. Then define generators and relations properly. ♣

In general if we have a finitely generated group we write

$$G = \langle g_1, \ldots, g_n | R_1, \cdots R_r \rangle \tag{10.2}$$

where $R_i$ are words in the letters of $\mathcal{S}$ which will be set to 1. **ALL** other relations, that is, all other identities of the form $W = 1$ are supposed to be consequences of these relations.

**Remark**: It is convenient to exclude the unit 1 from $\mathcal{S}$ so that we can reduce some obvious redundancies in our words. Also we need to be careful about what we mean by a word in the generators with letters drawn from $\mathcal{S}$. Some people intend that such words contain any integer power $s^n$ of any generator $s \in \mathcal{S}$, where $s^0 = 1$ and, if $n < 0$, this means $(s^{-1})^{|n|}$. So if two consecutive letters are $s^n s^m$ we understand this as $s^{n+m}$ for any integers $n, m$. Alternatively, we can, for each generator $s$ introduce another generator $t$, which will play the role of $s^{-1}$ and then impose another relation $st = ts = 1_G$. A generating set that contains such an element $t$ for every generator $s$ is said to be *symmetric*.

**Example 3.1**: If $\mathcal{S}$ consists of one element $a$ then $F(\mathcal{S}) \cong \mathbb{Z}$. The isomorphism is given by mapping $n \in \mathbb{Z}$ to the word $a^n$.

**Example 3.2**: The most general group with one generator and one relation must be of the form:

$$\langle a | a^N = 1 \rangle \tag{10.3}$$

where $N$ is an integer and by replacing $a \to a^{-1}$ we can assume it is a positive integer. As an exercise, prove that this group is isomorphic to $\mathbb{Z}_N$ and $\mu_N$.

[129] See Jacobsen, *Basic Algebra I*, sec. 1.11 for a more precise definition. Note that if $R_\alpha = 1$ then also $gR_\alpha g^{-1} = 1$ and we can then insert $g^{-1}g$ between all the letters in $R_\alpha$ to get another identity. And then we can take products of expressions like $gR_\alpha g^{-1}$ to produce new identities. Technically - using concepts we will explain below - one considers the free group on the generating set. Then within that free group consider the normal closure: The smallest normal subgroup of the free group containing all the words $R_\alpha$. Then the group defined by generators and relations is the quotient of the free group by this normal subgroup.

**Example 3.3**: *Free groups.* If we impose no relations on the generating set $\mathcal{S}$ then we obtain what is known as the *free group on* $\mathcal{S}$, denoted $F(\mathcal{S})$. If $\mathcal{S}$ consists of one element then we just get $\mathbb{Z}$, as above. However, things are completely different if $\mathcal{S}$ consists of more than one element. For example, suppose we have two elements $a, b$. Then $F(\mathcal{S})$ is very complicated. A typical element looks like one of

$$\begin{aligned} &a^{n_1} b^{m_1} \cdots a^{n_k} \\ &a^{n_1} b^{m_1} \cdots b^{m_k} \\ &b^{n_1} a^{m_1} \cdots a^{n_k} \\ &b^{n_1} a^{m_1} \cdots b^{m_k} \end{aligned} \tag{10.4}$$

where $n_i, m_i$ are nonzero integers (positive or negative). Three nice general results on free groups are:

1. Two free groups $F(\mathcal{S}_1)$ and $F(\mathcal{S}_2)$ are isomorphic iff they have the same cardinality.
2. *Nielsen-Shreier theorem*: Any subgroup of a free group is free.
3. Every group has a presentation in terms of generators and relations. For the group $G$ we can consider the free group $F(G)$ with $\mathcal{S} = G$ as a set. There is then a natural homomorphism $\varphi : F(G) \to G$ where we take a word in elements of $G$ and map concatenation of letters to group multiplication in $G$. As we will see in Section *** below, the kernel of the homomorphism is a normal subgroup $K(G)$ so that $G \cong F(G)/K(G)$ and $K(G)$ are the relations in this presentation. This presentation can be incredibly inefficient and useless. (Think, for example, of Lie groups.)

Combinatorial group theorists use the notion of a *Cayley graph* to illustrate groups presented by generators and relations. Assuming that $1 \notin \mathcal{S}$ the Cayley graph is a graph whose vertices correspond to all group elements in $G$ and the oriented edges are drawn between $g_1$ and $g_2$ if there is an $s \in \mathcal{S}$ with $g_2 = g_1 s$. We label the edge by $s$. (If $\mathcal{S}$ is symmetric we can identify this edge with the edge from $g_2$ to $g_1$ labeled by $s^{-1}$.) For the free group on two elements this generates the graph shown in Figure 13.

**Example 3.4**: *Coxeter groups*: Let $m_{ij}$ by an $n \times n$ symmetric matrix whose entries are positive integers or $\infty$, such that $m_{ii} = 1$, $1 \leq i \leq n$, and $m_{ij} \geq 2$ or $m_{ij} = \infty$ for $i \neq j$. Then a *Coxeter group* is the group with generators and relations:

$$\langle s_1, \ldots, s_n | \forall i, j : (s_i s_j)^{m_{ij}} = 1 \rangle \tag{10.5}$$

where, if $m_{ij} = \infty$ we interpret this to mean there is no relation.

Note that since $m_{ii} = 1$ we have

$$s_i^2 = 1 \tag{10.6}$$

Quite generally, a group element that squares to 1 is called an *involution*. So all the generators of a Coxeter group are involutions. It then follows that if $m_{ij} = 2$ then $s_i$ and $s_j$ commute. If $m_{ij} = 3$ then the relation can also be written:

$$s_i s_j s_i = s_j s_i s_j \tag{10.7}$$

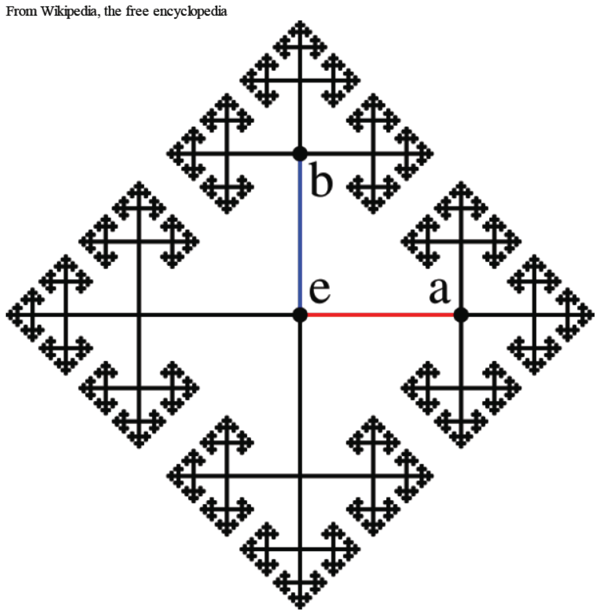



**Figure 13:** The Cayley graph for the free group on 2 generators $a$ and $b$.

A theorem of Coxeter's from the 1930's gives a classification of the finite Coxeter groups. [130] Coxeter found it useful to describe these groups by a diagrammatic notation: We draw a graph whose vertices correspond to the generators $s_i$. We draw an edge between vertices $i$ and $j$ if $m_{ij} \geq 3$. By convention the edges are labeled by $m_{ij}$ and if $m_{ij} = 3$ then the standard convention is to omit the label.

It turns out that the *finite* Coxeter groups can be classified. The corresponding Coxeter diagrams are shown in Figure 14.

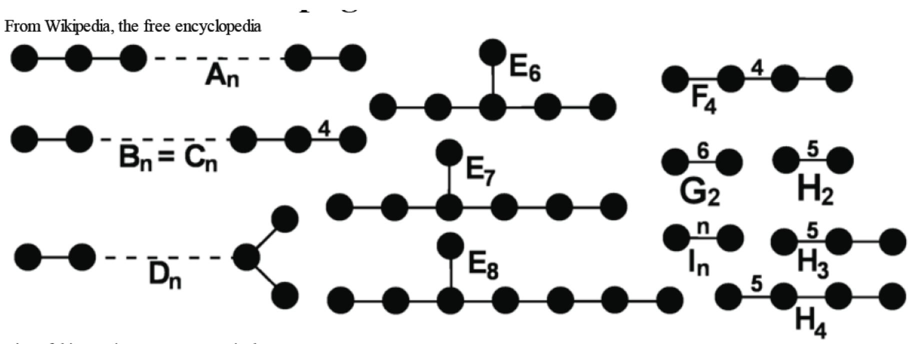



**Figure 14:** Coxeter's list of finite Coxeter groups. They are finite groups of reflections in some Euclidean space.

The finite Coxeter groups turn out to be isomorphic to concrete groups of *reflections* in some Euclidean space. That is, finite subgroups of $O(N)$ for some $N$. That is, there is

[130] For a quick summary see the expository note by D. Allcock at https://web.ma.utexas.edu/users/allcock/expos/reflec-classification.pdf.

some vector space $\mathbb{R}^N$ and collection of vectors $v_i \in \mathbb{R}^N$ with inner products

$$v_i \cdot v_j = -2\cos(\frac{\pi}{m_{i,j}}) \tag{10.8}$$

so that the group generated by reflections in the plane orthogonal to the vectors $v_i$:

$$P_{v_i} : v \mapsto v - \frac{2v \cdot v_i}{v_i \cdot v_i} v_i \tag{10.9}$$

is a finite group isomorphic to the Coxeter group with matrix $m_{i,j}$. (Note that since $m_{i,i} = 1$ we have $v_i^2 = 2$ and $P_{v_i}(v) = v - (v \cdot v_i)v_i$.)

Note that, if $P_v$ is the Euclidean reflection in the plane orthogonal to $v$ then $P_{v_1} \circ P_{v_2}$ is just rotation in the plane spanned by $v_1, v_2$ by an angle $2\theta$ where the angle between $v_1$ and $v_2$ is $\theta$. To prove this, note that $P_{v_1} \circ P_{v_2}$ clearly leaves all vectors in the plane orthogonal to $v_1, v_2$ fixed. Now represent vectors in a 2-dimensional Euclidean plane by complex numbers, but view $\mathbb{C}$ as a real vector space. WLOG take $v = e^{\mathrm{i}\theta}$. Then $P_v$ is the transformation:

$$P_v : z \mapsto -e^{2\mathrm{i}\theta}\bar{z} \tag{10.10}$$

Note that this is a linear transformation of real vector spaces: $P_v(a_1 z_1 + a_2 z_2) = a_1 P_v(z_1) + a_2 P_v(z_2)$ if $a_1, a_2 \in \mathbb{R}$. To check this formula note that if $z = e^{\mathrm{i}\theta}$ then $P_v(z) = -z$ and if $z = \mathrm{i}e^{\mathrm{i}\theta}$ is in the orthogonal hyperplane to $v$ then $P_v(z) = z$.

Now if $v_a = e^{\mathrm{i}\theta_a}$, $a = 1, 2$, it is an easy matter to compute:

$$\begin{aligned} z &\xrightarrow{P_{v_2}} -e^{2\mathrm{i}\theta_2}\bar{z} \\ &\xrightarrow{P_{v_1}} -e^{2\mathrm{i}\theta_1}\,\overline{(-e^{2\mathrm{i}\theta_2}\bar{z})} \\ &= e^{2\mathrm{i}(\theta_1-\theta_2)}z \end{aligned} \tag{10.11}$$

*So: The product of reflections in the hyper-planes orthogonal to two vectors at an angle $\theta$ is a rotation by an angle $2\theta$ in the plane spanned by the two vectors.*

We will meet some of these groups again later as Weyl groups of simple Lie groups. We have, in fact, already met two of these groups! The case $A_n$ turns out to be isomorphic to the symmetric group $S_{n+1}$. [131] In this case we have seen that the elementary generators $\sigma_i = (i, i+1)$, $1 \le i \le n$ indeed satisfy the Coxeter relations:

$$\begin{aligned} \sigma_i^2 &= 1 \\ (\sigma_i\sigma_{i+1})^3 &= 1 \qquad 1 \le i \le n-1 \\ (\sigma_i\sigma_j)^2 &= 1 \qquad |i-j| > 1 \end{aligned} \tag{10.12}$$

Now consider the standard basis $e_i$ for $\mathbb{R}^{n+1}$, $1 \le i \le n+1$ and consider the vectors:

$$\alpha_i = e_i - e_{i+1} \tag{10.13}$$

[131] The notation here is standard but exceedingly unfortunate and confusing!!! Here $A_n$ does <u>*NOT*</u> refer to the alternating group! It refers to Cartan's classification of simple Lie groups and the Coxeter group with this label is in fact isomorphic to $S_{n+1}$.

which have the inner products: $\alpha_i^2 = 2$ and $\alpha_i \cdot \alpha_{i\pm1} = -1$ (so they are at angle $2\pi/3$) and all other inner products vanish. This is summarized in the matrix:

$$\alpha_i \cdot \alpha_j = C_{ij} = 2\delta_{i,j} - \delta_{i,j+1} - \delta_{i,j-1} \tag{10.14}$$

Then the map $s_i \to P_{\alpha_i}$ is an isomorphism of the Coxeter group $A_n$ with a subgroup of $O(n+1)$. Moreover, one computes that

$$P_{\alpha_i}(e_j) = \begin{cases} e_j & j \neq i, i+1 \\ e_{i+1} & j = i \\ e_i & j = i+1 \end{cases} \tag{10.15}$$

So, referring to equation (6.10) we see that this is just the permutation action of $\sigma_i$ on the standard basis of $\mathbb{R}^{n+1}$. This makes clear that the Coxeter group is isomorphic to the symmetric group $S_{n+1}$.

♣A presentation of the Monster in terms of generators and relations is known.(Atlas) Give it here? ♣

**Remarks**

1. One very practical use of having a group presented in terms of generators and relations is in the construction of homomorphisms. If one is constructing a homomorphism $\phi : G_1 \to G_2$, then it suffices to say what elements the generators map to. That is, if $g_i$ are generators of $G_1$ we can fully specify a homomorphism by choosing elements $g_i' \in G_2$ (not necessarily generators) and declaring

$$\phi(g_i) = g_i' \quad . \tag{10.16}$$

However, we cannot choose the $g_i'$ arbitrarily. Rather, the $g_i'$ must satisfy the same relations as the $g_i$. This puts useful constraints on what homomorphisms you can write down. For example, using this idea you can prove that there is no nontrivial homomorphism $\phi : \mathbb{Z}_N \to \mathbb{Z}$.

2. In general it is hard to say much about a group given a presentation in terms of generators and relations. For example, it is not even obvious, in general, if the group is the trivial group! This is part of the famous "word problem for groups." There are finitely presented groups where the problem of saying whether two words represent the same element is undecidable! [132] However, for many important finitely presented groups the word problem can be solved. Indeed, the word problem was first formulated by Max Dehn in 1911 and solved by him for the surface groups discussed below.

♣It would be more effective here to give an example of a set of generators and relations that is actually isomorphic to the trivial group - but not obviously so. ♣

3. Nevertheless, there are four Tietze transformations (adding/removing a relation, adding/removing a generator) which can transform one presentation of a group to a different presentation of an isomorphic group. It is a theorem [REF!] that any two presentations can be related by a finite sequence of Tietze transformations. How is

[132]The Wikipedia article on "Word problem for groups," is useful.

this compatible with the previous remark? The point is that the number $f(n)$ of such transformations needed to transform a presentation of the trivial group with $n$ relations into the trivial presentation grows faster than any recursive function of $n$.

4. It turns out that the case of Coxeter groups $B_n = C_n$ are isomorphic to the group of symmetric permutations $\mathcal{W}B_n \subset S_{2n}$ discussed in card-shuffling. The Coxeter diagrams are very similar to the *Dynkin diagrams* that are used to label finite dimensional simple Lie algebras over the complex numbers except that $H_n$ and $I_2(n)$ for $n \geq 7$ are not Weyl groups of Lie algebras.

---

**Exercise** *Homomorphisms involving* $\mathbb{Z}_N$ *and* $\mathbb{Z}$
a.) Write a nontrivial homomorphism $\mu : \mathbb{Z} \to \mathbb{Z}_N$.
b.) Show that there is no nontrivial homomorphism $\mu : \mathbb{Z}_N \to \mathbb{Z}$. [133]
c.) Find the most general homomorphism $\mu : \mathbb{Z} \to \mathbb{Z}$.
d.) Find the most general homomorphism $\mu : \mathbb{Z}_N \to \mathbb{Z}_N$.

---

---

**Exercise** *One generator and many relations*
Suppose a group has a single generator $g$ but many relations. Describe this group. [134]

---

---

**Exercise** *Abelianizations*
Consider the free group on 2 generators. What is the abelianization? [135]

---

---

**Exercise** *The Regular n-gon*

Consider the symmetry group of the regular $n$-gon in the plane. Show that this is isomorphic to the Coxeter group with matrix

$$m_{ij} = \begin{pmatrix} 1 & n \\ n & 1 \end{pmatrix} \tag{10.17}$$

[133] *Answer*: Since $\mathbb{Z}_N$ can be generated by one element, say $\bar{1}$, it suffices to say what the value of $\phi(\bar{1})$ is. The trivial homomorphism takes the generator to zero: $\phi(\bar{1}) = 0 \in \mathbb{Z}$ and hence takes every element to zero. On the other hand, if $\phi(\bar{1}) = k$ is a nonzero integer, then $Nk = N\phi(\bar{1}) = \phi(N\bar{1}) = \phi(\bar{0}) = 0$, a contradiction. So there is no nontrivial homomorphism.

[134] *Answer*: Let the generator be $g$. The relations must be of the form $g^{k_1} = 1, \dots, g^{k_n} = 1$ for some integers $k_1, \dots, k_n$, and WLOG we can assume they are positive. Then it is not hard to see, using the Chinese remainder theorem (see below) that the group is isomorphic to $\mathbb{Z}_N$ where $N = gcd(k_1, ..., k_n)$.

[135] *Answer*: $\mathbb{Z}^2$.

---

**Exercise** *Simple Roots Of* $SU(n+1)$

a.) Verify equation (10.14). The matrix $C_{ij}$ is known as a *Cartan matrix* of $SU(n+1)$.

b.) Show that the vectors $\alpha_i$ are all orthogonal to the all-one vector: $v = (1, \dots, 1)$ and that they span the orthogonal complement of $v$.

c.) Show that the permutation representation of $S_{n+1}$ separately preserves $v$ and the orthogonal complement of $v$. Thus, $\mathbb{R}^{n+1}$ gives what is known as a *reducible representation* of $S_{n+1}$.

d.) Compute the action of $P_{\alpha_i}$ on $\alpha_j$. Give the matrix representation relative to the ordered basis $\{\alpha_1, \dots, \alpha_n\}$. [136]

---

---

**Exercise** Show that

$$\langle a, b | a^3 = 1, b^2 = 1, abab = 1 \rangle \tag{10.18}$$

is a presentation of $S_3$

---

---

**Exercise**

Consider the group with presentation:

$$\langle T, S | (ST)^3 = 1, S^2 = 1 \rangle \tag{10.19}$$

Is this group finite or infinite?

This group plays a very important role in string theory.

---

---

**Exercise** *Bounds on the minimal number of generators of a finite group*

Suppose we have a set of finite groups $G_1, G_2, G_3, \dots$ with a minimal set of generators $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots$ of cardinality $|\mathcal{S}_k| = k$. Show that $2|G_k| \le |G_{k+1}|$ and hence as $k \to \infty$ the order $|G_k|$ must grow at least as fast as $2^k$.

[136] *Answer* The matrix is a diagonal matrix of 1's except on the $3 \times 3$ block for rows and columns $i-1, i, i+1$ where it looks like

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

**Remark**: Denote the smallest cardinality of a set of generators of $G$ by $d(G)$. If $G$ is a finite and transitive permutation subgroup of $S_n$ (meaning it acts transitively on some set $X$) then there is a constant $C$ such that

$$d(G) \leq C\frac{n}{\sqrt{\log n}} \tag{10.20}$$

and if $G$ is a primitive permutation group, meaning that it acts on a set $X$ such that it does not preserve any nontrivial disjoint decomposition of $X$, then there is a constant $C$ so that if $n \geq 3$:

$$d(G) \leq C\frac{\log n}{\sqrt{\log\log n}} \tag{10.21}$$

Moreover, these results are asymptotically the best possible. For a review of such results see. [137]

---

---

**Exercise** *Generators And Relations For Products Of Groups*

Suppose you are given groups $G_1$ and $G_2$ in terms of generators and relations. Write a set of generators and relations for the product group $G_1 \times G_2$. [138]

---

### 10.1 Example Of Generators And Relations: Fundamental Groups In Topology

Presentations in terms of generators and relations is very common when discussing the *fundamental group* of a topological space $X$.

This subsection assumes some knowledge of topological spaces and the idea of a homotopy. Without trying to be too precise we choose a basepoint $x_0 \in X$ and let $\pi_1(X, x_0)$ be the set of closed paths in $X$, beginning and ending at $x_0$ where we identify two paths if they can be continuously deformed into each other. We can define a group multiplication by concatenation of paths. Inverses exist since we can run paths backwards. The following subsubsection contains more precise definitions.

♣Need to fix pictures and change $p_0$ to $x_0$. ♣

#### 10.1.1 The Fundamental Group Of A Topological Space

Choose a point $x_0 \in X$. The fundamental group $\pi_1(X, x_0)$ based at $x_0$ is, as a set, the set of homotopy classes of closed curves.

That is we consider <u>continuous</u> maps:

$$f : ([0, 1], \{0, 1\}) \to (X, \{x_0\}) \tag{10.22}$$

These define paths in $X$ with beginning and ending point fixed at $x_0$. The path must be traveled in time 1.

[137] F. Menegazzo, "The Number of Generators of a Finite Group," Irish Math. Soc. Bulletin 50 (2003) pp. 117-128

[138] *Answer*: If $G_1 = \langle g_i | R_i \rangle$ and $G_2 = \langle h_a | S_a \rangle$ then $G_1 \times G_2 = \langle g_i, h_a | R_i, S_a, g_i h_a g_i^{-1} h_a^{-1} = 1 \rangle$.

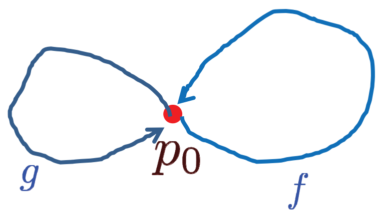



**Figure 15:** Two loops $f, g$ with basepoint at $x_0$.

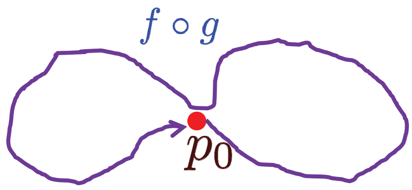



**Figure 16:** The concatenation of the looops $f \star g$. Note that the "later" loop is written on the right. This is generally a more convenient convention when working with homotopy and monodromy. In order for $f \star g$ to be a map from $[0,1]$ into $X$ we should run each of the individual loops at "twice the speed" so that at time $t = 1/2$ the loop returns to $x_0$. However, in homotoping $f \star g$ there is no reason why the point at $t = 1/2$ has to stay at $x_0$.

We say that two such paths $f_0, f_1$ are *homotopic* if there is a continuous map

$$F : [0,1] \times [0,1] \to X \tag{10.23}$$

such that

1. $F(0,t) = f_0(t)$ and $F(1,t) = f_1(t)$
2. $F(s,0) = F(s,1) = x_0$

If we define $f_s(t) := F(s,t)$ and consider $f_s(t)$ as a path in $t$ at fixed $s$ then, as we vary $s$ we are describing a path of paths.

Now, homotopy of paths in $X$ is an equivalence relation. [139] We denote by $[f]$ the equivalence class of a path $f$ and we denote the set of such equivalence class by $\pi_1(X, x_0)$. We will see that this set has a natural and beautiful group structure.

[139] See section 2 above for this notion.

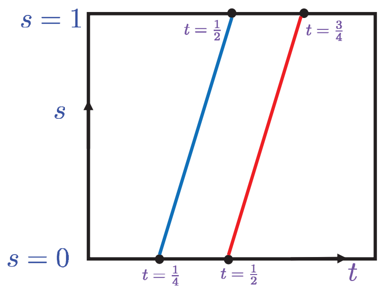



**Figure 17:** The homotopy demonstrating that loop concatenation is an associative multiplication on homotopy equivalence classes of closed loops. The blue line is $s = 4t - 1$ and the red line is $s = 4t - 2$.

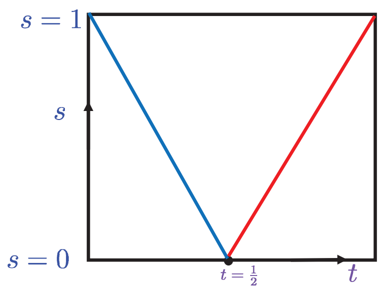



**Figure 18:** The homotopy demonstrating that the loop $g(t) = f(1-t)$ provides a representative for the inverse of $[f(t)]$.

We can define a group structure on $\pi_1(X, x_0)$ by concatenating curves as in Figure 16 and rescaling the time variable so that it runs from 0 to 1. In equations we have

$$f_1 \star f_2(t) := \begin{cases} f_1(2t) & 0 \le t \le \frac{1}{2} \\ f_2(2t-1) & \frac{1}{2} \le t \le 1 \end{cases} \tag{10.24}$$

**Remarks**

1. Note that we are composing successive paths *on the right*. This is slightly nonstandard but a nice convention when working with monodromy and path ordered expontentials of gauge fields - one of the main physical applications.

2. Note well that $(f_1 \star f_2) \star f_3$ is $NOT$ the same path as $f_1 \star (f_2 \star f_3)$. This observation ultimately leads to the notion of $A_\infty$ spaces.

3. For the moment we simply notice that if we mod out by homotopy then we have a well-defined product on homotopy classes in $\pi_1(X, x_0)$

$$[f_1] \cdot [f_2] := [f_1 \star f_2] \tag{10.25}$$

and the virtue of passing to homotopy classes is that now the product (10.25) is in fact associative. The proof is in Figure 17. Written out in excruciating detail the homotopy is

$$F(s,t) = \begin{cases} f_1(\frac{4}{s+1}t) & 0 \leq t \leq \frac{s+1}{4} \\ f_2(4t - (s+1)) & \frac{s+1}{4} \leq t \leq \frac{s+2}{4} \\ f_3(\frac{4}{2-s}(t - \frac{s+2}{4})) & \frac{s+2}{4} \leq t \leq 1 \end{cases} \tag{10.26}$$

4. Since we have an associative product on $\pi_1(X, x_0)$ we are now ready to define a group structure. The identity element is clearly given by the (homotopy class of the) constant loop: $f(t) = x_0$. If a homotopy class is represented by a loop $f(t)$ then the inverse is represented by running the loop backwards: $g(t) := f(1-t)$. The two are joined at $t = 1/2$, and since this is in the open interval $(0, 1)$ the image can be deformed away from $x_0$. See Figure 18. In equations, there is a homotopy of $f \star g$ with the constant loop given by

$$F(s,t) = \begin{cases} f(2t) & t \leq \frac{1-s}{2} \\ f(1-s) & \frac{1-s}{2} \leq t \leq \frac{1+s}{2} \\ f(2-2t) & \frac{1+s}{2} \leq t \leq 1 \end{cases} \tag{10.27}$$

> Thus, with the group operation defined by concatenation in the sense of (10.25) the set of homotopy classes $\pi_1(X, x_0)$ is a group. It is known as the fundamental group based at $x_0$.

5. A connected space such that $\pi_1(X, x_0)$ is the trivial group is called *simply connected.*

**A Basic Example: The Fundamental Group Of The Circle**:

The first and most basic example of a nontrivial fundamental group is the fundamental group of the circle. It should be intuitively clear that

$$\pi_1(S^1, x_0) \cong \mathbb{Z} \tag{10.28}$$

which just measures the number of times the path winds around the circle. The sign of the integer takes into account winding clockwise vs. counterclockwise.

Let us amplify a little on how one proves this basic fact. We will just give the main idea. For a thorough and careful proof see A. Hatcher's book on Algebraic Topology, or Section 13.2 of http://www.physics.rutgers.edu/∼gmoore/511Fall2014/Physics511-2014-Ch2-Topology.pdf

We have a standard map

$$p : \mathbb{R} \to S^1 \tag{10.29}$$

given by $p(x) = e^{2\pi i x}$. Note that the inverse image of any phase $e^{2\pi i s}$ with $s \in \mathbb{R}$ is the set of real numbers $s + \mathbb{Z}$. So $p$ is the map that identifies the orbits of the $\mathbb{Z}$-action by translation on $\mathbb{R}$ with $S^1$. That is $S^1 \cong \mathbb{R}/\mathbb{Z}$. Now, suppose we have a map $\bar{f} : [0,1] \to S^1$. We claim that there is a map $f : [0,1] \to \mathbb{R}$ so that $\bar{f}$ "factors through $p$," meaning that there is a map $f : [0,1] \to \mathbb{R}$ such that:

$$\bar{f} = p \circ f \tag{10.30}$$

The problem of finding such a map $f$ is nicely expressed in terms of diagrams. One is trying to complete the following diagram to make it a commutative diagram by finding a suitable map $f$ to use on the dashed line in:

$$\begin{array}{ccc} & & \mathbb{R} \\ & \overset{f}{\nearrow} & \downarrow p \\ [0,1] & \xrightarrow{\bar{f}} & S^1 \end{array} \tag{10.31}$$

In other words, given $\bar{f}$ one is trying to find a map $f : [0,1] \to \mathbb{R}$ so that

$$e^{2\pi i f(x)} = \bar{f}(x) \tag{10.32}$$

So, $f(x)$ is a logarithm, but a logarithm is not single-valued. Nevertheless, in sufficiently small open sets of $[0,1]$ (small enough that the image of $\bar{f}(x)$ does not "wrap" around the circle) one can choose an unambiguous logarithm. Let us fix (WLOG) $\bar{f}(0) = 1$. Then $f(0)$ can be any integer, $n_0$. Once we choose that integer the branch of the logarithm is fixed in some small open set $[0,\epsilon)$. Now we continue choosing open sets along the interval so that we can choose an unambiguous logarithm and we fix branches successively as we move from one open set to the next along the positive direction. The net result is an unambiguous map $f : [0,1] \to \mathbb{R}$ satisfying (10.32). Now if $\bar{f}(1) = 1$, then $f(1) = n_1$, with $n_1 \in \mathbb{Z}$. The integer $n_1 - n_0$ only depends on $\bar{f}$, and not on the particular choice $n_0$, and is the winding number of the map. Moroever, the winding number is continuous as a function of $\bar{f}$ and hence only depends on the homotopy class.

1. It is not hard to prove that

$$\pi_1(X \times Y, (x_0, y_0)) \cong \pi_1(X, x_0) \times \pi_1(Y, y_0) \tag{10.33}$$

   So the fundamental group of the $n$-dimensional torus is isomorphic to the $n$-dimensional lattice $\mathbb{Z}^n$.

2. If $F : X \to Z$ is a continuous map of topological spaces and takes $x_0 \in X$ to $z_0 \in Z$ then we can define $F_* : \pi_1(X, x_0) \to \pi_1(Z, z_0)$ simply by $F_*[f] := [F \circ f]$. This can be shown to be a group homomorphism. In particular, if $F$ is a homotopy equivalence, then it is a group isomorphism.

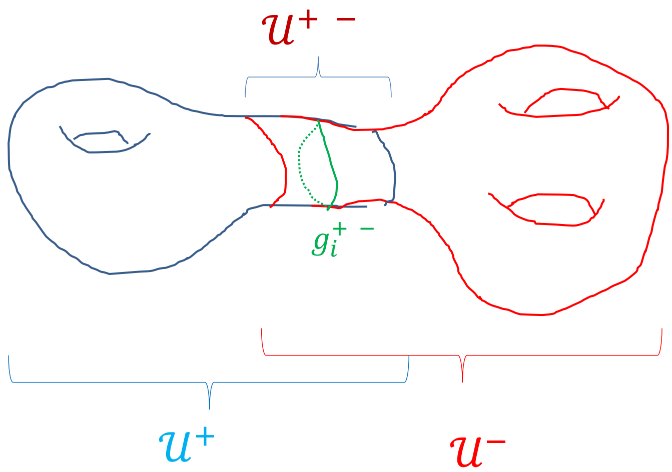



**Figure 19:** Illustrating the Seifert-VanKampen theorem. The green curve has a homotopy class in $\mathcal{U}^{+-}$ that is one of the generators of $\pi_1(\mathcal{U}^{+-})$. Now it must separately be a word $W_i^+$ in the generators of $\pi_1(\mathcal{U}^+)$ and $W_i^-$ in the generators of $\pi_1(\mathcal{U}^-)$ so in $\pi_1(X)$ there must be a relation of the form $W_i^+ = W_i^-$.

3. In algebraic topology books a major result which is proved is the *Seifert-van Kampen theorem*. This is an excellent illustration of defining groups by generators and relations. The theorem can be useful because it allows one to compute $\pi_1(X, x_0)$ by breaking up $X$ into simpler pieces. Specifically, suppose that $X = U^+ \cup U^-$ is a union of two open path-connected subsets and that $U^{+-} := U^+ \cap U^-$ is also path-connected and contains $x_0$. See Figure 19. Now suppose we know presentations of the fundamental groups of the pieces $U^+, U^-, U^{+-}$ in terms of generators and relations:

$$\begin{aligned} \pi_1(U^+, x_0) &\cong \langle g_1^+, \dots, g_{n^+}^+ | R_1^+, \dots, R_{m^+}^+ \rangle \\ \pi_1(U^-, x_0) &\cong \langle g_1^-, \dots, g_{n^-}^- | R_1^-, \dots, R_{m^-}^- \rangle \\ \pi_1(U^{+-}, x_0) &\cong \langle g_1^{+-}, \dots, g_{n^{+-}}^{+-} | R_1^{+-}, \dots, R_{m^{+-}}^{+-} \rangle \end{aligned} \tag{10.34}$$

Then the recipe for computing $\pi_1(X, x_0)$ is this: Denote the injection $\iota^+ : U^{+-} \to U^+$ and $\iota^- : U^{+-} \to U^-$. Then the generators of $\pi_1(U^{+-}, x_0)$ push forward to words in $g_i^+$ or $g_i^-$, respectively:

$$\begin{aligned} \iota_*^+(g_i^{+-}) &:= W_i^+ \qquad i = 1, \dots, n^{+-} \\ \iota_*^-(g_i^{+-}) &:= W_i^- \qquad i = 1, \dots, n^{+-} \end{aligned} \tag{10.35}$$

Finally, we have the presentation:

$$\pi_1(X, x_0) \cong \langle g_1^+, \dots, g_{n^+}^+, g_1^-, \dots, g_{n^-}^- | R_\alpha \rangle \tag{10.36}$$

where the relations $R_\alpha$ include the *old relations*

$$R_1^+, \dots, R_{m^+}^+, R_1^-, \dots, R_{m^-}^- \tag{10.37}$$

and a set of *new relations*:

$$W_1^+(W_1^-)^{-1}, \ldots, W_{n^{+-}}^+(W_{n^{+-}}^-)^{-1} \tag{10.38}$$

It is obvious that these are relations on the generators. What is not obvious is that these are the only ones. Note that in the final presentation the generators $g_i^{+-}$ and the relations $R_i^{+-}$ have dropped out of the description.

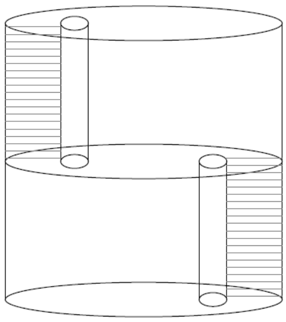


**Figure 20:** The house that Bing built. Taken from M. Freedman and T. Tam Nguyen-Pham, "Non-Separating Immersions Of Spheres and Bing Houses," which describes nice mathematical properties of this house.

---

**Exercise**
Show that if $X = S^n$ with $n > 1$ then $\pi_1(X, x_0)$ is the trivial group.

---

---

**Exercise**
Does the fundamental group depend on a choice of basepoint $x_0$ ?

---

---

**Exercise**
What is the fundamental group of Serin Physics Laboratory?

---

**Exercise** *The house that Bing built*

Show that the house in Figure 20 can be shrink-wrapped with a single balloon so that the complement of the balloon in $\mathbb{R}^3$ is connected and simply connected.

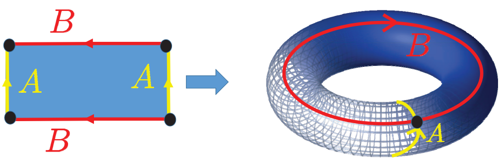



**Figure 21:** Right: Cutting a torus along the A and B cycles the surface falls apart into a rectangle, shown on the left. Conversely, gluing the sides of the rectangle together produces a torus with distinguished closed curves $A, B$.

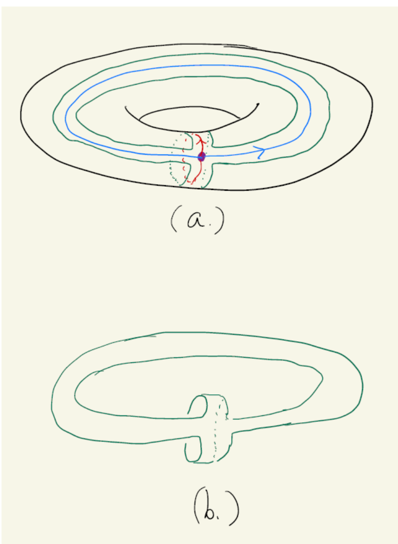



**Figure 22:** Illustrating the Seifert-VanKampen theorem for the torus: The "tubular neighborhood" - the green region - of the cutting curves, shown in (b) is homotopy equivalent to a one-point union of two circles. The latter space has a $\pi_1$ which is a free group on two generators. The boundary of the green region contracts into the remainder of the surface - which can be deformed to a disk. Therefore the group commutator $[a, b] = 1$.

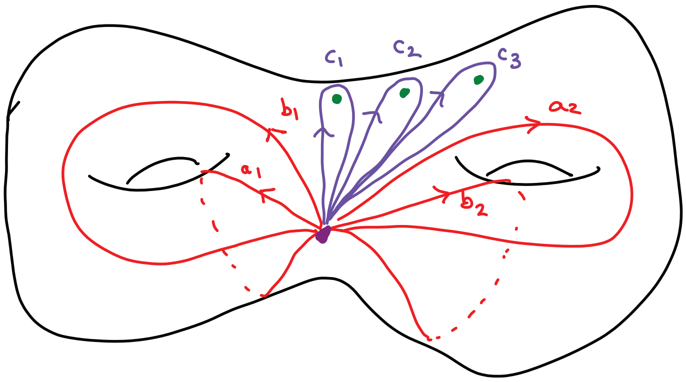



**Figure 23:** A collection of closed paths at $x_0$ which generate the fundamental group of a two-dimensional surface with two handles and three (green) holes.

### 10.1.2 Surface Groups: Compact Two-Dimensional Surfaces

The fundamental groups of two-dimensional surfaces, known as *surface groups* and braid groups turn out to provide a very rich set of examples of groups defined by generators and relations.

The simplest example of a nontrivial surface group is the torus. Let $a, b$ denote the homotopy classes of the cycles $A, B$ shown in Figure 21. One can convince oneself that these generate the fundamental group: Every closed curve based at $x_0$ can be homotoped to a word in $a^{\pm 1}$ and $b^{\pm 1}$. Now, if we cut the torus along the cycles the surface falls apart into a rectangle as shown in Figure 21. The edge of the rectangle represents the class $aba^{-1}b^{-1}$. A slightly different way of thinking about this is described in Figure 22.

**Definition**: In general, in group theory an expression of the form $g_1 g_2 g_1^{-1} g_2^{-1}$ is known as a *group commutator* and is sometimes denoted $[g_1, g_2]$. It should not be confused with the commutator of matrices $[A_1, A_2] = A_1 A_2 - A_2 A_1$.

Returning to the fundamental group of the torus, the group commutator $[a, b]$ it can be contracted inside the rectangle to a point. Therefore, the generators $a, b$ satisfy the relation:

$$aba^{-1}b^{-1} = 1 \tag{10.39}$$

so this means

$$ab = ba \tag{10.40}$$

In fact, this is the only relation and therefore:

$$\pi_1(T^2, x_0) \cong \mathbb{Z} \oplus \mathbb{Z}. \tag{10.41}$$

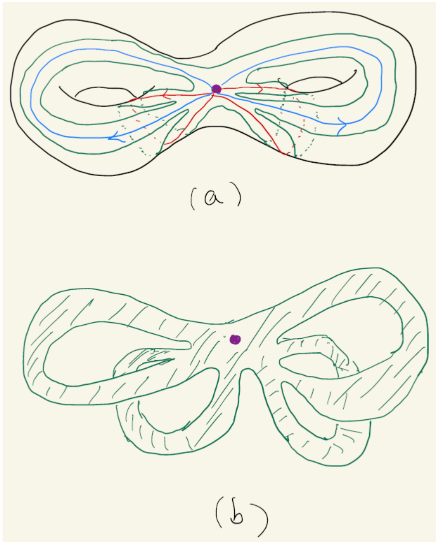



**Figure 24:** Illustrating the Seifert-VanKampen theorem for a genus two surface: The "tubular neighborhood" - the green region - of the cutting curves, shown in (b) is homotopy equivalent to a one-point union of 4 circles. The latter space has a $\pi_1$ which is a free group on four generators which we can call $a_1, b_1, a_2, b_2$. The boundary of the green region is a single circle homotopic to $[a_1, b_1][a_2, b_2]$. But it contracts into the remainder of the surface - which can be deformed to a disk. Therefore by the Seifert-van Kampen theorem the presentation of $\pi_1$ of the genus two surface has a single relation $[a_1, b_1][a_2, b_2] = 1$.

The above ideas generalize nicely. Let us consider the case of a genus 2, or 2-handled, surface shown in Figure 24. The fundamental group can be presented as a group with four generators $a_1, b_1, a_2, b_2$ and one relation:

$$[a_1, b_1][a_2, b_2] = 1 \tag{10.42}$$

Now let us consider a more complicated surface, perhaps with punctures as shown in Figure 25. By cutting along the paths shown there the surface unfolds to a presentation by gluing as in Figure 26:

From these kinds of constructions one can prove [140] that the fundamental group of an orientable surface with $g$ handles and $p$ punctures will be

$$\pi_1(S, x_0) = \langle a_i, b_i, c_s | \prod_{i=1}^{g} [a_i, b_i] \prod_{s=1}^{p} c_s = 1 \rangle \tag{10.43}$$

[140] See, for example, W. Massey, *Introduction to Algebraic Topology*, Springer GTM

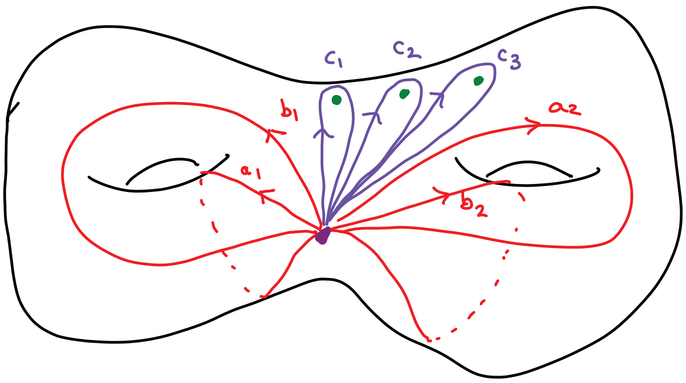



**Figure 25:** A collection of closed paths at $x_0$ which generate the fundamental group of a two-dimensional surface with two handles and three (green) holes.

There is only one relation so this is very close to a free group! In fact, for $p \geq 1$ we can solve for one generator $c_s$ in terms of the rest so the group is just a free group on $2g+p-1$ generators. When there are no punctures the group is not a free group. Groups of the form (10.43) are sometimes called *surface groups*.

As mentioned above, a flat connection amounts to a representation of this group - so one is searching for matrices $A_i, B_i, C_s$ such that

$$\prod_i (A_i B_i A_i^{-1} B_i^{-1}) \prod_s C_s = 1 \tag{10.44}$$

---

**Exercise** *The First Homology Group*

Recall the exercise on commutator subgroups and abelianization. See equation (7.72) above.

Consider a surface group of the type given in (10.43). The abelianization of this group is isomorphic to the *first homology group* $H_1(S)$ where $S$ is the punctured surface. Compute this group. [141]

[141]One can define higher homology groups $H_k(X)$ of a topological space $X$ but these in general are not Abelianiations of the higher homotopy groups $\pi_k(X)$, even though both groups are Abelian. Homology and homotopy groups measure different aspects of the topology of a space.

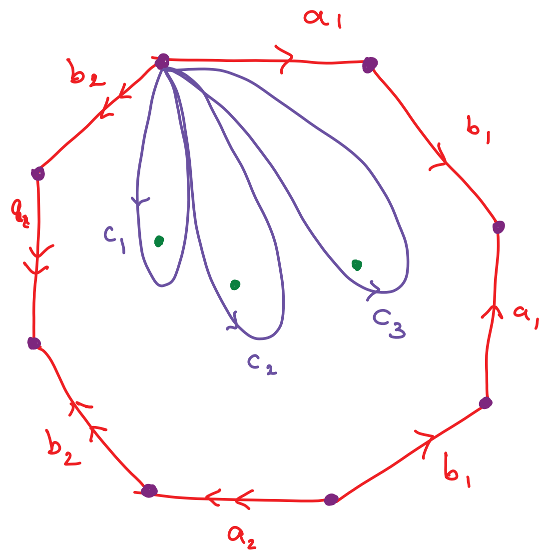



**Figure 26:** When the directed edges are identified according to their labels the above surface reproduces the genus two surface with three punctures. Since the disk is simply connected we derive one relation on the curves shown here.

---

---

**Exercise** *Fundamental group of the Klein bottle*

A very interesting unorientable surface is the Klein bottle. Its fundamental group has two natural presentations in terms of generators and relations. One is

$$\langle a, b | a^2 = b^2 \rangle \tag{10.45}$$

and the other is

$$\langle g_1, g_2 | g_1 g_2 g_1 g_2^{-1} = 1 \rangle \tag{10.46}$$

Show that these two presentations are equivalent.

---

**Exercise**

Use the Seifert-van Kampen theorem to relate the fundamental group of a torus to that of a torus with a disk cut out.

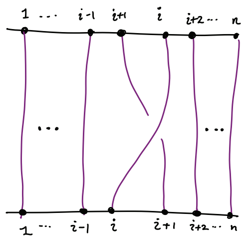



**Figure 27:** Pictorial illustration of the generator $\sigma_i$ of the braid group $B_n$.

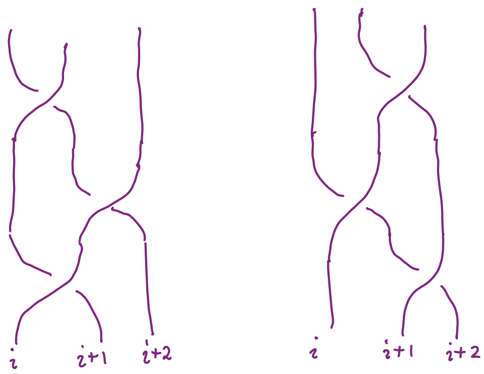



**Figure 28:** Pictorial illustration of the Yang-Baxter relation.

### 10.1.3 Braid Groups And Anyons

Let us modify Figure 2 and Figure 1 to include an under-crossing and overcrossing of the strands. So now we are including more information - the topological configuration of the strands in three dimensions. In an intuitive sense, which we will not make precise here we obtain a group called the $n^{th}$ *braid group*. It is generated by the overcrossing $\tilde{\sigma}_i$ of strings $(i, i+1)$, for $1 \leq i \leq n-1$ and may be pictured as in Figure 27. Note that $\tilde{\sigma}_i^{-1}$ is the undercrossing.

Now one verifies the relations

$$\tilde{\sigma}_i\tilde{\sigma}_j = \tilde{\sigma}_j\tilde{\sigma}_i \qquad\qquad |i-j| \geq 2 \tag{10.47}$$

and

$$\tilde{\sigma}_i\tilde{\sigma}_{i+1}\tilde{\sigma}_i = \tilde{\sigma}_{i+1}\tilde{\sigma}_i\tilde{\sigma}_{i+1} \tag{10.48}$$

where the relation (10.48) is illustrated in Figure 28.

The braid group $\mathcal{B}_n$ may be defined as the group generated by $\tilde{\sigma}_i$ subject to the relations (10.47)(10.48):

$$\mathcal{B}_n := \langle \tilde{\sigma}_1, \dots, \tilde{\sigma}_{n-1} | \tilde{\sigma}_i\tilde{\sigma}_j\tilde{\sigma}_i^{-1}\tilde{\sigma}_j^{-1} = 1, |i-j| \geq 2; \tilde{\sigma}_i\tilde{\sigma}_{i+1}\tilde{\sigma}_i = \tilde{\sigma}_{i+1}\tilde{\sigma}_i\tilde{\sigma}_{i+1} \rangle \tag{10.49}$$

The braid group $\mathcal{B}_n$ may also be defined as the fundamental group of the space of configurations of $n$ unordered points on the disk $D$. We first consider the set:

$$\{(x_1, \dots, x_n) | x_i \in D \qquad x_i \neq x_j \quad i \neq j\} \tag{10.50}$$

Then we observe that there is a group action of $S_n$ on this set. Note that this set is not simply connected: For example if we let $x_1$ loop around $x_2$ holding all other $x_i$ fixed it should be intuitively clear that the loop cannot be deformed to the trivial loop. That is even more clear if you view the looping process as taking place in time on particles in a plane.

Now we consider the space of orbits under this group action:

$$\mathcal{C}_n := \{(x_1, \dots, x_n) | x_i \in D \qquad x_i \neq x_j \quad i \neq j\}/S_n \tag{10.51}$$

There are new nontrivial loops here where, for example, $x_i$ and $x_j$ exchange places, all other $x_k$ staying fixed.

♣Since we must quotient by $S_n$ this needs to be moved to the section on group actions on spaces. ♣

Note that the "only" difference from the presentation of the symmetric group is that we do *not* put any relation like $(\tilde{\sigma}_i)^2 = 1$. Indeed, $\mathcal{B}_n$ is of infinite order because $\tilde{\sigma}_i^n$ keeps getting more and more twisted as $n \to \infty$.

---

**Exercise** *Homomorphisms Between Braid And Symmetric Groups*

a.) Define a homomorphism $\mu : \mathcal{B}_n \to S_n$.

b.) Can you define a homomorphism $s : S_n \to \mathcal{B}_n$ so that $\mu \circ s$ is the identity transformation?

---

**Remarks**

1. In the theory of integrable systems the relation (10.48) is closely related to the "Yang-Baxter relation." It plays a fundamental role in integrable models of 2D statistical mechanics and field theory.

2. One interesting application of permutation groups to physics is in the quantum theory of identical particles. It was a major step in the development of quantum theory when Einstein and Bose realized that a system of $n$ identical kinds of particles (photons, for example, or atomic Nuclei of the same isotope) are in fact *indistinguishable.* [142] In mathematical terms, there is a group action of $S_n$ on a set of $n$ indistinguishable particles leaving the physical system "the same." In quantum mechanics this translates into the statement that the Hilbert space of a system of $n$ indistinguishable particles should be a representation of (a central extension of) $S_n$. There are many different representations of $S_n$ (we have already encountered three different ones). Most of them are higher dimensional. Particles transforming in higher-dimensional representations are said to satisfy "parastatistics." (This idea goes back to B. Green in the 1950's and Messiah and Greenberg in the 1960's.) However, remarkably, in relativistically invariant theories in spacetimes of dimension larger than 3 particles are either bosons or fermions. This is related to the classification of the *projective representations* of $SO(d,1)$, where $d$ is the number of spatial dimensions, for relativistic systems and to representations of $SO(d)$ for nonrelativistic systems. (We will discuss projective representations in section **** below.) Now, when discussing projective representations the fundamental group of $SO(d,1)$ and $SO(d)$ becomes important. In fact $\pi_1(SO(d,1)) \cong \pi_1(SO(d))$ for $d \geq 2$. However, there is a fundamental difference between $d \leq 2$ and $d > 2$. The essential point is that the fundamental group $\pi_1(SO(2)) \cong \mathbb{Z}$ is infinite while $\pi_1(SO(d)) \cong \mathbb{Z}_2$ for $d \geq 3$. A consequence of this, and other principles of physics is that in $2+1$ and $1+1$ dimensions, particles with "anyonic" statistics can exist. [143] Anyons are defined by the property that, if we just consider the wavefunction of two identical such particles, $\Psi(z_1, z_2)$ where $z_1, z_2$ are points in the plane and then we adiabatically switch their positions using the kind of braiding that defines $\tilde{\sigma}$ then

$$\tilde{\sigma} \cdot \Psi(z_2, z_1) = e^{\mathrm{i}\theta} \Psi(z_1, z_2) \tag{10.52}$$

---

[142] See Chapters 24 and 25 in the book *Einstein and the Quantum* by A.D. Stone for a nice historical account of the importance of this discovery in the development of quantum mechanics.

[143] The possible existence of anyons was pointed out by Leinaas and Myrheim in 1977. Another early reference are the papers of Goldin, Melnikof and Sharp. The term "anyon" was invented in F. Wilczek, "Quantum Mechanics of Fractional-Spin Particles". Physical Review Letters 49 (14): 957-959. For comments on the early history of these ideas see "The Ancestry of the 'Anyon' " in Physics Today, August 1990, page 90.

Unlike bosons and fermions where $\theta = 0, \pi \mathrm{mod} 2\pi$, respectively, for "anyons" the phase can be anything - hence the name. There are even physical realizations of this theoretical prediction in the fractional quantum Hall effect. Moreover, quantum wavefunctions should transform in representations of the braid group. The law (10.52) leads to the 1-dimensional representation $\tilde{\sigma} \to e^{i\theta}$ but there can also be more interesting "nonabelian representations." That is, there can be interesting irreducible representations of dimension greater than one, and if wavefunctions transform in such representations there can be *nonabelian statistics*. The particles should be called *nonabelions*. There are some theoretical models of fractional quantum Hall states in which this takes place. [144] Nonabelions are of potentially great importance because of their possible use in quantum computation, an observation first made by A. Kitaev.

Here are some sources for more material about anyons:

♣Need to keep updating and refining this reference list. ♣


1. There are some nice lecture notes by John Preskill, which discuss the potential relation to quantum computation and quantum information theory: http://www.theory.caltech.edu/˜preski
2. A. Stern, "Anyons and the quantum Hall effect: A pedagogical review". Annals of Physics 323: 204; arXiv:0711.4697v1.
3. A. Lerda, *Anyons: Quantum mechanics of particles with fractional statistics* Lect.Notes Phys. M14 (1992) 1-138
4. A. Khare, *Fractional Statistics and Quantum Theory*,
5. G. Dunne, *Self-Dual Chern-Simons Theories.*
6. David Tong, "Lectures on the Quantum Hall Effect," e-Print: arXiv:1606.06687
7. S. Burton, "A Short Guide To Anyons and Modular Functors," 1610.05384
8. J. Pachos, *Introduction To Topological Quantum Computation*
9. K. Beer, "From Categories To Anyons: A Travelogue," 1811.06670
10. E.C. Rowell and Z. Wang, "Mathematics of Topological Quantum Computation," Bulletin American Math. Soc., 55, 183
11. Z. Wang, "Topological Quantum Computation," http://web.math.ucsb.edu/∼zhenghwa/data/course/c
12. L. Kong and Z.-H. Zhang, "An invitation to topological orders and category theory," arXiv:2205.05565


[144]Important work on the compatibility of nonabelions with spin-statistics theorems was done by Jurg Fröhlich. Perhaps the first concrete proposal of a physical system with nonabelions is that of G.W. Moore and N. Read, "Nonabelions in the fractional quantum hall effect." Nuclear Physics B360, 1991. It was inspired, in part, by the work of G. Moore and N. Seiberg, "Polynomial Equations For RCFT" and "Classical and Quantum Conformal Field Theory," Commun.Math.Phys. 123 (1989) 177 which gave the first description of a modular tensor category.

**Exercise** *Relation Of The Braid Group And The Symmetric Group*

Let $\mathcal{B}_n$ be a braid group. Compute the kernel of the natural homomorphism $\phi : \mathcal{B}_n \to S_n$ and show that there is an exact sequence

$$1 \to \mathbb{Z}^{n-1} \to \mathcal{B}_n \to S_n \to 1 \tag{10.53}$$

### 10.1.4 Fundamental Groups Of Three-Dimensional Manifolds

The fundamental groups of surfaces are easy to write down, and surfaces can be classified. The situation is completely different in dimensions three and larger where the problem is much harder. Of course, we could take products like $S^1 \times \Sigma$, where $\Sigma$ is a two-dimensional manifold. But many more possibilities arise.

Let us start by considering a circle $S^1 \subset S^3$. It is probably helpful to think of $S^3$ as $\mathbb{R}^3$ with the boundary at infinity identified to a point. Now the tubular neighborhood of the circle is diffeomorphic to $S^1 \times D^2$ and has boundary $S^1 \times S^1$. Imagine cutting out this tubular neighborhood - what remains? It is some three-manifold with a single boundary which is also a torus $S^1 \times S^1$. If you draw a solid torus you will notice that one of the homotopically nontrivial loops on the torus becomes homotopically trivial.

FIGURE HERE OF SOLID TORUS AND A-CYCLE.

The other cycle is nontrivial. Thus, it must be that $S^3$ can be obtained by gluing together two solid tori (bagels). However, $S^3$ is simply connected, so the cycle which is nontrivial in one bagel must become trivial in the other, and vice versa. This suggests that there are topologically interesting diffeomorphisms of the torus that are used when gluing together the two bagels: Note that if you glue with the identity transformation you get $S^2 \times S^1$, not $S^3$.

Indeed a vast number of new constructions comes of three-manifolds comes from ideas of "surgery." An important point is that there are typically many nontrivial diffeomorphisms of a surface. Let us see this with the case of a torus.

♣This repeats some material from around equation (7.68) ♣

We consider the torus to be the space of orbits $\mathbb{R}^2/\mathbb{Z}x\mathbb{Z}$ with coordinates $(\sigma^1, \sigma^2)$ with $\sigma^i \sim \sigma^i + 1$. Then consider the transformation

$$\begin{aligned} \sigma^1 &\to a\sigma^1 + b\sigma^2 \\ \sigma^2 &\to c\sigma^1 + d\sigma^2 \end{aligned} \tag{10.54}$$

This will be consistent with the identifications $\sigma^i \sim \sigma^i + 1$ iff $a, b, c, d \in \mathbb{Z}$, and will be invertible if

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{Z}) \tag{10.55}$$

Finally, it will preserve orientation if it is in $SL(2, \mathbb{Z})$, that is, if $ad - bc = 1$. We claim this diffeomorphism cannot be smoothly deformed to the identity if the matrix is not the identity. To see this consider its action on the homotopically nontrivial loops on the torus.

So, we can generalize the above construction by considering an arbitrary knot $K \subset S^3$. Again, its tubular neighborhood is topologically just a solid torus, as is the complement - but the two bagels must be glued together by a nontrivial diffeomorphism.

If we consider a higher genus surface $\Sigma$ there is similarly an infinite number of nontrivial self-diffeomorphisms.

The idea of a solid torus can be generalized to the notion of a *handlebody*: A connected three-manifold with an orientable surface $\Sigma$ as boundary. It can also be gotten from a 3-ball by choosing pairs of disks on the boundary and attaching these to the ends of solid tubes. ♣Be more precise ♣

FIGURE

Note that in a genus $g$ handlebody, $g$ of the generators of $\pi_1$ become topologically trivial and the $\pi_1$ is just the free group on $g$ generators.

We can now generalize the above construction by taking two genus $g$ handlebodies, together with a diffeomorphism $\phi \in \mathrm{Diff}(\Sigma)$ and glue the handlebodies together. The resulting three-manifold is said to have a Heegaard decomposition. This means we can find an embedded closed oriented surface $\Sigma \subset Y$ so that $Y$ is the gluing of two handlebodies (bordisms of $\Sigma$ to the emptyset) using a diffeomorphism of $\Sigma$. Such a Heegard decomposition strongly constraints the fundamental group, thanks to the Seifert-van-Kampen theorem: A set of generators is given by a set of generators of the two handlebodies. The fundamental group of the handlebody for $\Sigma$ of genus $g$ is just a free group on $g$ generators. (The "A-cycles" contract to a point, leaving the "B-cycles" with no relation.) After gluing to the other handlebody there will be $g$ relations expressing the contractibility of the new "A-cycles" of the second handlebody.

In fact, any three-manifold has a Heegaard decomposition: By a theorem of Moise, every 3-manifold can be triangulated. Take the 1-skeleton and thicken it to get a handlebody. What is less obvious is that the complement of this handlebody is also a handlebody, but this can be shown. [145]

This shows that every three-manifold admits a Heegaard decomposition. The trouble is, there are a lot of nontrivial diffeomorphisms of a surface to itself, and it can be hard to recognize two equivalent 3-folds constructed from different Heegaard splittings.

Thus, any three-manifold $Y$, $\pi_1(Y, y_0)$ admits a presentation where the number of generators is equal to the number of relations. The existence of such a presentation is not possible for a general finitely generated group. So, at least in three dimensions it is not true that any finitely generated group is the fundamental group of some three-dimensional manifold. Despite a great deal of progress in the past decades on the topology of three-manifolds there is no simple and useful classification of the possible fundamental groups of three-dimensional manifolds - in strong contrast to the two-dimensional situation.

[145] To show this it is necessary to introduce the dual triangulation: The vertices and barycenters of 3-simplices are exchanged and the dual edges go from a barycenter of a 3-simplex to the barycenters of the faces. One then shows that the complement of the handlebody of the 1-skeleton is homeomorphic to a neighborhood of the dual 1-skeleton, and is therefore a handlebody. For a careful discussion see https://web.math.ucsb.edu/∼mgscharl/papers/MorningsideNotes.pdf.

### 10.1.5 Fundamental Groups Of Four-Dimensional Manifolds

A basic fact is the following theorem of Markov:

**Theorem** Any finitely generated group $G$ is the fundamental group of some four-manifold.

*Proof*: Suppose the group $G$ has presentation:

$$G \cong \langle g_1, \dots, g_n | R_1, \dots, R_m \rangle \tag{10.56}$$

We aim to produce a four-manifold $M_4$ with fundamental group $G$. First, consider the free group on one generator $\langle g \rangle \cong \mathbb{Z}$. A good manifold that has this as a fundamental group is $X_4 = S^1 \times S^3$. Now let us consider $\tilde{M}_4 := X_4 \# \cdots \# X_4$. Then

$$\pi_1(\tilde{M}_4) \cong \langle g_1, \dots, g_n \rangle \tag{10.57}$$

is the free group with generators $g_i$ corresponding to the simple loops around the $S^1$ factor in each summand (extended to some common basepoint). Now, each relation $R_\alpha$ is a word in the $g_i$ and thus can be represented by some closed based loop $\ell_\alpha \subset \tilde{M}_4$. We can take the $\ell_\alpha$ to be nonintersecting, by simple codimension arguments. Now, take a tubular neighborhood $N(\ell_\alpha)$ of $\ell_\alpha$. By our discussion above of the local picture of submanifolds it is diffeomorphic to $N(\ell_\alpha) \cong S^1 \times D^3$, where $D^3$ is the 3-dimensional ball. The boundary is thus $\partial N(\ell_\alpha) \cong S^1 \times S^2$. This is also the boundary of $D^2 \times S^2$. So, glue in a copy of $D^2 \times S^2$ along the boundary of $N(\ell_\alpha)$. This procedure is known as *surgery.* Now the loop $S^1$ in $S^1 \times D^3$ (which was representing the word $R_\alpha$) becomes contractible! Thus it is a relation on the generators $g_i$ in the new manifold. We can choose the tubular neighborhoods around the different loops $\ell_\alpha$ to be nonintersecting, and hence we can perform surgeries on each of these loops without interference. If we do this for all the loops we produce our manifold $M_4$. By the Seifert-van Kampen theorem it follows that the fundamental group of $M_4$ is exactly $G$. ♠

♣Need to explain why it is a free group and not a free abelian group! ♣

Since finitely presented groups cannot be classified it follows that four-manifolds cannot be classified, even up to homotopy type.

## 11. Some Representation Theory

One of the main motivations from physics for studying representation theory stems from *Wigner's theorem* discussed below. The basic upshot of Wigner's theorem is that, in quantum mechanics, if $G$ is a group of symmetries of a physical system then the Hilbert space of the theory will be a representation space of $G$ and in fact will define a unitary representation: [146] For every symmetry operation $g \in G$ there is a unitary operator $U(g)$ acting on the Hilbert space $\mathcal{H}$ so that

$$U(g_1)U(g_2) = U(g_1 g_2) \tag{11.1}$$

[146] We will need to amend this in two ways to be completely accurate: First, classical symmetries in general are represented projectively on a quantum Hilbert space. Second, one must allow symmetries to be represented by both unitary and anti-unitary operators in general.

The use of representation can be very powerful and have far-reaching consequences. A few examples:

1. The use of representation theory greatly aids in the diagonalization of physical observables, such as Hamiltonians.
2. Quantum states can be classified according to their symmetry types. This has important applications to selection rules governing what kind of transition amplitudes can be nonzero.
3. Conservation laws are associated with symmetry operations.
4. The very formulation of Lagrangians and actions makes heavy use of representation theory. For example, in relativistically invariant field theory the fields form a representation of the Poincaré group (induced from the action on spacetime) and one wishes to make a Lorentz invariant density when forming a Lagrangian.

### 11.1 Motivation: Wigner's Theorem In Quantum Mechanics

The following material, while very important, assumes knowledge of some of the linear algebra from Chapter 2 and some familiarity with quantum mechanics. For further details see Chapter **** below. The reader should also consult Section 2 of [147]

Let us review, very briefly the most essential points of quantum mechanics:

The basic idea in quantum mechanics is that physical measurements are based on a pairing of a physical state and an observable to give a probability measure on the real line. If $\mathcal{S}$ is the set of physical states of a physical system and $Obs$ is the set of physical observables, then to a physical state $\rho \in \mathcal{S}$ and a physical observable $O \in Obs$ we obtain a probability distribution $\wp_{\rho,O}$ on the real line. The value $\wp_{\rho,O}(E)$ where $E \subset \mathbb{R}$ is a measureable set (such as $E = (a, b)$) is the probability that a perfect experimental observation of $O$ in the state $\rho$ will produce a value in $E$.

This general idea is given substance in the Dirac-von Neumann axioms of quantum mechanics. Those axioms posit that to a physical system we associate a complex Hilbert space $\mathcal{H}$ such that

1. Physical states are identified with traceclass positive operators on $\mathcal{H}$, denoted $\rho$ such that $\rho$ has trace one. States are often called *density matrices*. We denote the space of physical states by $\mathcal{S}$. [148]

[147] http://www.physics.rutgers.edu/∼gmoore/695Fall2013/CHAPTER1-QUANTUMSYMMETRY-OCT5.pdf

[148] As explained in the Linear Algebra User's Manual, positive operators can be defined as operators $A$ such that $(\psi, A\psi) \geq 0$ for every $\psi \in \mathcal{H}$. Such operators are always self-adjoint. Indeed, any operator $A$ such that $(\psi, A\psi) \in \mathbb{R}$ for all $\psi \in \mathcal{H}$ is self-adjoint. To see this note that $(\psi, A\psi)^* = (\psi, A\psi)$ and hence $(\psi, A\psi) = (\psi, A^\dagger\psi)$. Now apply this equation to $\psi_1 + z\psi_2$ for $z = 1$ and $z = \sqrt{-1}$ and add the resulting equations to deduce that $(\psi_1, A\psi_2) = (\psi_1, A^\dagger\psi_2)$ for all pairs $\psi_1, \psi_2 \in \mathcal{H}$. Choose an ON basis for $\mathcal{H}$ to deduce that $A = A^\dagger$.

2. Physical observables are identified with self-adjoint operators on $\mathcal{H}$. We denote the set of (bounded) self-adjoint operators on $\mathcal{H}$ by $B(\mathcal{H})$. In general *Obs* can also include unbounded self-adjoint operators, but here there are technical subtleties since the domain is typically not all of $\mathcal{H}$, and the choice of domain can have an important impact on the nature of the operator.

3. The *Born rule* states that when measuring the observable $O$ in a state $\rho$ in a perfect experiment the probability of measuring value $e \in E \subset \mathbb{R}$, where $E$ is a Borel-measurable subset of $\mathbb{R}$, is
$$\wp_{\rho,O}(E) = \mathrm{Tr} P_O(E)\rho. \tag{11.2}$$
where $P_O$ is the projection-valued-measure associated to the self-adjoint operator $O$ by the spectral theorem. For example, if $\mathcal{O}$ has a complete discrete spectrum $\{\lambda_i\}$ of eigenvalues so that
$$\mathcal{O} = \sum_i \lambda_i P(\lambda_i) \tag{11.3}$$
where $P(\lambda_i)$ is the projection operator onto the eigenspace with eigenvalue $\lambda_i$ then
$$P_O(E) = \sum_{\lambda_i \in E} P(\lambda_i) \tag{11.4}$$
When the spectrum of $O$ is more complicated, e.g. if there is a continuous spectrum, one can still define $P_O(E)$, but the story is more involved. See Chapter 2, the Linear Algebra User's Manual.

4. There are further axioms regarding time-development, and so on, but the above is all we need for the present discussion.

Given this setup up the natural notion of a general "symmetry" in quantum mechanics is the following:

**Definition** An *automorphism* of a quantum system is a pair of bijective maps $s_1 : \mathcal{S} \to \mathcal{S}$ and $s_2 : \mathcal{O} \to \mathcal{O}$ and where $s_2$ is real linear on $\mathcal{O}$ such that $(s_1, s_2)$ preserves probability measures:
$$\wp_{s_1(\rho), s_2(O)} = \wp_{\rho,O} \tag{11.5}$$
This set of mappings forms a group which we will call the group of *quantum automorphisms.*

♣Need to state some appropriate continuity properties. ♣

While this is conceptually straightforward, it is an unwieldy notion of symmetry. We will now simplify it considerably, ending up with the crucial theorem known as *Wigner's theorem.*

We begin by noting that the space of density matrices is a convex set. The convexity means that if $\rho_1, \rho_2$ are density matrices then for all $0 \le t \le 1$
$$t\rho_1 + (1-t)\rho_2 \tag{11.6}$$
is a density matrix. Given a convex set on defines an *extremal point* to be a point in the set which cannot be written in the above form with $0 < t < 1$. By definition, the *pure*

*states* are the extremal points of $\mathcal{S}$. The pure states are just the dimension one projection operators.

Pure states are often referred to in the physics literature as "rays in Hilbert space" for the following reason:

If $\psi \in \mathcal{H}$ is a nonzero vector then it determines a line

$$\ell_\psi := \{z\psi | z \in \mathbb{C}\} := \psi\mathbb{C} \tag{11.7}$$

Note that the line does not depend on the normalization or phase of $\psi$, that is, $\ell_\psi = \ell_{z\psi}$ for any nonzero complex number $z$. Put differently, the space of such lines is projective Hilbert space

$$\mathbb{P}\mathcal{H} := (\mathcal{H} - \{0\})/\mathbb{C}^* \tag{11.8}$$

Equivalently, this can be identified with the space of rank one projection operators. Indeed, given any line $\ell \subset \mathcal{H}$ we can write, in Dirac's bra-ket notation: [149]

$$P_\ell = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle} \tag{11.9}$$

where $\psi$ is any nonzero vector in the line $\ell$.

It is possible to argue (see the above reference) that such a symmetry maps pure states to pure states. Moreover, using the trick that a state is a self-adjoint operator, hence also an observable, one can show that the group of quantum automorphisms is ismorphic to the group of 1-1 maps on the pure states that preserves overlaps of pure states. For two pure states, thought of as rank one projection operators, the overlap function is

$$\mathfrak{o}(P_1, P_2) = \mathrm{Tr}(P_1 P_2) = \frac{|\langle\psi_1|\psi_2\rangle|^2}{\parallel \psi_1 \parallel^2 \parallel \psi_2 \parallel^2} \tag{11.10}$$

The the above claim is that the group $\mathrm{Aut}(QM)$ of quantum automorphisms is the group of bijective (continuous) maps $s : \mathbb{P}\mathcal{H} \to \mathbb{P}\mathcal{H}$ such that

$$\mathfrak{o}(s(P_1), s(P_2)) = \mathfrak{o}(P_1, P_2) \tag{11.11}$$

**Example**: Let us consider this case of a single Qbit, namely $\mathcal{H} = \mathbb{C}^2$. First we write the most general general density matrix. Any $2 \times 2$ Hermitian matrix is of the form $a + \vec{b} \cdot \vec{\sigma}$ where $\vec{\sigma}$ is the vector of "Pauli matrices":

$$\begin{aligned} \sigma^1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma^2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ \sigma^3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{11.12}$$

[149] We generally denote inner products in Hilbert space by $(x_1, x_2) \in \mathbb{C}$ where $x_1, x_2 \in \mathcal{H}$. Our convention is that it is complex-linear in the second argument. However, we sometimes write equations in Dirac's bra-ket notation because it is very popular. In this case, identify $x$ with $|x\rangle$. Using the Hermitian structure there is a unique anti-linear isomorphism of $\mathcal{H}$ with $\mathcal{H}^*$ which we denote $x \mapsto \langle x|$. Sometimes we denote vectors by Greek letters $\psi, \chi, \dots$, and scalars by Latin letters $z, w, \dots$. But sometimes we denote vectors by Latin letters, $x, w, \dots$ and scalars by Greek letters, $\alpha, \beta, \dots$.

$a \in \mathbb{R}$ and $\vec{b} \in \mathbb{R}^3$. Now a density matrix $\rho$ must have trace one, and therefore $a = \frac{1}{2}$. Then the eigenvalues are $\frac{1}{2} \pm |\vec{b}|$ so positivity means it must have the form

$$\rho = \frac{1}{2}(1 + \vec{x} \cdot \vec{\sigma}) \tag{11.13}$$

where $\vec{x} \in \mathbb{R}^3$ with $\vec{x}^2 \le 1$.

The extremal states, corresponding to the rank one projection operators are therefore of the form

$$P_{\vec{n}} = \frac{1}{2}(1 + \vec{n} \cdot \vec{\sigma}) \tag{11.14}$$

where $\vec{n}$ is a unit vector. This gives the explicit identification of the pure states with elements of $S^2$. Moreover, we can easily compute:

$$\mathrm{Tr} P_{\vec{n}_1} P_{\vec{n}_2} = \frac{1}{2}(1 + \vec{n}_1 \cdot \vec{n}_2) \tag{11.15}$$

Note that $\vec{n}_1 \cdot \vec{n}_2 = \cos(\theta)$ where $\theta$ is the angle between $\vec{n}_1$ and $\vec{n}_2$. Thus

$$\mathfrak{o}(P_1, P_2) = \cos^2(\frac{1}{2} d(P_1, P_2)) \tag{11.16}$$

where $d(P_1, P_2) = d(\hat{n}_1, \hat{n}_2)$ is the distance between two unit vectors in the unit radius sphere. Therefore the group of quantum automorphisms for a single Qbit is the same as the group of isometries of $S^2$ with its round metric. This group is well known to be the full orthogonal group $O(3)$.

In general for $\mathcal{H} = \mathbb{C}^{N+1}$ the space of pure states is $\mathbb{CP}^N$. This space has a natural homogeneous metric known as the Fubini-Study metric. When it is suitably normalized the overlap function $\mathfrak{o}$ for two projectors $P_1, P_2$ is nicely related to the Fubini-Study distance $d(P_1, P_2)$ between those projectors by

$$\mathfrak{o}(P_1, P_2) = \left(\cos \frac{d(P_1, P_2)}{2}\right)^2 \tag{11.17}$$

Once again, the group of quantum automorphisms is again the group of isometries of the Fubini-Study metric.

Now, it is hard to work with the space of rank-one projection operators since $\mathbb{CP}^N$ is not a linear space: The sum of two one-dimensional projection operators is typically not even proportional to a projection operator. (It is true that the sum of two orthogonal projection operators is a projection operator - but it would not be rank-one.) It would be much nicer to work with linear operators acting on Hilbert space. A fundamental theorem of quantum mechanics known as *Wigner's theorem* states that $\mathrm{Aut}(QM)$ is indeed a quotient of a certain group of operators $\mathrm{Aut}(\mathcal{H})$ that act on a Hilbert space.

A unitary operator on $\mathcal{H}$ is a $\mathbb{C}$-linear operator $u : \mathcal{H} \to \mathcal{H}$ that preserves norms:

$$\parallel u\psi \parallel = \parallel \psi \parallel \tag{11.18}$$

There is a homomorphism $\pi : U(\mathcal{H}) \to Aut(QM)$ where a unitary operator $u$ on $\mathcal{H}$ is mapped to the transformation of rank one projection operators:

$$\pi(u) : P \mapsto u P u^{\dagger} \tag{11.19}$$

for $u \in U(\mathcal{H})$. Note that since $u^\dagger = u^{-1}$ it follows that the overlap function is preserved:

$$\mathfrak{o}(\pi(u)(P_1), \pi(u)(P_2)) = \mathfrak{o}(P_1, P_2) \tag{11.20}$$

If we consider this homomorphism for the case of one Qbit we get $\pi : U(2) \to O(3)$, but it is not hard to see that we only map into the subgroup $SO(3) \subset O(3)$. It is natural to ask if there are operators on $\mathcal{H}$ that map to the reflection-rotation operators in $O(3)$. Indeed there are, but we need to include anti-unitary operators:

An anti-unitary operator on $\mathcal{H}$ is an $\mathbb{C}$-anti-linear operator [150] $a : \mathcal{H} \to \mathcal{H}$ that preserves norms:

$$\parallel a\psi \parallel = \parallel \psi \parallel \tag{11.21}$$

Now, for an anti-unitary operator we can still define

$$\pi(a) : P \mapsto aPa^\dagger \tag{11.22}$$

and (see the footnote above for the definition of the adjoint of an anti-unitary operator) it is still true that $a^{-1} = a^\dagger$ so these transformations of pure states also preserve the overlap function:

$$\mathfrak{o}(\pi(a)(P_1), \pi(a)(P_2)) = \mathfrak{o}(P_1, P_2) \tag{11.23}$$

Now we note that the union of unitary and anti-unitary operators is a group, denoted $\mathrm{Aut}(\mathcal{H})$. The composition of unitary operators is clearly unitary. So $U(\mathcal{H})$ is a subgroup of $\mathrm{Aut}(\mathcal{H})$. The composition of a unitary and antiunitary operator is antiunitary, and the composition of antiunitaries is unitary so we have an exact sequence

$$1 \to U(\mathcal{H}) \to \mathrm{Aut}(\mathcal{H}) \xrightarrow{\phi} \mathbb{Z}_2 \to 1 \tag{11.24}$$

where

$$\phi(g) = \begin{cases} +1 & g \text{ unitary} \\ -1 & g \text{ anti} - \text{unitary} \end{cases} \tag{11.25}$$

Altogether we have described a homomorphism $\pi : \mathrm{Aut}(\mathcal{H}) \to \mathrm{Aut}(QM)$. Now, Wigner's theorem says that: *Every quantum automorphism is of the form $\pi(u)$ or $\pi(a)$.* So the homomorphism $\pi : \mathrm{Aut}(\mathcal{H}) \to \mathrm{Aut}(QM)$ is surjective. The kernel of $\pi$ can be thought of as possible c-number phases which can multiply the operator on Hilbert space representing a symmetry operation:

$$1 \ \to \ U(1) \ \to \ \mathrm{Aut}(\mathcal{H}) \ \xrightarrow{\pi} \mathrm{Aut}(QM) \ \to 1 \tag{11.26}$$

Here the $U(1)$ is the group of phases acting on quantum states: $\psi \to z\psi$ for $z \in U(1)$.

[150] See Chapter 2, the Linear Algebra User's Manual, for more on this. Briefly this means $a(\psi_1 + \psi_2) = a(\psi_1) + a(\psi_2)$ for any two vectors but $a(z\psi) = z^* a(\psi)$ for any scalar $z \in \mathbb{C}$. Note that we must then define the adjoint by $\langle a^\dagger \psi_1, \psi_2 \rangle := \langle \psi_1, a\psi_2 \rangle^*$ in a convention where the sesquilinear form is $\mathbb{C}$-antilinear in the first argument and linear in the second.

In general a given physical system we will not have the full group $\mathrm{Aut}(QM)$ acting as a group of symmetries because we typically think of a "symmetry" as something that preserves the dynamical laws of evolution. For example, the time-dynamics of a nonrelativistic quantum system is governed by the Schrödinger equation:

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{11.27}$$

where $H$ is a self-adjoint operator, the Hamiltonian. For time-independent Hamiltonians the unitary time evolution is governed by

$$U(t) = \exp[-\mathrm{i}\frac{t}{\hbar}H] \tag{11.28}$$

this will induce a flow on the space of physical states $\rho \mapsto U(t)\rho U(t)^{-1}$. (See section **** below for an important subtlety related to this when symmetries involve the reversal of time orientation.) Only a subgroup of $\mathrm{Aut}(QM)$ will commute with the resultant flows on the space of states. Thus our symmetry group $G$ will come with a homomorphism to the group of quantum automorphisms:

$$\begin{array}{ccccccccc} & & & & & & G & & \\ & & & & & & \downarrow{\scriptstyle\varphi} & & \\ 1 & \longrightarrow & U(1) & \xrightarrow{\iota} & \mathrm{Aut}(\mathcal{H}) & \xrightarrow{\pi} & \mathrm{Aut}(QM) & \longrightarrow & 1 \end{array} \tag{11.29}$$

Now for each $g \in G$ we have $\varphi(g) \in \mathrm{Aut}(QM)$ and by Wigner's theorem we may choose an associated unitary, or antiunitary, operator $U(g)$ acting on a Hilbert space such that $\pi(U(g)) = \varphi(g)$. Of course, $U(g)$ is not unique. We also have $\pi(zU(g)) = \varphi(g)$ for any $z \in U(1)$. This ambiguity will be important in what follows. Now $\varphi$ is a homomorphism into $\mathrm{Aut}(QM)$ but this just means that if $g_1, g_2 \in G$ then

$$\pi(U(g_1)U(g_2)) = \varphi(g_1)\varphi(g_2) = \varphi(g_1g_2) = \pi(U(g_1g_2)) \tag{11.30}$$

Since $\ker(\pi)$ is the group of phases acting as scalars on $\mathcal{H}$ we can only conclude that

$$U(g_1)U(g_2) = c(g_1, g_2)U(g_1g_2) \tag{11.31}$$

for some phase factor $c(g_1, g_2)$. These phase factors define a map $c : G \times G \to U(1)$. In general this map is not a homomorphism. Since we made a choice $U(g)$ for each $\varphi(g)$ one might trying changing the choice. That is we try another operator $\tilde{U}(g)$ so that $\pi(\tilde{U}(g)) = \pi(U(g)) = \varphi(g)$. Two such choices will be related by $\tilde{U}(g) = b(g)U(g)$ where $b$ is a map from $G$ to $U(1)$. In general $b : G \to U(1)$ will not be a homomorphism. We compute

$$\tilde{U}(g_1)\tilde{U}(g_2) = \tilde{c}(g_1, g_2)\tilde{U}(g_1g_2) \tag{11.32}$$

with

$$\tilde{c}(g_1, g_2) = \begin{cases} \frac{b(g_1)b(g_2)}{b(g_1g_2)}c(g_1, g_2) & \phi(U(g_1)) = 1 \\ \frac{b(g_1)b(g_2)^*}{b(g_1g_2)}c(g_1, g_2) & \phi(U(g_1)) = -1 \end{cases} \tag{11.33}$$

It can be shown that in general such redefinitions cannot set the function $c$ to be one. When this happens the map $U : G \to \mathrm{Aut}(\mathcal{H})$ is said to a be *projective representation of* $G$. For more on all the above - see the chapter below on extensions.

**Remark**. It is worth noting the we can associate a representation of some group to the above situation: We consider the set $\tilde{G} = U(1) \times G$ with the multiplication law:

$$(z_1, g_1) \cdot (z_2, g_2) := (z_1 z_2 c(g_1, g_2), g_1 g_2) \tag{11.34}$$

One thing we can note is that finite dimensional matrices are always associative! [151] So for all $g_1, g_2, g_3 \in G$

$$(U(g_1)U(g_2))U(g_3) = U(g_1)(U(g_2)U(g_3)) \tag{11.35}$$

and hence (assuming all three operators are unitary)

$$c(g_1, g_2)c(g_1 g_2, g_3) = \begin{cases} c(g_2, g_3)c(g_1, g_2 g_3) & \phi(U(g_1)) = 1 \\ c(g_2, g_3)^* c(g_1, g_2 g_3) & \phi(U(g_1)) = -1 \end{cases} \tag{11.36}$$

Using this equation one can show that the multiplication law (11.34) is indeed a group law (assume $\phi(U(g)) = 1$ for all $g$). The group $\tilde{G}$ is called a *central extension of* $G$. What we have described is the so-called pullback construction of representation theory. See the chapter on extensions below. The group $\tilde{G}$ is represented on $\mathcal{H}$ via

$$T((z, g)) := zU(g) \tag{11.37}$$

Thus, when studying symmetries in quantum mechanics, all roads lead to representation theory.

With the above as motivation, we now turn to a systematic description of group representation theory.

---

**Exercise** *Defining New Groups From Cocycles*

a.) Show that the multiplication law (11.34) on $\tilde{G} = U(1) \times G$ defines a group law if $c : G \times G \to U(1)$ satisfies the first line of (11.36).

b.) Suppose there is a homomorphism $\phi : G \to \mathbb{Z}_2$ and we define

$$z^{\phi(g)} := \begin{cases} z & \phi(g) = 1 \\ z^* & \phi(g) = -1 \end{cases} \tag{11.38}$$

Show that the multiplication law on $U(1) \times G$ given by

$$(z_1, g_1) \cdot (z_2, g_2) := (z_1 z_2^{\phi(g_1)} c(g_1, g_2), g_1 g_2) \tag{11.39}$$

defines a group law if $c : G \times G \to U(1)$ satisfies the second line of (11.36).

For help: See Section 15 below.

---

[151] And the same holds for linear operators on infinite-dimensional Hilbert spaces provided the domains are such that the composition of the three operators is well-defined. Since our operators $U(g)$ are bounded operators there will not be any problem with the domain of definition.

## 11.2 Some Basic Definitions

Let $V$ be a vector space over a field $\kappa$ and recall that $GL(V)$ denotes the group of all invertible linear transformations $V \to V$. It is also denoted as $\mathrm{Aut}(V)$ since it is the group of linear automorphisms of $V$ with itself.

**Definition 11.2.1** A *representation* of $G$ is a group homomorphism from $G$ to a group of the form $GL(V) = \mathrm{Aut}(V)$ where $V$ is a vector space over a field $\kappa$:

$$\begin{aligned} T : g &\mapsto T(g) \\ G &\to GL(V) \end{aligned} \tag{11.40}$$

$V$ is sometimes called the *representation space* or the *carrier space*. We will abbreviate the action of $T(g)$ on a vector $v \in V$ from $T(g)(v)$ to $T(g)v$ for readability.

Put differently, in terms of group actions, a representation of $G$ is a $G$ action on a vector space that respects the linear structure:

$$g \cdot (\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 g \cdot v_1 + \alpha_2 g \cdot v_2 \tag{11.41}$$

where $v_1, v_2 \in V$ are any vectors and $\alpha_1, \alpha_2 \in \kappa$ are any scalars.

If the vector space has an ordered basis then we get a matrix representation. For example, if $V$ is finite-dimensional then we can choose an ordered basis $\{v_1, \dots, v_n\}$ and the corresponding matrix representation $g \mapsto M(g) \in GL(n, \kappa)$ is defined by:

$$T(g)v_i = \sum_j M(g)_{ji} v_j \tag{11.42}$$

One easily checks that $T(g_1) \circ T(g_2) = T(g_1 g_2)$ implies $M(g_1)M(g_2) = M(g_1 g_2)$.

1. As a simple example, take $V = \kappa$, the standard one-dimensional vector space over $\kappa$ and let $T(g) = 1_V$ for all $g \in G$. This representation is known as the trivial representation.

2. We will often abbreviate "representation" to "rep." Moreover we sometimes refer to "the rep $(V, T)$" or simply by $V$, when we wish to stress the representation space. Or we might refer to "the rep $T$," when the rest of the data is understood.

♣Should it be $(V, T)$ or $(T, V)$? Depends on which is logically prior. Some would say $V$ is defined by the codomain of $T$. Others would say you first choose a $V$ and then define a $T$. Need to be consistent... ♣

3. The "dimension of the representation" is by definition the dimension $\dim V$ of the vector space $V$. This number can be finite or infinite. Sometimes representations are simply denoted by the dimension of the carrier space. This can be dangerous. For example, we will see below that there are many different one-dimensional representations of Abelian groups. The practice is more commonly used for discussing representations of groups like $SU(2)$, but also here one should be cautious. For example, there are $p(n)$ inequivalent $n$-dimensional representations of $SU(2)$, although, as we will see, there is a unique irreducible $n$-dimensional representation of $SU(2)$ - up to isomorphism. See below.

4. The notion of representation can be generalized by replacing $GL(V)$ by $GL(R)$ where $R$ is a ring. Then one often speaks of a *module for $G$*.

**Definition 11.2.2**. Let $(V_1, T_1)$ and $(V_2, T_2)$ be two representations of a group $G$. An *intertwiner* between these representations is a linear transformation $A : V_1 \to V_2$ such that, for all $g \in G$ the diagram

$$\begin{array}{ccc} V_1 & \xrightarrow{A} & V_2 \\ {\scriptstyle T_1(g)}\downarrow & & \downarrow{\scriptstyle T_2(g)} \\ V_1 & \xrightarrow{A} & V_2 \end{array} \tag{11.43}$$

commutes. Equivalently,

$$T_2(g)A = AT_1(g) \tag{11.44}$$

for all $g \in G$. Put differently: $A$ is a morphism of $G$ actions, and put yet another way: $A$ is an equivariant linear map of $G$ spaces. We denote the vector space of all intertwiners by $\mathrm{Hom}_G(V_1, V_2)$.

Note that if an intertwiner is invertible then $T^{-1}$ is also an intertwiner. So we have:

**Definition 11.2.3**. Two representations $(T_1, V_1)$ and $(T_2, V_2)$ are *equivalent* $(T_1, V_1) \cong (T_2, V_2)$ (usually abbreviated to $T_1 \cong T_2$ or $V_1 \cong V_2$) if there is an intertwiner $A : V_1 \to V_2$ which is an isomorphism. That is,

$$T_2(g) = AT_1(g)A^{-1} \tag{11.45}$$

for all $g \in G$.

**Examples**:

1. The general linear group $GL(n, \kappa)$ with $\kappa = \mathbb{R}, \mathbb{C}$ always has a family of one-dimensional real representations, labeled by $\mu \in \mathbb{C}$ given by

$$T(g) := |\det g|^{\mu} \tag{11.46}$$

This is a representation because:

$$T(g_1 g_2) = |\mathrm{det} g_1 g_2|^{\mu} = |\mathrm{det} g_1|^{\mu} |\mathrm{det} g_2|^{\mu} = T(g_1)T(g_2) \tag{11.47}$$

Note that for different $\mu$ these are inequivalent representations.

2. We saw that the connected component of the Lorentz group in $1+1$ dimensions is isomorphic to $\mathbb{R}$, $SO_0(1,1) \cong \mathbb{R}$ with boosts $B(\theta_1)B(\theta_2) = B(\theta_1 + \theta_2)$ with $\theta_i \in \mathbb{R}$. The *"spin-s" representation* is

$$T_s(B(\theta)) = e^{s\theta} \tag{11.48}$$

For different values of $s$ these one-dimensional representations are inequivalent.

Familiar notions of linear algebra generalize to representations:

1. *The direct sum $\oplus$ of representations.* The direct sum of $(T_1, V_1)$ and $(T_2, V_2)$ is the rep $(T_1 \oplus T_2, V_1 \oplus V_2)$ where the representation space is $V_1 \oplus V_2$ *and* the operators are:

$$((T_1 \oplus T_2)(g)) \left(v_1 \oplus v_2\right) := (T_1(g))(v_1) \oplus (T_2(g))(v_2) \tag{11.49}$$

If $\{v_1, \dots, v_n\}$ is an ordered basis for $V_1$ and $\{w_1, \dots, w_m\}$ is an ordered basis for $V_2$ then $\{(v_1, 0), \dots, (v_n, 0), (0, w_1), \dots, (0, w_m)\}$ is an ordered basis for $V_1 \oplus V_2$ and relative to this ordered basis the matrix representation is block diagonal:

$$M_{T_1 \oplus T_2}(g) = \begin{pmatrix} M_{T_1}(g) & 0 \\ 0 & M_{T_2}(g) \end{pmatrix} \tag{11.50}$$

2. Similarly, for the tensor product, the carrier space is $V_1 \otimes V_2$. See LAUM for a proper definition of the tensor product. Briefly, the tensor product will be a vector space where elements are linear combinations of symbols $v \otimes w$, with $v \in V$ and $w \in W$ subject to the rules that: [152]

$$(\alpha_1 v_1 + \alpha_2 v_2) \otimes w = \alpha_1 v_1 \otimes w + \alpha_2 v_2 \otimes w \tag{11.51}$$

$$v \otimes (\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 v \otimes w_1 + \alpha_2 v \otimes w_2 \tag{11.52}$$

In practical terms, for finite-dimensional vector spaces we can say that if $\{v_1, \dots, v_n\}$ is a basis for $V_1$ and $\{w_1, \dots, w_m\}$ is a basis for $V_2$ then the set of vectors of the form $v_i \otimes w_a$ form a basis for $V_1 \otimes V_2$. We take linear combinations of these vectors, subject to the rules (11.51) and (11.52). We can also take tensor products of operators. We set:

$$((T_1 \otimes T_2)(g)) \left(v \otimes w\right) := (T_1(g)v) \otimes (T_2(g)w) \tag{11.53}$$

for all $v \in V_1$ and $w \in V_2$ and then extend by $\kappa$-linearity.

If $\{v_1, \dots, v_n\}$ is an ordered basis for $V_1$ and $\{w_1, \dots, w_m\}$ is an ordered basis for $V_2$ then the matrix elements of $(T_1 \otimes T_2)(g)$ will be of the form

$$(M_1 \otimes M_2)(g)_{ia,jb} = (M_1(g))_{ij}(M_2(g))_{ab} \tag{11.54}$$

Note that while the set $\{v_i \otimes w_a\}$ forms a basis for $V_1 \otimes V_2$ it does not in any natural way define an ordered basis: One needs to make a further choice of how to order this basis to get a tensor product matrix. In general there are several reasonable choices of ordering. Once one has done this, there will be an ordering on pairs $(i, a)$ and (11.54) will define the matrix relative to this ordered basis.

[152] For the mathematically inclined: $V \otimes W$ is the universal vector space defined by these properties. See LAUM for more details.

3. Given a representation of $V$ we get a dual representation on the dual space $V^\vee := \mathrm{Hom}(V,\kappa)$ of linear maps $V \to \kappa$. We do this by demanding that the natural pairing between $V$ and $V^\vee$ is preserved by the group transformation:

$$\langle T^\vee(g)\ell, T(g)v\rangle = \langle \ell, v\rangle, \tag{11.55}$$

where $\ell \in V^\vee, v \in V$. If $V$ is finite dimensional and $\{v_1,\dots,v_n\}$ is an ordered basis then there is a natural ordered basis $\{v_1^\vee,\dots,v_n^\vee\}$ for $V^\vee$ and relative to these bases we have

$$(M_{T^\vee}(g)) = M_T(g)^{tr,-1} \tag{11.56}$$

4. If vector spaces $V$ and $W$ are carrier spaces of representations of a group $G$ then the vector space of linear transformations from $V$ to $W$, denoted $\mathrm{Hom}(V,W)$, is canonically a representation space for $G$. This can be seen in two ways: First, it is a general fact of linear algebra that, as a vector space, $\mathrm{Hom}(V,W)$ is naturally isomorphic to $V^\vee \otimes W$. As we have seen, if $V$ and $W$ are representation spaces then so is $V^\vee \otimes W$. There is a different point of view which leads to the same conclusion. We view a representation space as a vector space with a $G$-action which happens to be linear. We view $\mathrm{Hom}(V,W)$ as a subspace of the space of all maps from $V \to W$. Now the $G$-action on a map $\phi \in Hom(V,W)$ is determined by our general remarks about induced actions on function spaces from section 8.1.1 above:

$$(g\cdot\phi)(v) := g\left(\phi\left(g^{-1}\cdot v\right)\right) \tag{11.57}$$

Equation (11.57) defines a left $G$-action on $\mathrm{Hom}(V,W)$ which is linear, and hence $\mathrm{Hom}(V,W)$ is a representation. Put differently,

$$\tilde{T}(g)(\phi) := T_W(g)\circ\phi\circ T_V(g^{-1}) \tag{11.58}$$

is the linear action of $g$ on $\phi$. As an exercise the reader should prove that these two viewpoints are compatible. When $V$ and $W$ are finite dimensional vector spaces with ordered bases one can write out the matrix elements of the representation $\mathrm{Hom}(V,W)$. See the exercise below.

5. If $V$ is a complex vector space then the complex conjugate representation sends $g \to \bar{T}(g) \in GL(\bar{V})$. A *real representation* is one where $(\bar{T},\bar{V})$ is equivalent to $(T,V)$. If $\{v_i\}$ is an ordered basis for $V$ then there is a canonical ordered basis $\{\bar{v}_i\}$ for $\bar{V}$ and the matrices are related by

$$M_{\bar{T}}(g) = (M_T(g))^* \tag{11.59}$$

---

**Exercise** *New Matrix Representations From Old Ones*
Given a matrix representation of a group $g \to M(g)$ show that

a.) $g \to (M(g))^{tr,-1}$ is also a representation.

b.) Check the claim (11.57) above.

c.) If $M$ is a matrix representation in $GL(n,\mathbb{C})$ then $g \mapsto M(g)^*$ is also a representation.

d.) If $T$ is a real representation, then there exists an $S \in GL(n,\mathbb{C})$ such that for all $g \in G$:

$$M^*(g) = SM(g)S^{-1} \tag{11.60}$$

Warning: The matrix elements $M(g)_{ij}$ of a real representation might not be real numbers:

e.) Show that the defining two-dimensional representation of $SU(2)$ acting on $\mathbb{C}^2$ is a real representation, but there is no basis in which the matrix elements are all real. [153] ♣Part (e) was already discussed above. ♣

As we will explain later, real representations can be further distinguished as totally real and quaternionic (a.k.a. pseudoreal).

---

---

**Exercise** *Representation Matrices On* Hom$(V, W)$

Let $V = \mathbb{C}^n$ and $W = \mathbb{C}^m$ equipped with their standard ordered bases $e_i$, $i = 1, \dots, n$ and $e_a$, $a = 1, \dots, m$, respectively. Identify $\mathrm{Hom}(V, W) \cong \mathrm{Mat}_{m\times n}(\mathbb{C})$ and consider the basis $e_{a,i}$ for $\mathrm{Mat}_{m\times n}(\mathbb{C})$. Show that if $V$ and $W$ are representation spaces of $G$ then the representation on $\mathrm{Hom}(V, W)$ satisfies:

$$\tilde{T}(g)e_{ai} = M_{ba}(g)(M(g)^{tr,-1})_{ki}e_{bk} \tag{11.61}$$

---

---

**Exercise** *Factoring Representations*

Since a representation is a homomorphism $T : G \to GL(V)$ the kernel of $T$, $H := \ker(T)$ is a normal subgroup of $G$. Show that there is a representation $\bar{T} : G/H \to GL(V)$ on the same carrier space so that

$$\begin{array}{ccc} G & \xrightarrow{T} & GL(V) \\ \downarrow{\scriptstyle \pi} & \nearrow_{\bar{T}} & \\ G/N & & \end{array} \tag{11.62}$$

We say "the representation $T$ factors through $\bar{T}$."

[153] *Answer*: As we have noted, for all $u \in SU(2)$ we have $u^* = (\mathrm{i}\sigma^2)u(\mathrm{i}\sigma^2)^{-1}$. Suppose there were an $S$ so that for all $u \in SU(2)$ we had $SuS^{-1}$ a real $2 \times 2$ matrix. It is not hard to show that theis implies that $S^{-1}S^*\sigma^2 = z1$ for some phase $z$. But taking the complex conjugate of this equation shows that it is inconsistent. So no such matrix $S$ exists.

---

---

**Exercise** *Orderings In A Tensor Product Representation*

Consider two matrix representations of $\mathbb{Z}_2$ where the nontrivial element is represented as

$$M_1(\sigma) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad M_2(\sigma) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{11.63}$$

Show that with the ordering $\{11, 12, 21, 22\}$ we have the tensor product matrix

$$M_1(\sigma) \otimes M_2(\sigma) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{11.64}$$

but with the ordering $\{11, 21, 12, 22\}$ we have instead:

$$M_1(\sigma) \otimes M_2(\sigma) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \tag{11.65}$$

These are the two possible lexicographic orderings. Of course with the other 22 possible orderings of this four-element set we would get other matrices. In general any two matrices are related by conjugation by a permutation matrix.

---

♣Some duplication in this section with previous material ♣

## 11.3 Characters

We now return to a concept introduced around equations (7.33) and (7.34). For any finite-dimensional representation $T : G \to \mathrm{Aut}(V)$ of any group $G$ we can define the *character of the representation*, denoted $\chi_T$. It is a function on the group:

$$\chi_T : G \to \kappa \tag{11.66}$$

and it is defined by

$$\chi_T(g) := \mathrm{Tr}_V(T(g)) \tag{11.67}$$

Some useful general remarks about characters:

1. The character is independent of any choice of basis for $V$.

2. Equivalent representations define precisely the same character function. Therefore, if two representations have different character functions then they must be inequivalent.

3. $\chi_T(h^{-1}gh) = \chi_T(g)$ for all $g, h \in G$. In other words, $\chi_T(g)$ only depends on $g$ via its conjugacy class. In general, a function $F : G \to \mathbb{C}$ that only depends on conjugacy class, that is, that satisfies $F(h^{-1}gh) = F(g)$ for all $g, h \in G$ is known as a *class function*. Such functions "descend" to functions on the set of conjugacy classes of $G$.

4. $\chi_{T_1 \oplus T_2} = \chi_{T_1} + \chi_{T_2}$

5. $\chi_{T_1 \otimes T_2} = \chi_{T_1} \chi_{T_2}$

---

**Exercise** *Complex Conjugate Representation*

Show that the fundamental representation of $SU(2)$ is equivalent to its complex conjugate representation.

Show that the fundamental representation of $SU(3)$ is not equivalent to its complex conjugate representation

---

### 11.4 Unitary Representations

In physics, unitary representations play a distinguished role because of Wigner's theorem, as explained above.

**Definition 11.4.1**. Let $V$ be an inner product space. [154] A *unitary representation* is a representation $(V, T)$ such that $\forall g \in G$, $T(g)$ is a unitary operator on $V$, i.e.,

$$\langle T(g)v, T(g)v \rangle = \langle v, v \rangle \qquad \forall g \in G, v \in V \tag{11.68}$$

1. The canonical representation of $S_n$ on $\mathbb{R}^n$ or $\mathbb{C}^n$ is unitary.

2. The fundamental representations of $SO(n)$ and $U(n)$ are unitary.

**Definition 11.4.2**. If a rep $(V, T)$ is equivalent to a unitary rep then such a rep is said to be *unitarizable*.

**Example.** A simple example of non-unitarizable reps are the $\det^\mu$ reps of $GL(n, \kappa)$ with $\kappa = \mathbb{R}, \mathbb{C}$ and the "spin-$s$" representations $\rho_s$ of the Lorentz group $SO_0(1,1)$ described in section 11.2.

---

**Exercise**

a.) Show that if $T(g)$ is a rep on an inner product space then $T(g^{-1})^\dagger$ is a rep also.

[154] See Chapter 2, section 12

b.) Suppose $T : G \to GL(V)$ is a unitary rep on an inner product space $V$. Let $\{v_i\}$ be an ordered orthonormal basis for $V$. Show that the corresponding matrix rep $M(g)_{ij}$ is a unitary matrix rep. That is:

$$M : G \to U(\mathrm{dim}V) \tag{11.69}$$

is a homomorphism.

c.) Show that for a unitary matrix rep the transpose-inverse and complex conjugate representations are equal.

---

---

**Exercise** *Characters Of Unitarizable Representations*
Show that if $(V, T)$ is unitarizable then

$$\chi_\rho(g^{-1}) = \chi_\rho(g)^* \ . \tag{11.70}$$

---

### 11.5 Haar Measure, a.k.a. Invariant Integration

When proving facts about representations a very important tool is the notion of *invariant integration*. In many situations we would like to consider a group to be a "measure space" (see definition below) as well so that we can define the average value of (measurable) functions on the group. In this section we give examples and state some general results (without proof) of the existence of measures which are nicely compatible with the group structure.

If the group is finite then one natural measure assigns equal weight to each group element. Then the expectation value of a function $f : G \to \mathbb{C}$ is:

$$f \to \frac{1}{|G|} \sum_{g \in G} f(g) := \langle f \rangle \tag{11.71}$$

We can write, very suggestively:

$$\frac{1}{|G|} \sum_{g} f(g) := \int_G f(g) dg \tag{11.72}$$

It is useful to think about this construction a little more abstractly. The map $f \mapsto \int_G f(g)dg$ can be viewed as a linear functional on the vector space of complex valued functions on $G$, denoted $Map(G, \mathbb{C})$. Another natural linear functional on $Map(G, \mathbb{C})$ is the "evaluation map" associated to a particular group element $g_0$. It is defined by:

$$\mathrm{ev}_{g_0} : f \mapsto f(g_0) \tag{11.73}$$

As a measure this is the Dirac measure supported at $g_0$. Of course we could take various linear combinations of functionals of the form $\mathrm{ev}_{g_0}$ to get others. What is special about the measure (11.72) is that it satisfies the *left invariance* property:

$$\int_G f(hg) dg = \int_G f(g) dg \tag{11.74}$$

for all $h \in G$.

The left-action on $G$ induces an action $L_h^*$ on the vector space $Map(G,\mathbb{C})$: Let $L_h^*(f)$ denote the function on $G$ defined by:

$$(L_h^* f)(g) := f(hg) \tag{11.75}$$

Then $\langle L_h^*(f)\rangle = \langle f\rangle$.

Note that, in this case of a finite group, the measure is also right-invariant:

$$\int_G f(gh)dg = \int_G f(g)dg \tag{11.76}$$

For a finite group left-invariant and right-invariant measures are unique up to overall scale. Indeed, the most general measure will be of the form

$$\sum_{g\in G} \rho(g) f(g) \tag{11.77}$$

for some weight function $\rho(g)$. Left invariance implies that

$$\begin{aligned} \sum_{g\in G} \rho(g) f(g) &= \sum_{g\in G} \rho(g) f(hg) \\ &= \sum_{g\in G} \rho(h^{-1}g) f(g) \end{aligned} \tag{11.78}$$

Now apply the statement of left-invariance to the "Dirac function" at $g_0$, a.k.a. the characteristic function at $g_0$: [155]

$$\delta_{g_0}(g) := \begin{cases} 1 & g = g_0 \\ 0 & else \end{cases} \tag{11.79}$$

Then $\rho(g_0) = \rho(h^{-1}g_0)$ for every $g_0$ and every $h$. Therefore $\rho(g) = c$ is just some constant function on the group. It is now easy to check that the measure is also right-invariant.

The idea of a left- or right-invariant measure extends to continuous groups. For example for $G = \mathbb{R}$ the general measure is given by

$$\int_{-\infty}^{+\infty} f(x)\rho(x)dx \tag{11.80}$$

for some measure $\rho(x)dx$. By a similar argument to the above we learn that $\rho(x+y)dx = \rho(x)dx$ for all $x, y$, and hence $\rho(x)$ is a constant. Once again, there is a unique left-invariant measure, up to an overall constant, and that measure turns out to be right-invariant as well.

**Remark: Haar's Theorem**: In general, a *measure space* is a set $X$ together with a collection of subsets of $X$, denoted $\mathfrak{B}$ which includes $\emptyset$ and is closed under complement

[155] Do not confuse this function with the Dirac measure $\mathrm{ev}_{g_0}$.

and countable union. The set $\mathfrak{B}$ is the set of "measurable sets" and a measure is then a function [156]

$$\mu : \mathfrak{B} \to \mathbb{R}_+ \tag{11.81}$$

so that $\mu(\emptyset) = 0$ and

$$\mu(\cup_k E_k) = \sum_k \mu(E_k) \tag{11.82}$$

for countable unions of disjoint sets.

Given a group $G$ we can try to make it a measure space. Using the group structure we can ask that the measure and group multiplication be compatible in the sense that if $S \in \mathfrak{B}$ then the translation $gS$ is also a measureable set. Then the measure is said to be "left-invariant" if $\mu(gS) = \mu(S)$ for all measurable subsets $S \in \mathfrak{B}$. Haar's theorem guarantees the existence of a left invariant measure on topological groups under very mild conditions: The topological group $G$ must be "locally compact" and "Hausdorff" - two mild topological conditions on $G$. [157] The set $\mathfrak{B}$ is the smallest set so that $(G, \mathfrak{B})$ is a measure space and $\mathfrak{B}$ contains all open subsets of $G$. The Haar measure should be left-invariant and satisfy a few other technical conditions: It must be such that $\mu(K) < \infty$ for every compact subset $K$ of $G$ and it must satisfy some (very natural) "regularity conditions." Then Haar's theorem states that such measures exist and are unique up to multiplication by a positive constant. There is a similar definition for "right-invariant" measure, and again the theorem guarantees existence and uniqueness up to scale of right-invariant measures. In general the left- and right-invariant measures are not the same.

Given a Haar measure one can define integrals of some class of functions - known as measureable functions. For example, for $G = \mathbb{R}$, a Haar measure is just the usual Lebesgue measure, $dx$, up to an overall constant. Then we can only integrate Lebesgue-measurable functions, i.e. functions such that

$$\int_{-\infty}^{+\infty} f(x)dx \tag{11.83}$$

exists.

As we just mentioned, in general, the left- and right- invariant measures on a topological group need not coincide, even up to scale. An example is discussed in an exercise below. However, in the case of <u>compact</u> groups the left- and right-invariant measures, are

[156] Of course, given two measures we could take complex combinations to get "complex-valued measures." More generally, measures can take other kinds of values. For example, in the discussion of the Born rule above we encountered "projection-valued measures" where the measure space is the real line with Borel-measurable sets as the set $\mathfrak{B}$ and the value on such a set is the projection operator to the subspace of Hilbert space so that the spectrum of some given self-adjoint operator is in the space being "measured." But the basic notion of a measure space is based on a generalization of area and volume, so here we take the codomain to be $\mathbb{R}_+$.

[157] A topological space $X$ is *Hausdorff* if, for all pairs of distinct points $x_1, x_2 \in X$ there exist neighborhoods $x_1 \in \mathcal{U}_1$ and $x_2 \in \mathcal{U}_2$ such that $\mathcal{U}_1 \cap \mathcal{U}_2 = \emptyset$. In other words, open sets separate points. A topological space $X$ is *locally compact* if, for every $x \in X$ there is a compact neighborhood of $x$. That is, there is a compact subspace $K \subset X$ so that $x \in U \subset K$ for some open neighborhood $U$ of $x$. Notably, infinite-dimensional groups that appear naturally in quantum field theory are <u>not</u> locally compact. But nevertheless, we imagine there are measures on them. Thus, to be rigorous, one needs to extend Haar's theorem.

unique and coincide up to scale. For compact Lie groups the essential observation is that left-invariance shows the volume form must be proportional to $\mathrm{Tr}_V(g^{-1}dg \wedge \cdots g^{-1}dg)$ in some nontrivial representation of the Lie algebra. (Different representations give the same measure up to scale - as they must.) [158]

**Examples**:

1. $G = \mathbb{R}$: The most general Haar measure is of the form:

$$\int_{G=\mathbb{R}} f(g)dg := c \int_{-\infty}^{+\infty} dx f(x) \tag{11.84}$$

   where $c$ is a constant.

2. Recall that $SO_0(1,1) \cong \mathbb{R}$ so the invariant measure if we parametrize Lorentz boosts by $B(\theta)$ is just $d\theta$. Recall that the mass hyperboloid $k^\mu k_\mu + m^2 = 0$ has a free transitive action of the group of boosts. We can parametrize it by $k^1 = m \sinh\theta$ and the invariant measure must be $d\theta$, up to a constant. In terms of momentum $k^1$ we therefore can write the invariant measure as:

$$d\theta = \frac{dk^1}{\sqrt{(k^1)^2 + m^2}} = d^2k\delta(k^2 + m^2)\theta(k^0) \tag{11.85}$$

3. $G = \mathbb{Z}$: The most general Haar measure is of the form:

$$\int_{G=\mathbb{Z}} f(g)dg := c \sum_{n\in\mathbb{Z}} f(n) \tag{11.86}$$

   where $c$ is a constant.

4. Now let $G = \mathbb{R}^*_{>0}$ be the multiplicative group of positive real numbers. The most general Haar measure is of the form

$$\int_{G=\mathbb{R}^*_{>0}} f(g)dg := c \int_0^\infty f(x)\frac{dx}{x} \tag{11.87}$$

   where $c$ is a constant. Compare with (11.84).

[158] See Chapter **** for a discussion of the Maurer-Cartan form $g^{-1}dg$ and the notion of left- and right-invariant differential forms on Lie groups. Another, more sophisticated, but also more elegant proof was explained to me by Dan Freed: Recall that for a finite dimensional real vector space $V$ of dimension $n$ there is a $GL(n,\mathbb{R})$ torsor $\mathcal{B}(V)$ of the bases. Consider the character $|\det|$ on $GL(n,\mathbb{R})$. There is a one-dimensional line of equivariant functions $\mathcal{B}(V) \to \mathbb{R}$ transforming according to this character. There is a natural orientation on this line given by the positive functions and a nonzero function corresponds to a measure on $V$. It is unique up to scalar multiplication by a positive constant. Now let $V$ be the vector space of left-invariant vector fields on $G$. A measure on $V$ corresponds to a left-invariant measure on the group. Now right translation acts by a positive scalar, and this gives a homomorphism $G \to \mathbb{R}_{>0}$. But $G$ is compact so the only possible homomorphism is the trivial one. Therefore the measure is also right-invariant.

5. Similarly, consider $G = GL(n, \mathbb{R})$. We can take the matrix elements $g_{ij}$ to be coordinates on the open domain $\{g|\text{det}g \neq 0\} \subset M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. The usual Euclidean measure $\prod_{ij} dg_{ij}$ changes under $g \to g_0 g$ by

$$\prod_{ij} dg_{ij} \to |\text{det}g_0|^n \prod_{ij} dg_{ij} \tag{11.88}$$

and so the most general Haar measure is of the form

$$\int_{G=GL(n,\mathbb{R})} f(g)dg := c \int_{\text{det}g\neq 0} f(g)|\text{det}g|^{-n} \prod_{ij} dg_{ij} \tag{11.89}$$

where $c$ is a constant. The previous example is the special case of this construction with $n = 1$.

6. $G = U(1)$: Up to scale we have the Haar measure:

$$\int_{G=U(1)} f(g)dg := \frac{1}{2\pi\text{i}} \oint_{|z|=1} f(z)\frac{dz}{z} = \int_0^{2\pi} \frac{d\theta}{2\pi} g(\theta) \tag{11.90}$$

where $g(\theta) = f(e^{\text{i}\theta})$ and the range of the last integral is over any interval of length $2\pi$. Here the scale of the measure has been chosen so that the "volume" of the group is 1.

7. For the important case of $G = SU(2)$ we can write the Haar measure as follows. We again review the Euler angles: Every $SU(2)$ matrix can be written as:

$$g = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix} \tag{11.91}$$

for 2 complex numbers $\alpha, \beta$ with

$$|\alpha|^2 + |\beta|^2 = 1. \tag{11.92}$$

In this way we identify the group as a manifold as $S^3$. That manifold has no globally well-defined coordinate chart. The best we can do is define coordinates that cover "most" of the group but will have singularities in some places. (It is always important to be careful about those singularities when using explicit coordinates!) We can always write:

$$\begin{aligned} \alpha &= \zeta_1 \cos\theta/2 \\ \beta &= \zeta_2 \sin\theta/2 \end{aligned} \tag{11.93}$$

where $\zeta_1, \zeta_2$ are phases, and the magnitude is parametrized in a 1-1 fashion by taking $0 \leq \theta \leq \pi$. Next it is standard to parametrize the phases by:

$$\begin{aligned} \alpha &= e^{\text{i}\frac{1}{2}(\phi+\psi)} \cos\theta/2 \\ \beta &= \text{i}e^{\text{i}\frac{1}{2}(\phi-\psi)} \sin\theta/2 \end{aligned} \tag{11.94}$$

The virtue of this definition is that we can then write:

♣Sign of $\phi$ and $\psi$ flipped relative to previous conventions. ♣

$$
\begin{aligned}
g &= \begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix} \begin{pmatrix} \cos\theta/2 & i\sin\theta/2 \\ i\sin\theta/2 & \cos\theta/2 \end{pmatrix} \begin{pmatrix} e^{i\psi/2} & 0 \\ 0 & e^{-i\psi/2} \end{pmatrix} \\
&= e^{i\frac{1}{2}\phi\sigma^3} e^{i\frac{1}{2}\theta\sigma^1} e^{i\frac{1}{2}\psi\sigma^3}
\end{aligned}
\tag{11.95}
$$

and under the standard homomorphism $\pi : SU(2) \to SO(3)$ the angles $\theta, \phi, \psi$ become the Euler angles.

♣This is redundant with earlier discussion of ranges of Euler angles. ♣

We need to be a little careful about $\phi$ and $\psi$ since they are not defined at $\theta = 0, \pi$ and we also need to be careful about their ranges. The above expression is invariant under the transformations:

$$
\begin{aligned}
(\phi, \psi) &\to (\phi + 4\pi, \psi) \\
(\phi, \psi) &\to (\phi, \psi + 4\pi) \\
(\phi, \psi) &\to (\phi + 2\pi, \psi + 2\pi)
\end{aligned}
\tag{11.96}
$$

If we think of this as generating a group of transformations on $\mathbb{R}^2$ we can choose a fundamental domain in various ways, and then taking $(\phi, \psi)$ to be in that fundamental domain we will cover the group elements exactly once (away from $\theta = 0, \pi$). One standard fundamental domain is:

$$
\begin{aligned}
0 \le \phi < 2\pi \\
0 \le \psi < 4\pi
\end{aligned}
\tag{11.97}
$$

This is good because if we then act on the unit vector $e_3$, $(\phi, \theta)$ become the standard angular coordinates on $S^2$.

The normalized Haar measure for $SU(2)$ in these coordinates is [159]

$$
[dg] = \frac{1}{16\pi^2} d\psi \wedge d\phi \wedge \sin\theta d\theta \tag{11.98}
$$

8. Quite generally, measures "push forward." This means the following. First a function $f : X_1 \to X_2$ is called a measurable function $f : (X_1, \mathfrak{B}_1) \to (X_2, \mathfrak{B}_2)$ between two measure spaces if for all $S \in \mathfrak{B}_2$ we have $f^{-1}(S) \in \mathfrak{B}_1$. Now, suppose that $\mu_1 : \mathfrak{B}_1 \to \mathbb{R}_+$ is a measure. Then we define the pushforward measure $f_*\mu_1 : \mathfrak{B}_2 \to \mathbb{R}_+$ by the formula $(f_*\mu)(S) = \mu(f^{-1}(S))$ for $S \in \mathfrak{B}_2$. In general if $H \subset G$ is a subgroup of a Lie group then we can define an average over functions on $G$ where we just integrate over $H$:

$$
\pi_H(f)(g) := \int_H f(gh) dh \tag{11.99}
$$

where $dh$ is a left-invariant measure on $H$. Note that $\pi_H(f)$ is really just a function on the homogeneous space $G/H$. Then, there exists a measure $d\mu_{G/H}$ on $G/H$ so that

$$
\int_G f(g) dg = \int_{G/H} \pi_H(f) d\mu_{G/H} \tag{11.100}
$$

[159] We have chosen an orientation so that, with a positive constant, this is $\mathrm{Tr}_{\mathbf{2}}(g^{-1}dg)^3)$.

As an example, let $H$ be the subgroup of diagonal elements of $SU(2)$. The projection $\pi_H$ takes a function on $SU(2)$, expressed as a function of Euler angles, $f(\theta,\phi,\psi)$ to a function $\pi_H(f)$ on $SU(2)/U(1) \cong S^2$, expressed in terms of standard polar angles $(\theta,\phi)$. It now follows (just integrate (11.98) over $\psi$) that we recover

$$d\mu_{G/H} = \frac{1}{4\pi} \sin\theta d\theta d\phi \ , \tag{11.101}$$

the standard unit-volume measure on $S^2$ in the round metric. Similarly, one can derive the relativistic measure on the mass shell hyperboloid, regarded as $SO(1,d)/SO(d)$.

9. For much more about this, see Chapter 5 below. For a compact group the unit normalized Haar measure is unique. In the case of $SU(2)$, equation (11.98) is one way of writing the Haar measure, but this is not the most useful way to write it when integrating class functions. We encounter a different way of writing the measure, more suitable for integrating class functions, in Section 11.19.2 below, where the generalization to the higher rank classical matrix groups is given (a formula known as the Weyl Density formula).

10. More information can be found in Leopoldo Nachbin's *The Haar Integral* or Elizabeth Meckes, *The Random Matrix Theory of the Classical Compact Groups.*

A consequence of the existence of invariant integration is that every finite-dimensional representation of a compact group is equivalent to a unitary representation. This follows because you can use a suitable averaging procedure over the group to make the matrices unitary.

**Proposition** If $(T,V)$ is a rep of a *compact group* $G$ and $V$ is an inner product space then $(T,V)$ is unitarizable.

*Proof* We make essential use of the Haar measure. If $T$ is not already unitary with respect to the inner product $\langle\cdot,\cdot\rangle_1$ then we can define a new inner product by:

$$\langle v,w\rangle_2 := \int_G \langle T(g)v, T(g)w\rangle_1 \ dg \tag{11.102}$$

For a compact group $\langle T(g)v, T(g)w\rangle_1$ will be a nice continuous function, hence bounded, and therefore integrable. If $v \neq 0$ then $T(g)v \neq 0$ hence $\langle T(g)v, T(g)v\rangle_1 > 0$ and therefore $0 < \langle v,v\rangle_2 < \infty$. Here we used the fact that for a compact group the volume is finite. So $\langle\cdot,\cdot\rangle_2$ will be a good inner product. Then using the properties of the Haar measure it is easily checked that

$$\langle T(g)v, T(g)w\rangle_2 = \langle v,w\rangle_2 \tag{11.103}$$

and hence $T(g)$ is unitary w.r.t. the inner product $\langle\cdot,\cdot\rangle_2$ ♠

**Remarks**:

1. We saw that the representations $\det^{\mu}$ of $GL(n,\mathbb{R})$, $GL(n,\mathbb{C})$ and $T_s$ of $SO_0(1,1)$ are not unitarizable. What fails in the above argument is the infinite volume of these noncompact groups.

2. The above proposition has the following corollary in terms of matrix representations: Suppose that $M : G \to GL(n,\mathbb{C})$ is a matrix representation of a compact group $G$. Then, there exists an invertible matrix $A$ (independent of $g$) so that

$$U(g) = AM(g)A^{-1} \tag{11.104}$$

are unitary matrices for all $g \in G$. To see this suppose, quite generally that a vector space has two inner products, $\langle \cdot, \cdot \rangle_{1,2}$ and that $\{u_i^{(1)}\}$ is an ON basis for $\langle \cdot, \cdot \rangle_1$ and $\{u_i^{(2)}\}$ is an ON basis for $\langle \cdot, \cdot \rangle_2$. We can relate the two bases by

$$u_i^{(1)} = \sum_k A_{ki} u_k^{(2)} \tag{11.105}$$

Then, in general $A$ is not a unitary matrix. Indeed

$$\langle u_i^{(1)}, u_j^{(1)} \rangle_2 = (A^\dagger A)_{ij} \tag{11.106}$$

Now, suppose that $U$ is a unitary operator wrt the inner product $\langle \cdot, \cdot \rangle_2$. Choose an ordering of the basis $\{u_i^{(2)}\}$ and let $U_{ij}$ be the matrix for $U$ relative to this basis. The matrix with matrix elements $U_{ij}$ will be unitary. Next, choose an ordering of $\{u_i^{(1)}\}$. Then the matrix of $U$ relative to $\{u_i^{(1)}\}$ is

$$A^{-1} U A = U^{(1)} \tag{11.107}$$

which, in general will not be a unitary matrix. Put differently, if $U$ is a unitarizable operator, but, wrt some randomly chosen inner product it has matrix $U^{(1)}$ relative to an ON basis for that inner product, then unitarizability means there exists an invertible matrix $A$ so that $AU^{(1)}A^{-1}$ is unitary.

---

**Exercise** *Due Diligence*
Show that

$$\langle T(g)v, T(g)w \rangle_2 = \langle v, w \rangle_2 \tag{11.108}$$

---

---

**Exercise** *Computation For $SU(2)$*

a.) Using the parametrization by Euler angles show that

$$\int_{SU(2)} g_{\alpha\beta} dg = 0 \tag{11.109}$$

$$\int_{SU(2)} g_{\alpha\beta} g_{\gamma\delta} dg = \frac{1}{2}\epsilon_{\alpha\gamma}\epsilon_{\beta\delta} \tag{11.110}$$

We will later interpret these equations as special cases of the orthogonality relations of matrix elements in irreducible representations. (See section on the Peter-Weyl theorem below.)

b.) Using left-invariance and right-invariance show that

$$\int_{SU(2)} g_{\alpha_1\beta_1} \cdots g_{\alpha_n\beta_n} dg \tag{11.111}$$

can only be nonzero if $n$ is even and half of the $\alpha_1, \dots, \alpha_n$ are 1 (so the other half are 2) and similarly for $\beta_1, \dots, \beta_n$.

**Remark**: Integrals like (11.111)show up in many contexts in physics. One example is in the strong coupling expansion of lattice gauge theory. A second example is in the $k \to \infty$ limit of correlation functions of vertex operators in the WZW model, a famous 2d conformal field theory.

---

---

**Exercise** *An Example Where Left-Invariant And Right-Invariant Measures Do Not Agree*

Consider the group of upper triangular real matrices of the form

$$\begin{pmatrix} x & y \\ 0 & x^{-1} \end{pmatrix} \tag{11.112}$$

a.) Show that a left-invariant measure must be of the form $\frac{dxdy}{x^2}$ up to an overall constant.

b.) Show that a right-invariant measure must be of the form $dxdy$ up to an overall constant.

c.) In general, if $d\mu_L(g)$ is a left-invariant measure on $g$ then there exists a function, known as the *modular function* $\Delta : G \to \mathbb{R}_+$, so that

$$\int_G f(gh) d\mu_L(g) = \Delta(h) \int_G f(g) d\mu_L(g) \tag{11.113}$$

Accepting that such a function $\Delta$ exists show that it is a group homomorphism. [160]

[160] As far as I know the term "modular" has no relation to the use of "modular" in the modular group $SL(2,\mathbb{Z})$ and the theory of automorphic forms.

d.) If $G$ is compact show that $\Delta(h) = 1$. [161]

---

### 11.6 The Regular Representation

Let $G$ be a group. Then there is a left action of $G \times G$ on $G$: $(g_1, g_2) \mapsto L(g_1)R(g_2^{-1})$:

$$(g_1, g_2) \cdot g_0 = g_1 g_0 g_2^{-1} \tag{11.114}$$

and hence an induced action on $\mathrm{Map}(G, Y)$ for any $Y$. Now let $Y = \mathbb{C}$. Then $\mathrm{Map}(G, \mathbb{C})$ is a representation of $G \times G$ because the induced left-action:

$$((g_1, g_2) \cdot \Psi)(h) := \Psi(g_1^{-1} h g_2) \tag{11.115}$$

converts the vector space of functions $\Psi : G \to \mathbb{C}$ into a representation space for $G \times G$.

If we equip $G$ with a left- and right-invariant Haar measure then we can speak of $L^2(G)$, namely the Hilbert space based on the complex-valued functions such that

$$\langle f, f \rangle := \int_G |f(g)|^2 dg < \infty \tag{11.116}$$

Note that the $G \times G$ action preserves the $L^2$-property thanks to left- and right- invariance, and in fact the $G \times G$ action is unitary.

**Definition** The representation $L^2(G)$ is known as the *regular representation of* $G$.

Note that $L^2(G)$ is a representation of $G \times G$ although by restriction to the natural subgroups $G \times \{1\}$ and $\{1\} \times G$ it becomes a representation of $G$. So sometimes people speak of "the regular representation of $G$." More precisely, if we restrict to operations of the form:

$$(L(h) \cdot \Psi)(g) := \Psi(h^{-1} g) \tag{11.117}$$

We have the *left regular representation*, while

$$(R(h) \cdot \Psi)(g) := \Psi(gh) \tag{11.118}$$

defines the *right regular representation*. Both actions are left- actions on the function space, so the terminology is slightly confusing.

Suppose that $(T, V)$ is a representation of $G$. As we explained above the vector space of linear transformations $\mathrm{End}(V) := \mathrm{Hom}(V, V)$ of $V$ to itself is then also a representation. In fact, it is a representation of $G \times G$ because if $S \in \mathrm{End}(V)$ then we can define a linear left-action of $G \times G$ on $\mathrm{End}(V)$ by:

$$(g_1, g_2) \cdot S := T(g_1) \circ S \circ T(g_2)^{-1} \tag{11.119}$$

[161] We know that $\Delta(1) = 1$. From the definition $\Delta$ must be continuous. Since $G$ is compact $\Delta$ must assume its maximal and minimal value on some elements $g_{mx}$ and $g_{min}$ respectively. If $\Delta(g_{mx}) > 1$ then $\Delta(g_{mx}^2) = \Delta(g_{mx})^2 > \Delta(g_{mx})$, a contradiction. So $\Delta(g_{mx}) = 1$. Similarly, $\Delta(g_{min}) = 1$. Therefore $\Delta(g) = 1$ for all $g \in G$.

We now have two representations of $G\times G$. First, we defined $L^2(G)$ as a representation of $G\times G$ and now we have recognized $\mathrm{End}(V)$ as a carrier space for another representation of $G\times G$. These two representations are related as follows: If $V$ is finite-dimensional we have a map

$$\iota : \mathrm{End}(V) \to L^2(G) \tag{11.120}$$

The map $\iota$ takes a linear transformation $S : V \to V$ to the complex-valued function $\Psi_S : G \to \mathbb{C}$ defined by

$$\Psi_S(g) := \mathrm{Tr}_V(ST(g^{-1})) \tag{11.121}$$

That is

$$\iota(S) := \Psi_S \tag{11.122}$$

We claim that $\iota$ is a $G\times G$-equivariant map:

$$(h_1,h_2)\cdot\Psi_S = \Psi_{(h_1,h_2)\cdot S} \tag{11.123}$$

You are asked to prove this in an exercise below. Put differently, denoting by $T_{\mathrm{End}(V)}$ the representation of $G\times G$ on $\mathrm{End}(V)$ and $T_{Reg.Rep.}$ the representation of $G\times G$ on $\mathrm{Map}(G,\mathbb{C})$ we get a commutative diagram:

$$\begin{array}{ccc} \mathrm{End}(V) & \xrightarrow{\iota} & \mathrm{Map}(G,\mathbb{C}) \\ \downarrow{\scriptstyle T_{\mathrm{End}(V)}} & & \downarrow{\scriptstyle T_{Reg.Rep.}} \\ \mathrm{End}(V) & \xrightarrow{\iota} & \mathrm{Map}(G,\mathbb{C}) \end{array} \tag{11.124}$$

So, $\iota$ is an intertwiner.

If we choose an ordered basis $\{v_i\}$ for $V$ then the operators $T(g)$ are represented by matrices:

$$T(g)\cdot v_i = \sum_j M(g)_{ji} v_j \tag{11.125}$$

If we take $S = e_{ij}$ to be the matrix unit in this basis then $\Psi_S$ is the function on $G$ given by the matrix element $M(g^{-1})_{ji} = M^{tr,-1}(g)_{ij}$. So the $\Psi_S$'s are linear combinations of matrix elements of the representation matrices of $G$. (Replacing $V$ by its dual $V^\vee$ we will get the representation matrices $M(g)_{ij}$.) The advantage of (11.121) is that it is completely canonical and basis-independent.

See section 11.10 below for more about the regular representation.

---

**Exercise** *Due Diligence*

Prove equation (11.123). [162]

---

[162] *Answer*: This is a straightforward computation:

$$\begin{aligned} (h_1,h_2)\cdot\Psi_S(g) &= \Psi_S(h_1^{-1}gh_2) \\ &= \mathrm{Tr}_V(ST(h_2^{-1}g^{-1}h_1)) \\ &= \mathrm{Tr}_V(T(h_1)ST(h_2^{-1})T(g^{-1})) \\ &= \mathrm{Tr}_V((h_1,h_2)\cdot ST(g^{-1})) \\ &= \Psi_{(h_1,h_2)\cdot S}(g) \end{aligned} \tag{11.126}$$

**Exercise**

a.) Let $\delta_0, \delta_1, \delta_2$ be a basis of functions in the regular representation of $\mu_3$ which are 1 on $1, \omega, \omega^2$, respectively, and zero elsewhere. Show that $\omega$ is represented as

$$L(\omega) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{11.127}$$

b.) Show that

$$\begin{aligned} L(h) \cdot \delta_g &= \delta_{h \cdot g} \\ R(h) \cdot \delta_g &= \delta_{g \cdot h^{-1}} \end{aligned} \tag{11.128}$$

and conclude that for the left, or right, regular representation of a finite group the representation matrices in the $\delta$-function basis are permuation matrices.

## 11.7 Reducible And Irreducible Representations

### 11.7.1 Definitions

In general, bigger matrices are harder to work with than smaller matrices. Put differently, representations of a group on a vector space of "large dimension" are harder to understand and work with than representations on a vector space of "small dimension." So, presenting a representation as a direct sum of smaller ones is often a very useful simplification. In terms of matrices it corresponds to block diagonalization. We now investigate how that can be done systematically.

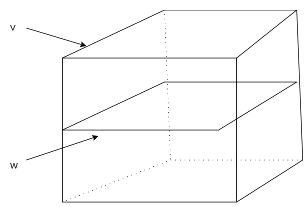



**Figure 29:** $T(g)$ preserves the subspace $W$.

**Definition**. Let $W \subset V$ be a linear subspace of the carrier space $V$ of a group representation $T : G \to GL(V)$. Then $W$ is *invariant* under $T$, a.k.a. an *invariant subspace* if $\forall g \in G, w \in W$

$$T(g)w \in W \tag{11.129}$$

This may be pictured as in 29.

**Examples**

1. Both $\{\vec{0}\}$ and $V$ are always invariant subspaces.

2. Consider the three-dimensional representation of $SO(2)$ as rotations around the $z$-axis. Then the vector subspace of the $xy$ plane at $z = 0$ is an invariant subspace. Note that the parallel planes at $z = z_0 \neq 0$ are invariant under the group action but are not linear subspaces.

3. Consider the canonical representation of $S_n$ on $\kappa^n$. The line through the all ones vector is an invariant subspace.

4. We saw above that spaces of functions of the form (11.121) for a fixed $V$ define invariant subspaces of the regular representation with action (11.115). Put differently $\iota(\mathrm{End}(V)) \subset L^2(G)$ is an invariant subspace. We can re-express this in more concrete terms as follows:

   Let $M : G \to GL(n, \kappa)$ be *any* matrix $n$-dimensional representation of $G$. For a fixed $i, j$ consider the matrix element $M_{ij}$ as a $\kappa$-valued function on $G$: So $M_{ij}$ is the function on $G$ whose value at $g \in G$ is just $M(g)_{ij} \in \kappa$. Now consider the linear span of functions where we fix $i$:

   $$\mathcal{R}_i := Span\{M_{ij}\}_{j=1,\dots,n} \tag{11.130}$$

   We claim this is an invariant subspace of $L^2(G)$ in the right regular representation: To see this, just check:

   $$\begin{aligned}(R(g) \cdot M_{ij})(h) &= M_{ij}(hg) \\ &= \sum_{s=1}^{n} M_{is}(h) M_{sj}(g)\end{aligned} \tag{11.131}$$

   which is equivalent to the equation on functions:

   $$\overbrace{R(g) \cdot M_{ij}}^{\text{function on G}} = \sum_{s=1}^{n_\mu} \underbrace{M_{sj}(g)}_{\text{matrix element for } \mathcal{R}_i} \overbrace{M_{is}}^{\text{function on G;vector in } \mathcal{R}_i} \tag{11.132}$$

   Similarly, suppose we fix $j$ and consider:

   $$\mathcal{L}_j := Span\{M_{ij}\}_{i=1,\dots,n} \tag{11.133}$$

   is an invariant subspace under the left-regular representation of $G$. Putting these together we see that the space

   $$\mathcal{LR} := Span\{M_{ij}\}_{i,j=1,\dots,n} \tag{11.134}$$

   is an invariant subspace under the left-action of $G \times G$ on $L^2(G)$ defined by (11.115). As a representation of $G$, under the left regular representation we have the decomposition:

   $$\mathcal{LR} \cong \oplus_{i=1}^{n} \mathcal{L}_i \tag{11.135}$$

   and similarly for $\mathcal{R}_i$. So, if $n > 1$ then we immediately conclude that $\mathcal{LR}$ is reducible as a representation of $G$.

**Remarks:**

1. If $(V,T)$ is a rep and $W \subset V$ is an invariant subspace we can define a smaller group representation $(T,W)$ called *restriction of* $T$ *to* $W$. We also say the $(T,W)$ *is a subrepresentation of* $(V,T)$. Strictly speaking we should write $T|_W$ but we will generally not write that out.

2. If $T$ is unitary on $V$ then it is unitary on $W$.

**Definition**. A representation $T$ is called *reducible* if there is an invariant subspace $W \subset V$, under $T$, which is nontrivial, i.e., such that $W \neq 0, V$, If $V$ is not reducible we say $V$ is *irreducible*. That is, in an irreducible rep, the only invariant subspaces are $\{\vec{0}\}$ and $V$. We often shorten the term "irreducible representation" to "irrep."

**Remarks**:

1. Given any nonzero vector $v \in V$, the linear span of $\{T(g)v\}_{g\in G}$ is an invariant subspace. In an irrep this will span all of $V$. Such a vector is called a *cyclic vector*. Caution!! The existence of a cyclic vector does not imply the representation is irreducible. Consider the vector $e_1$ in the permutation representation of $S_n$ on $\mathbb{R}^n$.

2. Suppose $(T,W)$ is a subrepresentation of $(T,V)$. Choose an ordered basis

$$\{w_1, \dots, w_k\} \tag{11.136}$$

for $W$. Then it can be completed to an ordered basis

$$\{w_1, \dots, w_k, u_{k+1}, \dots, u_n\} \tag{11.137}$$

for $V$. Let us write $w_i$, $i = 1, \dots, k$ for the basis vectors for $W$ and $u_a$, $a = k+1, \dots, n$ for a <u>choice</u> of a set of complementary ordered basis vectors for $V$. The matrix representation associated with such a choice of basis is defined by:

$$\begin{aligned} T(g)(w_i) &= (M_{11}(g))_{ji} w_j + (M_{12}(g))_{ai} u_a \\ T(g)(u_a) &= (M_{21}(g))_{ja} w_j + (M_{22}(g))_{ba} u_b \end{aligned} \tag{11.138}$$

where

$$\begin{aligned} &M_{11}(g) \in Mat_{k\times k} \\ &M_{12}(g) \in Mat_{k\times(n-k)} \\ &M_{21}(g) \in Mat_{(n-k)\times k} \\ &M_{22}(g) \in Mat_{(n-k)\times(n-k)} \end{aligned} \tag{11.139}$$

Because $W$ is an invariant subspace we have $M_{12} = 0$ so the matrix representation looks like:

$$M(g) = \begin{pmatrix} M_{11}(g) & 0 \\ M_{21}(g) & M_{22}(g) \end{pmatrix} \tag{11.140}$$

Since $M(g_1)M(g_2) = M(g_1 g_2)$ it follows that $M_{11}$ gives a matrix representation on $W$.

3. If $W \subset V$ is an invariant subspace then the quotient vector space (see Chapter two) $V/W$ is a representation of $G$ in a natural way:

$$T(g)(v + W) := T(g)(v) + W \tag{11.141}$$

The reader can check this is well-defined. The vectors $u_a$ above define a basis for $V/W$ of the form $u_a + W$, and relative to this basis the representation will look like $M_{22}$.

4. The main point of the discussion of equation (11.130) et. seq. above is that in general $L^2(G)$ is highly reducible.

**Examples**

1. Consider $G = \mathbb{Z}_2$. There are exactly two irreducible representations and they are both one-dimensional, so the carrier space for both representations is $V = \mathbb{R}$ if we work over the real field. We denote the trivial representation by $\rho_+$. The other irrep is

$$\rho_-(g) = \begin{cases} 1 & g = 1 \\ -1 & g \neq 1 \end{cases} \tag{11.142}$$

If we work over $\kappa = \mathbb{C}$ then the carrier space of the irreducible representations is $V = \mathbb{C}$. We can consider $V = \mathbb{C}$ to be a two-dimensional real vector space so under change of ground field an irreducible representation can become reducible.

2. Now consider the following two-dimensional real representation of $G = S_2 \cong \mathbb{Z}_2$:

$$\begin{aligned} 1 &\rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ (12) &\rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned} \tag{11.143}$$

is a 2-dimensional *reducible* rep on $\mathbb{R}^2$ because $W = \{(x, x)\} \subset \mathbb{R}^2$ is a nontrivial invariant subspace. Indeed, $\sigma^1$ is diagonalizable, so this rep is equivalent to

$$\begin{aligned} 1 &\rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ (12) &\rightarrow \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{11.144}$$

which is a direct sum of two $1 \times 1$ reps:

$$(T, V) \cong \rho_+ \oplus \rho_- \tag{11.145}$$

3. Let $G = U(1)$. Then, for any $n \in \mathbb{Z}$ we define the one-dimensional representation $\rho_n$ acting on the carrier space $V = \mathbb{C}$ as multiplication by

$$\rho_n(z) = z^n \tag{11.146}$$

for $z \in U(1)$. This is clearly an irreducible representation. Moreover, the representations are inequivalent for distinct integers. We will argue below that these are the only irreducible representations.

4. *Finite-dimensional representations of Abelian groups.* Choosing an ordered ON basis the matrices $M(g)$, $g \in G$ are commuting unitary matrices and, over the complex field, they can be Simultaneously diagonalized, by the spectral theorem. For example, a fd unitary representation of $U(1)$ will be a family of commuting matrices and we can choose a basis so that

$$M(z) = Diag\{z^{n_1}, \dots, z^{n_d}\} \tag{11.147}$$

so if $V \cong \mathbb{C}^d$ is the carrier space we would write

$$V \cong \rho_{n_1} \oplus \cdots \oplus \rho_{n_d} \tag{11.148}$$

On the other hand, if we look at finite dimensional representations of $SO(2)$ over the real numbers then the best we can do is block diagonalize into $2 \times 2$ blocks:

$$M(R(\theta)) = R(n_1\theta) \oplus R(n_2\theta) \oplus \cdots \oplus R(n_j\theta) \tag{11.149}$$

5. Now consider the nonabelian group $S_3$. There is a natural 3-dimensional permutation representation of $S_3$ we defined above: $T(\sigma)e_i = e_{\sigma(i)}$ where $e_i$ is the standard basis for $\mathbb{R}^3$. As we have noted, the all ones vector $u_0 := e_1 + e_2 + e_3$ spans an invariant subspace $L \subset \mathbb{R}^3$. We can choose basis vectors for the orthogonal complement

$$\begin{aligned} u_1 &:= e_1 - e_2 \\ u_2 &:= e_2 - e_3 \end{aligned} \tag{11.150}$$

and the subspace spanned by $u_1, u_2$ is also an invariant subspace. In fact, one easily computes that relative to this basis:

$$\begin{aligned} &M((12)) = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \qquad M((23)) = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \qquad M((13)) = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \\ &M((123)) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \qquad M((132)) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \end{aligned} \tag{11.151}$$

Note that this representation is clearly irreducible: Since it is a representation of a finite group it is completely reducible (see below). The reduction would have to give diagonal matrices. But diagonal matrices commute. The above matrices do not commute since $T(12)T(23) = T(123)$ is not the same as $T(23)T(12) = T(132)$. ♣"completely reducible" not yet define. ♣

Note that the basis $u_1, u_2$ is <u>not</u> an ON basis and the above matrices are <u>not</u> unitary matrices, even though this is a unitary representation. One could use the averaging procedure above to make a unitary representation, although this would be tedious. A better way to proceed is to note that the reflections and rotations in the plane that preserve an equilateral triangle form the group $S_3$. (We will discuss this in much more detail later.) If we label the vertices 123 in counter-clockwise order with vertex 3 on the $y$-axis, as in figure 39 then:

$$M((12)) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{11.152}$$

$$M((13)) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \tag{11.153}$$

$$M((23)) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \tag{11.154}$$

$$M((123)) = R(\frac{2\pi}{3}) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \tag{11.155}$$

$$M((132)) = R(\frac{-2\pi}{3}) = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \tag{11.156}$$

generates a unitary representation.

6. Consider the representation of $S_n$ on $\mathbb{R}^n$. Then the one-dimensional subspace $L = \{(x, \cdots, x)\}$ is a subrepresentation. Moreover we can take

$$L^\perp = \{(x_1, \cdots x_n) | \sum x_i = 0\}. \tag{11.157}$$

Then $L^\perp$ is an $(n-1)$-dimensional representation of $S_n$ and the representation is equivalent to a direct sum $L \oplus L^\perp$. So, if we choose a basis of the all ones vector and an orthogonal basis for $L^\perp$ the matrices will be block diagonal. It is not obvious, but it does follow from the general representation theory of the symmetric group that the representation $L^\perp$ of dimension $(n-1)$ is irreducible.

7. The $N$-dimensional fundamental representation of $SU(N)$ is irreducible.

### 11.7.2 Reducible vs. Completely reducible representations

It is always useful to discover a nontrivial subgrepresentation of a $G$ representation. However, such subrepresentations cannot always be "disentangled" or "decoupled" from the remainder of the representation. When this can be done we have a stronger form of reducibility known as complete reducibility:

**Definition**. A representation $(T,V)$ is called *completely reducible* if it is isomorphic to a direct sum of representations:

$$V \cong W_1 \oplus \cdots \oplus W_n \tag{11.158}$$

where the $W_i$ are irreducible reps. Thus, there is a basis in which the matrices look like:

$$M(g) = \begin{pmatrix} M_{11}(g) & 0 & 0 & \cdots \\ 0 & M_{22}(g) & 0 & \cdots \\ 0 & 0 & M_{33}(g) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \tag{11.159}$$

with each block $M_{ii}(g)$ being an irreducible matrix representation of $G$.

All the examples we studied above were completely reducible.

In real life it can and does actually happen that a group $G$ has representations which are reducible but not completely reducible. Reducible, but not completely reducible reps are sometimes called *indecomposable.*

**Example 1** An example of an indecomposible rep which is not completely reducible is the rep

$$A \to \begin{pmatrix} 1 & \log|\det A| \\ 0 & 1 \end{pmatrix} \tag{11.160}$$

of $GL(n,\mathbb{R})$.

**Example 2** Similarly, we can write an indecomposable representation of the connected component of the identity of the $1+1$ dimensional Lorentz group. If $B(\eta) \in SO_0(1,1)$ is the boost of rapidity $\eta$, with $\eta \in \mathbb{R}$, then:

$$T(B(\eta)) = \begin{pmatrix} 1 & \eta \\ 0 & 1 \end{pmatrix} \tag{11.161}$$

**Example 3** Let $G = GL(n,\kappa)$ and $H = \mathrm{Mat}_n(\kappa)$. We can define a group which, as a set is $H \times G$, but it has a "twisted" group multiplication law:

$$\mathbf{m}((h_1,g_1),(h_2,g_2)) := (h_1 + g_1 h_2 g_1^{tr}, g_1 g_2) \tag{11.162}$$

One can check that this really does define a group structure. It is a special case of the semi-direct product structure that we will study in more detail in Section **** . Now check that the following is a matrix representation:

$$T(h,g) := \begin{pmatrix} g & h g^{tr,-1} \\ 0 & g^{tr,-1} \end{pmatrix} \tag{11.163}$$

Because of the upper-triangular nature of the matrices we cannot simultaneously block-diagonalize all the matrices $T(h,g)$ but there is a nontrivial invariant subspace. This kind

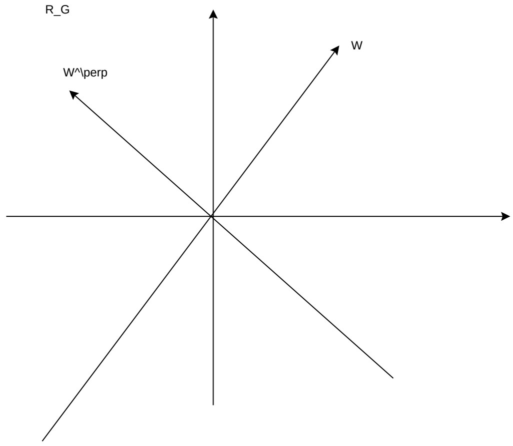



**Figure 30:** The orthogonal complement of an invariant subspace is an invariant subspace.

of construction gives the finite-dimensional indecomposable representations of the Poincaré group, the Euclidean group, and its crystallographic subgroups. ♣Explain this last sentence more. ♣

It is useful to have criteria for when this complication cannot occur:

**Proposition 11.7.2.1** Suppose that the representation $(V, T)$ is a unitary representation on an inner product space $V$ and $W \subset V$ is an invariant subspace then $W^{\perp}$ is an invariant subspace:

*Proof*: Recall that

$$y \in W^{\perp} \Leftrightarrow \forall x \in W, \quad \langle y, x \rangle = 0 \tag{11.164}$$

Let $g \in G$, $y \in W^{\perp}$. Compute

$$\begin{aligned} \langle T(g)y, x \rangle &= \langle y, T(g)^{\dagger} x \rangle \\ &= \langle y, T(g^{-1})x \rangle \end{aligned} \tag{11.165}$$

But, $T(g^{-1})x \in W$, since $W \subset V$ is an invariant subspace. Therefore: $\forall x \in W$, $\langle T(g)y, x \rangle = 0$. Therefore: $T(g)y \in W^{\perp}$ Therefore: $W^{\perp}$ is an invariant subspace. ♠

Therefore:

1. *Finite dimensional unitary representations are always completely reducible.* This follows from the above proposition together with induction on the dimension.

2. *Finite dimensional representations of compact groups are always completely reducible.*

3. *In particular, for a finite group the regular representation $L^2(G)$ is completely reducible as a representation of $G$ or of $G \times G$.*

4. In section 11.10 below we will show that for all compact groups $G$, $L^2(G)$ is completely reducible. We will also show that the irreducible representations are all finite

dimensional. Now if $G$ has positive dimension then $L^2(G)$ is an infinite-dimensional vector space. Therefore, in this case the decomposition of $L^2(G)$ into irreps will be an infinite direct sum. The specific form of that decomposition, which holds uniformly for all compact groups is the *Peter-Weyl theorem* described below.

5. Above we described the finite dimensional representations of the Poincaré and Euclidean groups (without proving these are the only ones, but they are the only ones). They are reducible but not fully reducible. It follows that they cannot be unitary. The Poincaré and Euclidean groups do have unitary irreducible representations, but they are infinite-dimensional.

**Isotypical Components**: Assume that the set of irreducible representations of $G$ (up to isomorphism) is a countable set. This turns out to be the case for all compact groups, but is typically not true for noncompact groups. For each isomorphism class of irreducible representation of $G$ choose a representative $(T^{(\mu)}, V^{(\mu)})$ where $\mu$ runs over the set of distinct irreducible representations. If $V$ is a completely decomposable representation then we can write

$$V \cong \oplus_{\mu} \oplus_{i=1}^{a_\mu} V^{(\mu)} \tag{11.166}$$

where $V^{(\mu)}$ is the carrier space of an irreducible representation of $G$, we are summing over all irreps in (11.166), and $a_\mu$ is the number of times that irrep appears in the decomposition. (We understand that if a particular irrep does not appear at all then we take $a_\mu = 0$.) When $a_\mu \neq 0$ we can write

$$\oplus_{i=1}^{a_\mu} V^{(\mu)} \cong \kappa^{a_\mu} \otimes_{\kappa} V^{(\mu)} \tag{11.167}$$

as vector spaces where $\kappa$ is the ground field. We can view $\kappa^{a_\mu} \otimes_\kappa V^{(\mu)}$ as a tensor product of $G$-representations where $T(g)$ acts as trivially on the first factor so it is of the form $1 \otimes T^{(\mu)}(g)$. With this understood, the summand $\kappa^{a_\mu} \otimes_\kappa V^{(\mu)}$ is called the *isotypical component of* $V$ belonging to $\mu$. It is useful to abbreviate $\kappa^{a_\mu} \otimes V^{(\mu)}$ to $a_\mu V^{(\mu)}$. This has the nice advantage that if $a_\mu = 0$ then we have the zero vector space and we can just drop it from the direct sum. In this notation then the decomposition into isotypical components can be written as:

$$V = \oplus a_\mu V^{(\mu)} \tag{11.168}$$

where $a_\mu$ are nonnegative integers. Note that if we denote $n_\mu := \dim_\kappa V^{(\mu)}$ then

$$\dim_\kappa V = \sum_\mu a_\mu n_\mu \ . \tag{11.169}$$

In terms of bases and matrices we can rephrase the above as follows: $V$ admits a basis of the form $\{\psi_{\mu,i,\alpha}\}$ where $\mu$ runs over the distinct (isomorphism classes of) irreps. For each irrep $V^{(\mu)}$ we have chosen a basis $\{v_1, \dots, v_{n_\mu}\}$ and operators $T^{(\mu)}(g)$ on $V^{(\mu)}$ have matrices $M^\mu(g)_{ij}$ with respect to this basis. Finally, $\alpha = 1, \dots, a_\mu$ is a degeneracy index.

The transformation operators $T(g)$ on $V$ are block diagonal in the basis $\{\psi_{\mu,i,\alpha}\}$ and only act on the $i$-index according to:

$$T(g) : \psi_{\mu,i,\alpha} \mapsto \sum_{j=1}^{n_\mu} M^\mu(g)_{ji}\psi_{\mu,j,\alpha} \tag{11.170}$$

**Examples**

1. Suppose $T : U(1) \to U(N)$ is a representation by diagonal matrices so that $T(z)$ is a diagonal matrix with $a_n$ entries of $z^n$, where $n \in \mathbb{Z}$. Then

$$(T, V) \cong \oplus_{n\in\mathbb{Z}} a_n \rho_n \tag{11.171}$$

2. In general, a representation of $\mathbb{Z}_2$ on a vector space is the same thing as a linear operator $T : V \to V$ such that $T^2 = 1$. Given such a $T$ one can form orthogonal projection operators $P_\pm = \frac{1}{2}(1 \pm T)$. Then $P_\pm$ are diagonalizable, because, $P_+$ say has kernel $V_+ \subset V$ and then the image $P_+(V)$ is the $+1$ eigenspace. So the isotypical decomposition is

$$V \cong V_+ \otimes \rho_- \oplus V_- \otimes \rho_+ \tag{11.172}$$

So, for example, if $V = \mathbb{R}^2$ and $T = \sigma^1$ then $a_+ = +1$ and $a_- = 1$.

3. Using the representation theory for $SU(2)$ discussed below we can say the following: In the famous case of the Hydrogen atom for each principal quantum number $n$ there is a representation of $SU(2)$ which has isotypical decomposition

$$\oplus_{j=0}^{n-1} V_j \tag{11.173}$$

where $V_j$ is the spin $j$ irreducible representation of $SU(2)$ and we sum over integer spins.

### 11.8 Schur's Lemmas

A very important remark about equivariant maps between irreducible representations is known as *Schur's lemma.* It is almost a tautology - but it is a very powerful tautology.

> Schur's Lemma 1: Let $G$ be <u>any</u> group. Let $\kappa$ be <u>any</u> field. Let $V_1, V_2$ be vector spaces over $\kappa$ such that they are carrier spaces of irreducible representations of $G$. If $A : V_1 \to V_2$ is an intertwiner between these two irreps then $A$ is either zero or an isomorphism of representations.

*Proof*: Note that the kernel and image of $A$ are invariant subspaces of $V_1$ and $V_2$, respectively. Recall these are defined by:

$$\ker A := \{v_1 \in V_1 | A(v_1) = 0\} \tag{11.174}$$

$$\text{im}A := \{v_2 \in V_2 | \exists v_1 \in V_1 \qquad v_2 = A(v_1)\} \tag{11.175}$$

The reader should check that these are linear subspaces because $A$ is linear and they are invariant subspaces because $A$ is an intertwiner.

Now, since $V_1$ is an irrep we conclude that $\text{ker}A$ is either the 0 vector space or all of $V_1$. Similarly, since $V_2$ is an irrep $\text{im}A$ is either 0 or the entire space $V_2$. Now, if $\text{ker}A = V_1$ then $A = 0$. If $A \neq 0$ then $\text{ker}A \neq V_1$ so therefore $\text{ker}A = 0$, therefore $A$ is injective. Moreover if $A \neq 0$ then there is a nonzero vector in $\text{im}A$, and therefore $\text{im}A = V_2$, so $A$ is surjective. Therefore, $A$ is an isomorphism. ♠

Note that it is a general fact of linear algebra that if $V$ is a vector space over a field $\kappa$ then the set of $\kappa$-linear transformations $V \to V$, denoted $\text{Hom}_\kappa(V, V)$ or $\text{End}_\kappa(V)$ is not only a linear space over $\kappa$ but is also an algebra. It is a vector space because the sum of linear transformations is a linear transformation and the product with a scalar is a linear transformation compatible with the sum. It is an algebra because the composition of linear transformations is linear, in a way compatible with the sum and scalar multiplication. See LAUM for more on algebras. Everything we just said can also be said for the $G$-equivariant linear operators $V \to V$ if $V$ is a $G$-representation. These are denoted by $\text{Hom}_\kappa^G(V, V)$ or $\text{End}_\kappa^G(V)$. Schur's Lemma 1 says that if $V$ is an irrep then this algebra has the following property: If $A \in \text{End}_\kappa^G(V)$ is nonzero then it is invertible. In general, an algebra where all the nonzero elements are invertible is called a *division algebra*.

For Schur's second lemma the field $\kappa$ becomes important and we just state it for $\kappa = \mathbb{C}$:

Schur's Lemma 2: Suppose $(V, T)$ is an irreducible representation of a group $G$ on a complex vector space $V$ by $\mathbb{C}$-linear transformations. Suppose $A : V \to V$ is a $\mathbb{C}$-linear intertwiner, i.e., $A$ is a $\mathbb{C}$-linear operator that commutes with $T(g)$ for all $g \in G$. Then $A$ is proportional to the identity transformation: There exists a scalar $\lambda \in \mathbb{C}$ such that for all $v \in V$, $A(v) = \lambda v$. In other words, $\text{End}_{\mathbb{C}}^G(V) \cong \mathbb{C}$, the one-dimensional algebra of complex numbers.

*Proof*: Since we are working over the complex field and $A$ is $\mathbb{C}$-linear, the operator $A$ has a nonzero eigenvector $Av = \lambda v$. That follows because the characteristic polynomial $p_A(x) = \det(x1 - A)$ is a polynomial in the complex field and has a root in the complex numbers. The eigenspace $C = \{w : Aw = \lambda w\}$ is therefore not the zero vector space. But it is also an invariant subspace. Therefore, it must be the entire carrier space. ♠.

**Remarks**:

1. *Degeneracy Spaces As Spaces Of Intertwiners.* Let us return to the isotypical decomposition of a completely reducible representation, (11.166). Let $\text{Hom}^G(V_1, V_2)$ denote the vector space of $G$-equivariant maps between two $G$-spaces $V_1, V_2$. Note that

$$\text{Hom}^G(V_1, V_2 \oplus V_3) \cong \text{Hom}^G(V_1, V_2) \oplus \text{Hom}^G(V_1, V_3) \tag{11.176}$$

and if $G$ acts trivially on $V_2$, say, then

$$\mathrm{Hom}^G(V_1, V_2 \otimes V_3) \cong V_2 \otimes \mathrm{Hom}^G(V_1, V_3) \tag{11.177}$$

As an example, consider:

$$\mathrm{Hom}^G(\kappa^{a_\mu} \otimes V^{(\mu)}, \kappa^{a_\nu} \otimes V^{(\nu)}) \tag{11.178}$$

where we recall that, as a representation of $G$ the group operators on $\kappa^{a_\mu} \otimes V^{(\mu)}$ are of the form $1 \otimes T^{(\mu)}(g)$. Then, taking the ground field to be $\kappa = \mathbb{C}$ we can say

$$\begin{aligned}\mathrm{Hom}^G(\kappa^{a_\mu} \otimes V^{(\mu)}, \kappa^{a_\nu} \otimes V^{(\nu)}) &\cong \mathrm{Hom}^G(\kappa^{a_\mu}, \kappa^{a_\nu}) \otimes \mathrm{Hom}^G(V^{(\mu)}, V^{(\nu)}) \\ &\cong \mathrm{Hom}^G(\kappa^{a_\mu}, \kappa^{a_\nu}) \otimes \delta_{\mu,\nu}\mathbb{C} \\ &\cong \delta_{\mu,\nu}\mathrm{Hom}(\kappa^{a_\mu}, \kappa^{a_\nu}) \\ &\cong \delta_{\mu,\nu}\mathrm{Mat}_{a_\nu \times a_\mu}(\mathbb{C})\end{aligned} \tag{11.179}$$

For a general field we can say:

$$\begin{aligned}\mathrm{Hom}^G(V^{(\mu)}, V) &\cong \oplus_\nu \mathrm{Hom}^G(V^{(\mu)}, \kappa^{a_\nu} \otimes V^{(\nu)}) \\ &\cong \oplus_\nu \kappa^{a_\nu} \otimes \mathrm{Hom}^G(V^{(\mu)}, V^{(\nu)}) \\ &= \kappa^{a_\mu} \otimes \mathrm{Hom}^G(V^{(\mu)}, V^{(\mu)})\end{aligned} \tag{11.180}$$

In the second line we used the fact that $G$ acts trivially on $\kappa^{a_\nu}$. If we work over $\kappa = \mathbb{C}$ then we just showed $\mathrm{Hom}^G(V^{(\mu)}, V^{(\mu)}) \cong \mathbb{C}$ and hence we have a better interpretation of the degeneracy space: It is the linear space of $G$-invariant maps from $V^{(\mu)} \to V$. Indeed, note that there is a canonical equivariant map:

$$\mathrm{Hom}^G(V^{(\mu)}, V) \otimes V^{(\mu)} \to V \tag{11.181}$$

given by $A \otimes u \mapsto A(u)$. So we have a canonical $G$-equivariant map

$$\oplus_\mu \mathrm{Hom}_G(V^{(\mu)}, V) \otimes V^{(\mu)} \to V \tag{11.182}$$

and complete reducibility is the statement that this is an isomorphism.

2. *Block diagonalization of Hamiltonians.* Suppose $\mathcal{H}$ is a physical Hilbert space which is a representation of a group $G$ so that $\mathcal{H}$ is completely reducible into isotypical components:

$$\mathcal{H} \cong \oplus_\mu \mathcal{H}^{(\mu)} \tag{11.183}$$

That is $\mathcal{H}^{(\mu)} \cong D_\mu \otimes V^{(\mu)}$ where $V^{(\mu)}$ is an irreducible representation and $D_\mu$ is the *degeneracy space* (isomorphic to $\kappa^{a_\mu}$ in terms of our previous discussion). The sum in (11.183) is over the distinct irreps of $G$. Suppose now that $H$ is a $\mathbb{C}$-linear operator that commutes with the $G$-action. A typical application is where $H$ is a Hamiltonian, so having a $G$-symmetry of the dynamics means that

$$H : \mathcal{H} \to \mathcal{H} \tag{11.184}$$

is an intertwiner. By Schur's lemma we therefore have under this isomorphism

$$H \cong \oplus_\mu h^{(\mu)} \otimes 1_{V(\mu)} \tag{11.185}$$

In terms a basis compatible with the isotypical decomposition it means that $h^{(\mu)}$ acts only on the degeneracy space, and hence the Hamiltonian has been partially (or sometimes completely) diagonalized. Group theory alone cannot provide any more information about the nature of the Hermitian operator $h^{(\mu)}$ on the degeneracy space.

As we have discussed above, in terms of bases the isotypical decomposition implies there is a basis $\{\psi_{\mu,i,\alpha}\}$ for $\mathcal{H}$ where $\mu$ runs over the distinct (isomorphism classes of) irreps, $i = 1, \dots, n_\mu$ are indices transforming under $G$, and $\alpha = 1, \dots, a_\mu$ labels the degeneracy of the representation. With respect to such a basis an intertwiner has matrix elements of the form:

$$H_{(\mu_1,i_1,\alpha_1),(\mu_2,i_2,\alpha_2)} = \delta_{\mu_1,\mu_2}\delta_{i_1,i_2}h^{\mu}_{\alpha_1,\alpha_2} \tag{11.186}$$

Here the $\delta_{\mu_1,\mu_2}$ factor follows from Schur's Lemma 1, the $\delta_{i_1,i_2}$ factor follows from Schur's Lemma 2, and $h^{\mu}_{\alpha_1,\alpha_2}$ is an arbitrary matrix in the degeneracy indices. The group theory of $G$ gives no information on $h^{\mu}_{\alpha_1,\alpha_2}$. It is an arbitrary $a_\mu \times a_\mu$ matrix which must be determined by other means in a given physical problem.

3. *Selection Rules.* The above observations also leads also lead to *selection rules.* If $\psi_1, \psi_2$ are two states in different isotypical components and $\mathcal{O}$ is an operator that commutes with the $G$-action then the transition amplitude

$$\langle \psi_1, \mathcal{O}\psi_2 \rangle = 0 \tag{11.187}$$

The reason is that if $\psi_1 \in \mathcal{H}^{(\mu)}$ and $\psi_2 \in \mathcal{H}^{(\nu)}$ with $\mu \neq \nu$ and $\mathcal{O}$ is an intertwiner then $\mathcal{O} = \oplus_\mu \mathcal{O}^{(\mu)}$ where $\mathcal{O}^{(\mu)} \in \mathrm{End}(D_\mu) \otimes 1_{V(\mu)}$ and in particular preserves the isotypical component.

As a very easy example of this general statement suppose that $\mathcal{H}$ is a Hilbert space with a $\mathbb{Z}_2$ action. As we saw, this is equivalent to an operator $T : \mathcal{H} \to \mathcal{H}$ so that $T^2 = 1$. Then we can write the isotypical decomposition

$$\begin{aligned} \mathcal{H} &= \mathcal{H}_+ \oplus \mathcal{H}_- \\ &= D_+ \otimes \rho_+ \oplus D_- \otimes \rho_- \end{aligned} \tag{11.188}$$

where $T = +1$ on $\mathcal{H}_+$ and $T = -1$ on $\mathcal{H}_-$. If $\psi_+ \in \mathcal{H}_+$ and $\psi_- \in \mathcal{H}_-$ then $\langle \psi_+, \psi_- \rangle = 0$. One (of many) ways to prove this is to note that

$$\langle \psi_+, \psi_- \rangle = \langle \psi_+, T^2\psi_- \rangle = \langle T\psi_+, T\psi_- \rangle = -\langle \psi_+, \psi_- \rangle \tag{11.189}$$

If $T\mathcal{O} = \mathcal{O}T$ then since $T$ preserves $\mathcal{H}_+$ and $\mathcal{H}_-$ the selection rule follows.

4. *Algebra Of Intertwiners For $\kappa = \mathbb{R}$.* Schur's lemma over other fields can lead to more complicated possibilities. As we noted above, it follows from Schur's Lemma

1 that $\mathrm{Hom}_{\kappa}^{G}(V^{(\mu)}, V^{(\mu)})$ is a *division algebra* over $\kappa$, meaning that it is an algebra in which all nonzero elements are invertible. When we take $\kappa = \mathbb{R}$ we therefore get a division algebra over $\mathbb{R}$. A nice mathematical theorem (see LAUM) states that there are only three division algebras over the real numbers: $\mathbb{R}, \mathbb{C}, \mathbb{H}$. As an example of the case $\mathbb{R}$ just consider the real linear transformations between the irreducible real representations of $\mathbb{Z}_2$. Here the carrier space is $\mathbb{R}$ and any linear transformation from $\mathbb{R}$ to itself commutes. The space of linear transformations is canonically: $\mathrm{Hom}(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}$, meaning that the general linear transformation is just multiplication by a real number. So $\mathrm{Hom}_{\mathbb{R}}^{\mathbb{Z}_2}(\rho_\mu, \rho_\nu) \cong \delta_{\mu,\nu}\mathbb{R}$ where $\mu, \nu \in \{\pm\}$.

Similarly, if we take $G = U(1)$ the irreducible representations are one complex dimensional and $V \cong \mathbb{C}$. The irreps are $\rho_n(z) = z^n$. We can view $\mathbb{C}$ as a two-dimensional real vector space, and the <u>real</u> linear transformations from $\mathbb{C} \to \mathbb{C}$ involve complex conjugation. The general real linear transformation would act on the "vector" $z \in \mathbb{C}$ as

$$\begin{pmatrix} Re(z) \\ Im(z) \end{pmatrix} \to \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} Re(z) \\ Im(z) \end{pmatrix} \tag{11.190}$$

If we use a general $2 \times 2$ real matrix and re-express the transformation in terms of $z$ we find that the transformation would take $z$ to a real linear combination of $z$ and $\bar{z}$ and therefore would not commute with multiplication by $e^{\mathrm{i}n\theta}$ which is the group operator $T_n(e^{\mathrm{i}\theta})$ acting on $z$. Only the $\mathbb{C}$-linear ones can be intertwiners. From this it follows that $\mathrm{Hom}_{\mathbb{R}}^{U(1)}(\rho_n, \rho_m) \cong \delta_{n,m}\mathbb{C}$.

Now, to give an example where the algebra of real intertwiners is $\mathbb{H}$ we turn to the fundamental representation of $SU(2)$. This is of course the $SU(2)$ matrix just acting on a vector in $\mathbb{C}^2$. Again, by writing all complex numbers in terms of their real and imaginary parts, e.g. by writing

$$\begin{pmatrix} z \\ w \end{pmatrix} \to \begin{pmatrix} Re(z) \\ Im(z) \\ Re(w) \\ Im(w) \end{pmatrix} \in \mathbb{R}^4 \tag{11.191}$$

we can express the fundamental representation of $SU(2)$ in terms of $4 \times 4$ real matrices acting on vectors in $\mathbb{R}^4$. When we compute $\mathrm{Hom}_{\mathbb{R}}^{SU(2)}(V, V)$ with $V$ the fundmaental representation we are computing the algebra of $\mathbb{R}$-linear operators that commute with all the group operators. We can find this algebra of operators very nicely using the concept of the quaternions. See LAUM for a general discussion of quaternions. For our remarks here we take a low-brow viewpoint and identify the quaternions, denoted $\mathbb{H}$, with the set of $2 \times 2$ complex matrices $q$ of the form

$$q = \begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix} \tag{11.192}$$

with $u, v \in \mathbb{C}$. That is, for us, we define:

$$\mathbb{H} := \{\begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix} |u, v \in \mathbb{C}\} \subset \mathrm{Mat}_{2\times 2}(\mathbb{C}) \tag{11.193}$$

Note that $\mathbb{H}$ is a <u>real</u> vector space: If we add such matrices or multiply them by real numbers we get another matrix of the same general form. However, it is not naturally a complex vector space: If we multiply a matrix of this form by $\sqrt{-1}$ we do <u>not</u> obtain a matrix of the same general form. Note that $\mathbb{H}$ is also an algebra: If we multiply two matrices of this form we get another one. You can check that directly, by matrix multiplication, or you can note that the general matrix $q \in \mathbb{H}$ can be written as

$$q = x_4 1 + \sum_{k=1}^{3} x_k \mathrm{i}\sigma^k \tag{11.194}$$

where $\sigma^k$ are the Pauli matrices. Now recall that

$$\sigma^k \sigma^\ell = \delta^{k,\ell} + \mathrm{i}\epsilon^{k\ell m}\sigma^m \tag{11.195}$$

and using this one shows that the product of two elements of $\mathbb{H}$ is an element of $\mathbb{H}$, from which one can deduce that it is an algebra. Note well that $SU(2) \subset \mathbb{H}$ consists of the elements of determinant one, or equivalently, such that $|u|^2 + |v|^2 = 1$.

Now note that the first column of a matrix in $\mathbb{H}$ is just a vector in $\mathbb{C}^2$, and the second column is determined from the first. So, from a vector in $\mathbb{C}^2$ we uniquely determine a matrix in $\mathbb{H}$, and vice versa. Now the left action of $SU(2) \subset \mathbb{H}$ on $\mathbb{H}$ preserves $\mathbb{H}$. Note that the left action of an $SU(2)$ matrix on a vector in $\mathbb{C}^2$ is equivalent to the left action of a matrix on the corresponding element of $\mathbb{H}$. So we can think of the fundamental reprsentation of $SU(2)$ in terms of the left action of $SU(2)$ on $\mathbb{H}$. So stated this way, the action of

$$g = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \tag{11.196}$$

on

$$\begin{pmatrix} z \\ w \end{pmatrix} \in \mathbb{C}^2 \tag{11.197}$$

is expressed as the transformation

$$\begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix} \mapsto \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix} = \begin{pmatrix} \tilde{z} & -\bar{\tilde{w}} \\ \tilde{w} & \bar{\tilde{z}} \end{pmatrix} \tag{11.198}$$

Now fix some quaternion $q_0$. Consider the transformation $A_{q_0} : \mathbb{H} \to \mathbb{H}$ defined by

$$A_{q_0} : q \in \mathbb{H} \mapsto q q_0 \in \mathbb{H} \tag{11.199}$$

Again this makes sense since $\mathbb{H}$ is an algebra. Moreover it is a real linear transformation of the real vector space $\mathbb{H}$: Note that $(\alpha_1 q_1 + \alpha_2 q_2)q_0 \mapsto \alpha_1 q_1 q_0 + \alpha_2 q_2 q_0$. We can express this transformation as a transformation of $2 \times 2$ complex matrices:

$$A_{q_0} : \begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix} \mapsto \begin{pmatrix} z & -\bar{w} \\ w & \bar{z} \end{pmatrix} \begin{pmatrix} u_0 & -\bar{v}_0 \\ v_0 & \bar{u}_0 \end{pmatrix} := \begin{pmatrix} \tilde{z} & -\overline{\tilde{w}} \\ \tilde{w} & \overline{\tilde{z}} \end{pmatrix} \tag{11.200}$$

In terms of the first column this is the transformation

$$A_{q_0} : \begin{pmatrix} z \\ w \end{pmatrix} \mapsto \begin{pmatrix} u_0 z - v_0 \bar{w} \\ v_0 \bar{z} + u_0 w \end{pmatrix} \tag{11.201}$$

If $v_0 \neq 0$ then this transformation is <u>not</u> $\mathbb{C}$-linear in $z, w$. However, if we write this as a transformation on $\mathbb{R}^4$ using (11.191) then $A_{q_0}$ <u>is</u> $\mathbb{R}$-linear. Moreover, the transformations of type (11.198) commute with transformations of type (11.199) or equivalently of type (11.200) because one acts with (unit) quaternions on the left and the other acts with (arbitrary) quaternions on the right. Therefore $A_{q_0}$ are $\mathbb{R}$-linear intertwiners. In fact, it is not hard to see that these are the most general $\mathbb{R}$-linear intertwiners. So we learn that, with $V$ the fundamental representation of $SU(2)$ we have the real algebra of intertwiners:

$$\mathrm{Hom}_{\mathbb{R}}^{SU(2)}(V, V) \cong \mathbb{H} \ . \tag{11.202}$$

---

**Exercise** *What's Wrong With This Picture?*

Here is an alleged proof that the only nontrivial irreducible representation of $SU(2)$ is the fundamental representation. Where is the wrong step? [163]

a.) Suppose $V$ is a unitary finite-dimensional representation of $SU(2)$. Choose any nonzero vector $v \in V$.

b.) Consider the $SU(2)$ elements $\mathrm{i}\sigma^1, \mathrm{i}\sigma^2, \mathrm{i}\sigma^3$. They act by operators $I, J, K$ on $V$ so consider the vectors $Iv, Jv, Kv$.

c.) Every element of $SU(2)$ is of the form $x_4 1 + x_k \mathrm{i}\sigma^k$ with $x_\mu x_\mu = 1$.

d.) Therefore the action of $SU(2)$ preserves the linear subspace $W \subset V$ generated by $v, Iv, Jv, Kv$. But this is just the fundamental representation!

e.) Now since $W$ is an invariant subspace so is $W^\perp$. Proceed by induction to find that $V$ is just a direct sum of fundamental representations.

---

[163] *Answer*: The wrong step is (d). It assumes that $T(x_4 1 + x_k \mathrm{i}\sigma^k) = x_4 T(1) + x_k T(\mathrm{i}\sigma^k)$. That is not the case for the general representation of $SU(2)$. (Just consider the tensor product of the fundamental with itself.) However, this <u>is</u> a valid argument for the general irreducible representation of the group algebra, which is just the algebra $\mathbb{H}$.

### 11.9 Pontryagin Duality

In this section we introduce the beautiful idea of the *Pontryagin dual* of an Abelian group. [164] We will use it in section 11.17.2, and again we will use it to give a very general construction of interesting Heisenberg extensions of Abelian groups.

**Definition**: Let $S$ be an Abelian group. The *Pontryagin dual* group $\widehat{S}$ is defined to be the group of homomorphisms $\mathrm{Hom}(S, U(1))$. Note that if $\chi_1, \chi_2 \in \mathrm{Hom}(S, U(1))$ then the product $\chi_1 \cdot \chi_2$ is defined to be

$$(\chi_1 \cdot \chi_2)(s) := \chi_1(s)\chi_2(s) \tag{11.203}$$

So $\chi_1 \cdot \chi_2$ is also a homomorphism $S \to U(1)$. This product makes $\widehat{S}$ into an Abelian group.

**Remarks**:

1. The Pontryagin dual group $\widehat{S}$ can also be thought of as the group of all complex one-dimensional unitary representations of $S$. It follows from Schur's lemma that all irreducible finite dimensional complex representations of an Abelian group are one-dimensional.

   Note that the adjective complex is essential here. After all the defining representation of the Abelian group $SO(2)$ is $\mathbb{R}^2$ and is irreducible as a representation over $\mathbb{R}$.

2. Elements of the group $\widehat{S}$ are also called characters.

3. It is best to discuss Pontryagin duality in the context of topological groups. In this case we should only consider the continuous characters $\chi : S \to U(1)$. For the duality theorem below we should consider *locally compact Abelian groups*. Examples of locally compact Abelian groups are $\mathbb{R}^n$, tori, lattices, and finite Abelian groups with compact topology. An infinite-dimensional Hilbert space is a topological Abelian group under addition, but it is not locally compact.

Note that, for a fixed $s \in S$ we can define a homomorphism $\mathrm{Hom}(\widehat{S}, U(1))$ by

$$\hat{s} : \chi \mapsto \chi(s) \tag{11.204}$$

Note that $\hat{s}$ is just the evaluation map $\mathrm{ev}_s$ discussed previously. The map $s \mapsto \hat{s}$ is a homomorphism $S \to \widehat{\widehat{S}}$. The main theorem is:

**Theorem**[Pontryagin-van Kampen duality]. If $S$ is a locally compact Abelian group then the canonical homomorphism $S \to \widehat{\widehat{S}}$ is in fact an isomorphism:

$$\widehat{\widehat{S}} \cong S \tag{11.205}$$

[164] The transliteration from the Cyrillic to the Latin alphabets takes various forms. Another common one is Pontrjagin.

For a proof see, for example, the book on representation theory by A.A. Kirillov.

**Example 1**: Consider $S = \mathbb{Z}/n\mathbb{Z}$, thought of additively. To determine $\chi \in \mathrm{Hom}(S, U(1))$ it suffices to determine $\chi(\bar{1})$, since $\chi(\bar{\ell}) = \chi(\bar{1})^{\ell}$ for any $\ell \in \mathbb{Z}$. Put $\chi(\bar{1}) = \omega \in U(1)$. But now we need to impose the relation $\chi(\bar{n}) = \chi(\bar{0}) = 1$. This implies $\omega^n = 1$, so $\omega$ is an $n^{th}$ root of unity. So the most general element of $\widehat{\mathbb{Z}/n\mathbb{Z}}$ is

$$\chi(\bar{\ell}) = \chi_{\omega}(\bar{\ell}) := \omega^{\ell} \tag{11.206}$$

where $\omega$ is an $n^{th}$ root of unity. Moreover $\chi_{\omega_1}\chi_{\omega_2} = \chi_{\omega_1\omega_2}$ so $\widehat{\mathbb{Z}/n\mathbb{Z}}$ is identified in this way with the multiplicative group of $n^{th}$ roots of unity $\mu_n$. In this way we see that, as abstract groups

$$\widehat{\mathbb{Z}/n\mathbb{Z}} = \mu_n \cong \mathbb{Z}/n\mathbb{Z} \tag{11.207}$$

So a finite cyclic group is self-dual.

**Example 2**: Consider $\mathbb{R}$, additively. Then if $\chi \in \widehat{\mathbb{R}}$ we have $\chi(x+y) = \chi(x)\chi(y)$ so $\chi(x) = e^{ax}$ for some constant $a$. For $\chi$ to be valued in $U(1)$ we must have $a = \mathrm{i}k$ with $k \in \mathbb{R}$ and hence

$$\chi(x) = \chi_k(x) := e^{\mathrm{i}kx} \tag{11.208}$$

moreover,

$$\chi_k\chi_\ell = \chi_{k+\ell} \tag{11.209}$$

and hence $\widehat{\mathbb{R}} \cong \mathbb{R}$. In an entirely similar way $\widehat{\mathbb{R}^n} \cong \mathbb{R}^n$

**Example 3**: Consider $S = \mathbb{Z}$. To determine $\chi \in \mathrm{Hom}(S, U(1))$ it suffices to determine $\chi(1)$. Choose any phase $\xi \in U(1)$ and set $\chi(1) = \xi$. Then it must be that, for all $n \in \mathbb{Z}$:

$$\chi(n) = \chi_{\xi}(n) := \xi^n \tag{11.210}$$

Moreover $\chi_{\xi_1}\chi_{\xi_2} = \chi_{\xi_1\xi_2}$. Thus,

$$\widehat{\mathbb{Z}} \cong U(1) \tag{11.211}$$

**Example 4**: Consider $S = U(1)$. To determine $\chi \in \mathrm{Hom}(S, U(1))$ it might help to think of $U(1) \cong \mathbb{R}/\mathbb{Z}$. We know from the Pontryagin dual of $\mathbb{R}$ that $\chi$ should be of the form

$$\chi(x+\mathbb{Z}) = \exp[\mathrm{i}k(x+\mathbb{Z})] \tag{11.212}$$

for some real number $k$. However, for this to be well-defined we must have $k = 2\pi n$ with $n \in \mathbb{Z}$. Therefore $\chi$ must be of the form

$$\chi_n(x+\mathbb{Z}) = \exp[2\pi \mathrm{i} n x] \tag{11.213}$$

Or, if we think of $S = U(1)$ multiplicatively as complex numbers of modulus one, then we can say that every character on $U(1)$ is of the form:

$$\chi_n(\xi) := \xi^n \tag{11.214}$$

for some $n \in \mathbb{Z}$. Therefore,

$$\widehat{U(1)} \cong \mathbb{Z} \tag{11.215}$$

Comparing (11.211) and (11.215) we verify the general result (11.205).

**Example 5**: *Tori.* Consider the group $G = \mathbb{Z}^d$. It will be useful to consider a free $G$ action on affine Euclidean space $\mathbb{E}^d$. This defines a subset $\Gamma \subset \mathbb{E}^d$ known as a *lattice* (sometimes called an *embedded lattice*). The quotient space $\mathbb{E}^d/\Gamma$ has a natural basepoint, namely the coset of $\Gamma$ and, as a group it is isomorphic to $U(1)^d$. By the same arguments as above its Pontryagin dual will be isomorphic to $\mathbb{Z}^d$.

There is a nice way to think about the Pontryagin duality between lattices and tori. Suppose $\Gamma$ is a lattice in $\mathbb{R}^d$. Using the Euclidean norm we can define another lattice, the *dual lattice*

♣Really we should do this using the dual vector space... ♣

$$\Gamma^\vee = \{g \in \mathbb{R}^d | g \cdot \gamma \in \mathbb{Z} \qquad \forall \gamma \in \Gamma\} \tag{11.216}$$

Note we can identify $\Gamma^\vee \cong \mathrm{Hom}(\Gamma, \mathbb{Z})$ since any $g \in \Gamma^\vee$ defines a homomorphism $\phi_g$ whose values are: $\phi_g(\gamma) = g \cdot \gamma$. Also, as an abstract Abelian group of course $\Gamma^\vee \cong \mathbb{Z}^d$, but the above definition identifies it as a specific subgroup of $\mathbb{R}^d$.

The unitary irreps of $\Gamma$ are represented by points in the torus $T^\vee := \mathbb{R}^d/\Gamma^\vee$. Note that $T^\vee$ is a torus, as a manifold, and is isomorphic to the group $U(1)^d$, as an Abelian group. For any $\bar{k} \in T^\vee$ we can write a formula for the character $\chi_{\bar{k}}$ of the form:

$$\chi_{\bar{k}}(\gamma) = \exp[2\pi \mathrm{i} k \cdot \gamma] \tag{11.217}$$

where $k$ is any representative of $\bar{k}$, that is $\bar{k} = k + \Gamma^\vee$. We call any vector $k \in \mathbb{R}^d$ which projects to $\bar{k}$ under the projection $\pi : \mathbb{R}^d \to T^\vee$ a <u>lift</u> of $\bar{k}$. Note that the above formula is well-defined because if we choose any two lifts $k_1$ and $k_2$ of $\bar{k}$ then $k_1 = k_2 + g$ with $g \in \Gamma^\vee$, and then $g \cdot \gamma \in \mathbb{Z}$ for all $\gamma \in \Gamma$. So

$$\widehat{\Gamma} \cong \mathbb{R}^d/\Gamma^\vee \cong U(1)^d \tag{11.218}$$

Conversely, the Pontryagin dual of the torus $\mathbb{R}^d/\Gamma^\vee$ can naturally be identified with $\Gamma$ by the same formula:

$$\chi_\gamma(\bar{k}) = \exp[2\pi \mathrm{i} k \cdot \gamma] \tag{11.219}$$

---

**Exercise** *Pontryagin Dual Of The Prüfer Groups*

Recall that the Prüfer groups $Pr(p)$ are defined for each prime $p$ as the union over all $n$ of roots of unity of order $p^n$.

What is the Pontryagin dual of $Pr(p)$? (Give it the discrete topology.) [165]

[165] *Answer*: Use the isomorphism $U(1) \cong \mathbb{R}/\mathbb{Z}$. One needs to say what is the image of $p^{-n}$. So $\chi(\frac{1}{p^n}) = \exp[2\pi \mathrm{i} a_n/p^n]$, for some integer $a_n$ because $\chi(p^{-n})$ must itself be a $(p^n)^{th}$ root of unity. Note we can regard $a_n \in \mathbb{Z}/p^n\mathbb{Z}$. Now note that

$$\exp[2\pi \mathrm{i}\frac{a_{n-1}}{p^{n-1}}] = \chi(\frac{1}{p^{n-1}}) = \chi(\frac{1}{p^n})^p = \exp[2\pi \mathrm{i}\frac{a_n}{p^{n-1}}] \tag{11.220}$$

So the Pontryagin dual is the subgroup of $\prod_n \mathbb{Z}/p^n\mathbb{Z}$ consisting of sequences $a_n$ so that $a_n$ projects to $a_{n-1}$. This is known as the group of $p$-adic integers.

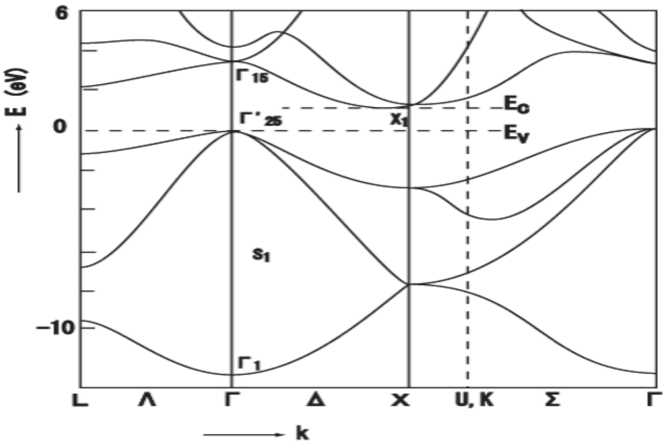



**Figure 31:** Example of a bandstructure. (For silicon.) On the horizontal axis the structure is plotted as a function of $k$ along lines inside the Brillouin torus. The letters refer to points where the (cubic) crystallographic group has fixed points. $\Gamma$ denotes the identity element $\bar{k} = 0$ where the full cubic symmetry group is restored.

### 11.9.1 An Application Of Pontryagin Duality: Bloch's Theorem And Band Structure

The Pontryagin duality between an embedded lattice $\Gamma \subset \mathbb{R}^d$ and the torus $T^\vee = \mathbb{R}^d/\Gamma^\vee$ has a very significant application in condensed matter physics known as *Bloch's theorem.*

In the one-electron approximation to the Schrodinger problem of electrons in a crystal one considers the Schrödinger Hamiltonian on $L^2(\mathbb{R}^d)$ of the form:

$$H = -\frac{\hbar^2}{2m}\nabla^2 + U(x) \tag{11.221}$$

where the potential $U : \mathbb{R}^d \to \mathbb{R}$ is assumed to be invariant under some crystallographic group (see below). In particular, it is invariant under translation by a lattice $\Gamma \subset \mathbb{R}^d$. For example, if we take into account the Coulomb interaction between the electron and a collection of ions of charge $Z_i e$ at positions $x_i \in C$, where $C$ is a crystal then

$$U(x) = \sum_i \frac{-Z_i e^2}{|x - x_i|} \tag{11.222}$$

But for the statement we are going to make all we need is that $U(x+\gamma) = U(x)$ for $\gamma \in \Gamma$.

Now, the Hilbert space $\mathcal{H}$ is a unitary representation of the translation group on $\mathbb{R}^d$. Translation by $a \in \mathbb{R}^d$ is represented by

$$\rho(a) = \exp[\mathrm{i}a \cdot \hat{p}/\hbar] \tag{11.223}$$

where $\hat{p} = -\mathrm{i}\hbar\vec{\nabla}$ as is standard in quantum mechanics. When considering a crystal with the above Hamiltonian only the subgroup $\Gamma \subset \mathbb{R}^d$ commutes with $H$. So we consider $\mathcal{H}$ as a unitary representation of $\Gamma$ with operators:

$$\rho(\gamma) = \exp[\mathrm{i}\gamma \cdot \hat{p}/\hbar] \tag{11.224}$$

where $\hat{p} = -\mathrm{i}\hbar\vec{\nabla}$ as is standard in quantum mechanics. Note that

$$\rho(\gamma_1)\rho(\gamma_2) = \rho(\gamma_1 + \gamma_2) \tag{11.225}$$

We have classified the one-dimensional representations above so if $\psi \in \mathcal{H}$ were to be in a one-dimensional representation then it would have to be quasi-periodic:

$$\psi(x + \gamma) = \chi_{\bar{k}}(\gamma)\psi(x) \tag{11.226}$$

where $\bar{k} \in \mathbb{R}^d/\Gamma^\vee$. In this context the Pontryagin dual torus $T^\vee$ is known as the *Brillouin torus*.

The notion of isotypical decomposition described above does extend to some noncompact groups with continuous families of irreducible representations, such as lattices. The isotypical decomposition of the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d)$ is thus expected to be of the form of an integral over the Brillouin torus:

$$\mathcal{H} \cong \oint_{T^\vee} d\bar{k}\mathcal{H}_{\bar{k}} \tag{11.227}$$

Here $\mathcal{H}_{\bar{k}}$ is a Hilbert space of wave functions satisfying the quasiperiodicity requirement (11.226). The Brillouin torus is a Riemannian manifold and inherits a natural measure. One can make sense of a continuous sum of Hilbert spaces.

There is an important technical complication with (11.227). Note that a quasiperiodic function (11.226) cannot actually be in $L^2(\mathbb{R}^d)$. The square of the wavefunction: $|\psi(x)|^2$ is a $\Gamma$-periodic function of $x$. Therefore $|\psi(x)|^2$ descends to a well-defined function on the real-space torus $T := \mathbb{R}^d/\Gamma$. The integral over a fundamental domain for $\Gamma$ will be some positive number, which is then multiplied by the infinite number of fundamental domains when integrating over $\mathbb{R}^d$. It follows that the integral over all of $\mathbb{R}^d$ is infinite. This is similar to the technical point that plane-waves on $\mathbb{R}^d$ - which are formally an ON basis - are not actually $L^2$ normalizable. Nevertheless, they are very useful formal tools for describing analysis in Hilbert space, and honest $L^2$-normalizable wavefunctions are made from "wavepackets" that are linear combinations of planewaves. In the present case the correct thing to do is this: For each point $\bar{k}$ in the Brillouin torus define the Hilbert space

$$\mathcal{H}_{\bar{k}} := \{\psi(x)|\psi(x+\gamma) = \chi_{\bar{k}}(\gamma)\psi(x) \qquad \int_T |\psi(x)|^2 dx < \infty\} \ . \tag{11.228}$$

Then when writing vectors in the space (11.227) on the RHS we should only consider sums ("wavepackets") so that the resulting function is in $L^2(\mathbb{R}^d)$.

Now the Hamiltonian cannot make transitions between different irreps in an isotypical decomposition, so we can reduce the eigenvalue problem to that of $H$ acting on $\mathcal{H}_{\bar{k}}$. Note that if $\psi$ obeys (11.226) then we can always write (noncanonically!!!)

$$\psi(x) = e^{2\pi\mathrm{i}k\cdot x}u_k(x) \tag{11.229}$$

where we have chosen a specific representative $k \in \mathbb{R}^d$ of the element $\bar{k}$ in the Brillouin torus and $u_k(x)$ is periodic in $x$, i.e. invariant under shifts of $x \to x + \gamma$ for $\gamma \in \Gamma$. In condensed matter physics the vector $k \in \mathbb{R}^d$ is known as a *reciprocal vector*. We stress that this decomposition is noncanonical. If $k_1$ and $k_2$ both represent the same $\bar{k}$ then $k_1 = k_2 + g$ for $g \in \Gamma^\vee$ and $u_{k_2}(x) = e^{2\pi \mathrm{i} g \cdot x} u_{k_1}(x)$. Note that both functions $u_{k_2}$ and $u_{k_1}$ are periodic under shifts by $\Gamma$. Therefore, they can both be considered as functions on the *real space torus* $T := \mathbb{R}^d/\Gamma$.

It is useful to write the eigenvalue problem as

$$H_k u_k(x) = E_k u_k(x) \tag{11.230}$$

where

$$H_k = e^{-2\pi \mathrm{i} k\cdot x} H e^{2\pi \mathrm{i} k \cdot x} \tag{11.231}$$

Note that here we had to make a definite choice of $k$ that projects to $\bar{k}$ in order to write the Hamiltonian $H_k$. However, if $k' = k + g$ where $g \in \Gamma^\vee$ then $H_{k'}$ is unitarily equivalent to $H_k$. Indeed $U = e^{2\pi \mathrm{i}\hat{x}\cdot g}$ is a nice unitary operator on the wavefunctions on the torus $T = \mathbb{R}^d/\Gamma$ that conjugates $H_k$ to $H_{k'}$, and therefore the spectrum of $H_k$ and $H_{k'}$ is the same.

$H_k$ is an Hermitian elliptic operator acting on the functions on a compact manifold. Explicitly, it works out to

$$H_k = -\frac{\hbar^2}{2m}\nabla^2 - 4\pi \frac{\hbar^2}{2m} k \cdot (i\nabla) + (U + \frac{\hbar}{2m} 4\pi^2 k^2) \tag{11.232}$$

and it acts on $L^2$ functions on the torus $T = \mathbb{R}^d/\Gamma$. This operator should be viewed as a perturbation of a Laplace operator on functions on a compact manifold. The latter has a discrete spectrum of eigenvalues. The spectrum of $h = -\nabla^2$ is $\{4\pi^2 g^2\}_{g\in\Gamma^\vee}$. The theory of elliptic Hermitian operators on compact manifolds shows that the lower order terms in (11.232) do not change this property and hence the operator (11.232) has a discrete set of eigenvalues $\{E_n(k)\}$. Note that while $H_k$ depends on $k$, the full spectrum of $H_k$ only depends on $\bar{k}$. This is not obvious from (11.232) but it follows from the unitary equvalence between $H_{k_1}$ and $H_{k_2}$ where $k_1$ and $k_2$ are two representatives of $\bar{k}$.

The eigenvalues vary continuously as functions of $\bar{k} \in \mathbb{R}^d/\Gamma^\vee$ to give what is called a *band structure*. See Figure 31.

## 11.10 Orthogonality Relations Of Matrix Elements And The Peter-Weyl Theorem

The very beautiful *Peter-Weyl theorem* states that, if $G$ is a compact group, then there is an isomorphism of $G \times G$ representations:

$$L^2(G) \cong \oplus_\mu \mathrm{End}(V^{(\mu)}) \tag{11.233}$$

where we sum over the isomorphism class of each irreducible representation exactly once, and for each irrep we choose a representative $(T^{(\mu)}, V^{(\mu)})$.

♣Notation is too heavy: Replace $V^{(\mu)}$ by $V^\mu$ etc. below. ♣

A useful preliminary result is the following:

**Proposition** Let $(T, V)$ be a unitary irreducible representation of a compact group $G$ on a complex vector space $V$. Then $V$ is finite dimensional.

*Proof*: Choose a nonzero vector $v \in V$ and define the operator $L : V \to V$ by saying that for $w \in V$

$$L(w) := \int_G T(g)v\langle T(g)v, w\rangle dg \tag{11.234}$$

Then a short computation (do it!) using the left-invariance of the measure $dg$ shows that

$$L(w)T(g_0) = T(g_0)L(w) \tag{11.235}$$

for all $g_0 \in G$. That is, $L$ is a self-intertwiner, and therefore, by Schur's lemma $L = \lambda Id$ where $Id$ is the identity operator. In fact, by taking the inner product of $L(v)$ with $v$ we get a formula for $\lambda$:

$$\lambda \parallel v \parallel^2 = \int_G |\langle v, T(g)v\rangle|^2 dg \tag{11.236}$$

Note that on the RHS, $|\langle v, T(g)v\rangle|^2$ is a continuous function on a compact space and hence is bounded as a function of $g$, and since $G$ has finite volume the RHS is finite.

Now, since $L = \lambda Id$ we have $Tr(L) = \lambda \dim V$. So we need only show that $\mathrm{Tr}(L) < \infty$. To do this, introduce an ON basis $\{v_i\}$ for $V$ and note that

$$\begin{aligned}
\mathrm{Tr}(L) &= \sum_i \langle v_i, L(v_i)\rangle \\
&= \sum_i \int_G \langle v_i, T(g)v\rangle\langle T(g)v, v_i\rangle dg \\
&= \sum_i \int_G |\langle v_i, T(g)v\rangle|^2 dg \\
&= \int_G \sum_i |\langle v_i, T(g)v\rangle|^2 dg \\
&= \int_G \parallel T(g)v \parallel^2 dg \\
&= \parallel v \parallel^2 \mathrm{vol}\,(G)
\end{aligned} \tag{11.237}$$

where in the last line we used that $T(g)$ is a unitary operator ♠

As a little corollary of the previous proof we actually get a formula for the dimension:

$$\dim V = \mathrm{vol}\,(G)\frac{\parallel v \parallel^4}{\int_G |\langle v, T(g)v\rangle|^2 dg} \tag{11.238}$$

valid for any Haar measure on $G$ and any nonzero vector $v \in V$.

The key to proving the Peter-Weyl theorem are the *orthogonality relations for matrix elements of irreps*:

**Theorem**: Let $G$ be a compact group, and define an Hermitian inner product on $L^2(G)$ by

$$\langle \Psi_1, \Psi_2\rangle := \int_G \Psi_1^*(g)\Psi_2(g)dg \tag{11.239}$$

where, WLOG we normalize the Haar measure so the volume of $G$ is one. Let $\{V^{(\mu)}\}$ be a set of representatives of the distinct isomorphism classes of irreducible unitary representations for $G$. For each representation $V^{(\mu)}$ choose an ON basis $w_i^{(\mu)}$, $i = 1, \dots, n_\mu$ with

$$n_\mu := \dim_{\mathbb{C}} V^{(\mu)} \; . \tag{11.240}$$

(We showed above that the irreps are all finite dimensional.) Then the matrix elements $M_{ij}^{\mu}(g)$ defined by

$$T^{(\mu)}(g) w_i^{(\mu)} = \sum_{j=1}^{n_\mu} M_{ji}^{\mu}(g) w_j^{(\mu)} \tag{11.241}$$

form a complete orthogonal set of functions on $L^2(G)$ so that

$$\langle M_{i_1,j_1}^{\mu_1}, M_{i_2,j_2}^{\mu_2} \rangle = \frac{1}{n_\mu} \delta^{\mu_1,\mu_2} \delta_{i_1,i_2} \delta_{j_1,j_2} \tag{11.242}$$

*Proof*: The proof is based on linear algebra and Schur's lemma. For any linear transformation $A : V^{(\mu)} \to V^{(\nu)}$ we can average using the Haar measure

$$\tilde{A} := \int_G T^{(\nu)}(g) A T^{(\mu)}(g^{-1}) dg \tag{11.243}$$

And then a small computation shows that $\tilde{A}$ is an intertwiner:

$$\begin{aligned} T^{(\nu)}(h)\tilde{A} &= \int_G T^{(\nu)}(hg) A T^{(\mu)}(g^{-1}) dg \\ &= \int_G T^{(\nu)}(g) A T^{(\mu)}(g^{-1}h) dg \\ &= \tilde{A} T^{(\mu)}(h) \end{aligned} \tag{11.244}$$

Therefore, by Schur's lemma, $\tilde{A} = \delta_{\mu,\nu} \hat{A}$ where $\hat{A}$ is a multiple of the identity transformation. It is useful at this point to choose an ordered basis and examine our conclusion: $\tilde{A} = \delta_{\mu,\nu} \hat{A}$, with $\hat{A}$ proportional to the identity, in that basis:

For any matrix $A \in \mathrm{Mat}_{n_\nu \times n_\mu}(\mathbb{C})$ we have

$$\int_G [dg] M_{ij}^{(\nu)}(g) A_{ja} M_{ab}^{(\mu)}(g^{-1}) = \delta_{\mu,\nu} c_A \delta_{i,b} \tag{11.245}$$

We can determine the constant $c_A$ by setting $\mu = \nu$ and $b = i$ and summing on $i$ to get

$$\mathrm{Tr}_{V^{(\mu)}}(A) = n_\mu c_A \tag{11.246}$$

where we normalized the volume of the group to 1.

Note that for the matrix unit $A = e_{jk}$ we have $\mathrm{Tr} e_{jk} = \delta_{jk}$ and hence

$$\int_G [dg] M_{ij}^{(\nu)}(g) M_{k\ell}^{(\mu)}(g^{-1}) = \frac{1}{n_\mu} \delta_{\mu,\nu} \delta_{jk} \delta_{i,\ell} \tag{11.247}$$

Equation (11.242) holds for the matrix elements relative to any ordered bases for the $V^{(\mu)}$. If we specialize to ON bases then the matrices $M^{(\mu)}(g)$ are unitary. It is now useful to define functions $\phi^{\mu}_{ij} \in L^2(G)$ by

$$\phi^{(\mu)}_{ij} : g \mapsto \sqrt{n_\mu} M^{(\mu)}_{ij}(g) \tag{11.248}$$

so that we have

$$\int_G [dg](\phi^{(\mu)}_{ij}(g))^* \phi^{(\nu)}_{k\ell}(g) = \delta_{\mu,\nu}\delta_{ik}\delta_{j\ell} \ . \tag{11.249}$$

We have now shown that $\{\phi^{\mu}_{ij}\}_{\mu,i,j}$ is an ON set of functions in $L^2(G)$, but we do not yet know it is an ON basis. Let $W$ be the closed subspace spanned by these functions. Then we know the orthogonal complement $\mathcal{N}$ in $L^2(G)$ is also a unitary representation. There are no infinite dimensional irreps so $\mathcal{N}$ must be reducible. So there must be a subrepresentation of $\mathcal{N}$ which is a direct sum of finite-dimensional irreps. If we have any collection of functions $\{f_j\}_{j=1}^{n_\mu}$ transforming as $V^\mu$ under the RRG then we can say

♣There is a gap in the argument here. We only proved finite dimensional reps are fully reducible into irreps. How do we know $\mathcal{N}$ isn't indecomposable with indecomposable subrep? Need to invoke some functional analysis to rule this out. ♣

$$R(g) \cdot f_j = \sum_k M^{\mu}_{kj}(g) f_k \tag{11.250}$$

so

$$f_j(hg) = \sum_k M^{\mu}_{kj}(g) f_k(h) \tag{11.251}$$

so putting $h = 1$ we get and equality of functions in $L^2(G)$

$$f_j(\cdot) = \sum_k f_k(1) M^{\mu}_{kj}(\cdot) \tag{11.252}$$

and hence $f_j$ must be in the span $W$. Thus the functions $\phi^{(\mu)}_{ij}$ are in fact an orthonormal basis for $L^2(G)$. ♠

There is a nice way of phrasing the above conclusion in terms of map $\iota : \mathrm{End}(V) \to L^2(G)$ in (11.122) and (11.121). Note that if we have any collection of finite-dimensional representations $\{V_\lambda\}$ of $G$ then we have

$$\iota : \oplus_\lambda \mathrm{End}(V_\lambda) \hookrightarrow L^2(G) \tag{11.253}$$

where we just add the functions $\iota(\oplus_i S_i) := \sum_i \Psi_{S_i}$. Thanks to the equivariance, the image of (11.253) is a $G \times G$-invariant subspace, i.e. a subrepresentation of $L^2(G)$. It follows from the above theorem that if we choose the collection $\{V_\lambda\}$ to be the set of distinct irreps of $G$, namely $\{V^{(\mu)}\}$ then $\iota$ is an isomorphism with $L^2(G)$ as a unitary $G \times G$ representation.

When working with finite groups then we have the following beautiful:

**Corollary**: If $G$ is a finite group then

$$|G| = \sum_\mu n_\mu^2 \tag{11.254}$$

*Proof*: On the one hand $L^2(G)$ clearly has a basis of delta-functions $\delta_g$ so the dimension is $|G|$ on the other hand $\mathrm{End}(V^{(\mu)})$ has dimension $n_\mu^2$

This fact is extremely useful: If you are trying to list the dimensions of the irreps of a finite group it is always a very useful check to verify (11.254). Experience shows that one easily can miss an irrep and realize the mistake by checking this identity.

♣SHOULD WE CHANGE $D$ TO $M$ FOR MATRIX ELEMENTS BELOW? ♣

**Example 1**: Let $G = \mathbb{Z}_2 = \{1, \sigma\}$ with $\sigma^2 = 1$. Then the general complex valued-function on $G$ is specified by two complex numbers $(\psi_+, \psi_-) \in \mathbb{C}^2$:

$$\Psi(1) = \psi_+ \qquad \Psi(\sigma) = \psi_- \tag{11.255}$$

This identifies $\mathrm{Map}(G, \mathbb{C}) \cong \mathbb{C}^2$ as a vector space. We found the irreps of any cyclic group above. For $\mathbb{Z}_2$ there are just two irreducible representations $V_\pm \cong \mathbb{C}$ with $\rho_\pm(\sigma) = \pm 1$. The matrix elements give two functions on the group $M^\pm$:

$$M^+(1) = 1 \qquad M^+(\sigma) = 1 \tag{11.256}$$

$$M^-(1) = 1 \qquad M^-(\sigma) = -1 \tag{11.257}$$

(Here and in the next examples when working with $1 \times 1$ matrices we drop the $\mu\nu$ subscript!) The reader can check they are orthonormal, and they are complete because any function $\Psi$ can be expressed as:

$$\Psi = \frac{\psi_+ + \psi_-}{2} M^+ + \frac{\psi_+ - \psi_-}{2} M^- \tag{11.258}$$

Note that if $V$ is any representation of $\mathbb{Z}_2$ with nontrivial element represented by $T(\sigma)$ then we can form orthogonal projection operators $P_\pm = \frac{1}{2}(1 \pm T(\sigma))$ onto direct sums of isotypical components. Note that we can write these projectors as:

$$P_\pm = \int_G (M^\pm(g))^* T(g) dg \tag{11.259}$$

**Example 2**: We can generalize the previous example slightly by taking $G = \mathbb{Z}/n\mathbb{Z} = \langle \omega | \omega^n = 1 \rangle$. Let us identify this group with the group of $n^{th}$ roots of unity and choose a generator $\omega = \exp[2\pi \mathrm{i}/n]$. Since $G$ is abelian all the representation matrices can be simultaneously diagonalized so all the irreps are one-dimensional. They are:

$V = \mathbb{C}$ and $\rho_m(\omega) = \omega^m$ where $m$ is an integer. Note that $m \sim m + n$ so the set of irreps is again labeled by $\mathbb{Z}/n\mathbb{Z}$ and in fact, under tensor product the set of irreps itself forms a group isomorphic to $\mathbb{Z}/n\mathbb{Z}$.

The matrix elements in the irrep $(\rho_m, V)$ are

♣These were called $\rho_m$ previously.... ♣

$$M^{(m)}(\omega^j) = \omega^{mj} = e^{2\pi \mathrm{i} \frac{mj}{n}} \tag{11.260}$$

Now we can check that indeed

$$\frac{1}{|G|} \sum_{g \in G} (M^{(m_1)}(g))^* M^{(m_2)}(g) = \delta_{m_1 - m_2 = 0 \mathrm{mod} n} \tag{11.261}$$

The decomposition of a function $\Psi$ on the group $G$ is known as the discrete Fourier transform: If $\Psi : \mathbb{Z}_n \to \mathbb{C}$ is any function we can write it as

$$\Psi = \sum_m \hat{\Psi}_m M^{(m)} \tag{11.262}$$

$$\hat{\Psi}_m = \int_{\mathbb{Z}_n} (M^{(m)}(g))^* \Psi(g) dg \tag{11.263}$$

**Example 3**: The theorem applies to all compact Lie groups. For example, when $G = U(1) = \{z||z| = 1\}$ then the invariant measure on the group is just $-\mathrm{i}\frac{dz}{z} = \frac{d\theta}{2\pi}$ where $z = e^{\mathrm{i}\theta}$:

$$\langle \Psi_1, \Psi_2 \rangle = \int_0^{2\pi} \frac{d\theta}{2\pi} (\Psi_1(\theta))^* \Psi_2(\theta) \tag{11.264}$$

Now, again since $G$ is abelian the irreducible representations are 1-dimensional and the unitary representations are $(\rho_n, V_n)$ where $n \in \mathbb{Z}$, $V_n \cong \mathbb{C}$ and

$$\rho_n(z) := z^n \tag{11.265}$$

Now, the orthonormality of the matrix elements is the standard orthonormality of $e^{\mathrm{i}n\theta}$ and the Peter-Weyl theorem specializes to Fourier analysis: An $L^2$-function $\Psi(\theta)$ on the circle can be expanded in terms of the matrix elements of the irreps:

$$\Psi = \sum_{\text{Irreps } \rho_n} \hat{\Psi}_n \rho_n \tag{11.266}$$

This is just the standard Fourier decomposition of a periodic function.

We stress that the Peter-Weyl theorem applies to all compact groups, not just Abelian ones. For this reason it is a basic result in the subject of "nonabelian Fourier analysis." Here is a simple nonabelian example:

**Example 4**: Let us consider $G = S_3$. It has order 6 so that $L^2(G)$ is a six-dimensional vector space. So far, we have discussed three different irreps of dimensions $1, 1, 2$. Are there any others? Note that $6 = 1^2 + 1^2 + 2^2$. So we conclude that there are no other irreps. The matrix elements are:

1. *The trivial representation*: $M^+(g) = 1$ for all $g \in S_3$
2. *The sign representation*:

$$\begin{gathered} M^-(1) = 1 \qquad\qquad M^-(123) = M^-(132) = 1 \\ M^-(12) = M^-(13) = M^-(23) = -1 \end{gathered} \tag{11.267}$$

3. *The two-dimensional representation.* See equation (11.152) et. seq. Consider the function $M^{(2)}_{11}$ on $S_3$. Its values are:

$$\begin{aligned} M^{(2)}_{11}(1) = M^{(2)}_{11}((12)) = 1 \\ M^{(2)}_{11}((13)) = M^{(2)}_{11}((23)) = -1/2 \\ M^{(2)}_{11}((123)) = M^{(2)}_{11}((123)) = -1/2 \end{aligned} \tag{11.268}$$

and so on for the other 3 matrix elements in the 2-dimensional representation. The reader should check that the function $M^{(2)}_{11}$ is orthogonal to the functions $M^+$ and $M^-$ and that $\sqrt{2}M^{(2)}_{11}$ has unit norm.

**Remarks**:

1. An important consequence of the Peter-Weyl theorem is that every compact Lie group is a matrix group, that is, is a subgroup of $U(N)$ for some $N$. This is not an obvious statement: There are simple examples of finite-dimensional Lie groups which are not matrix groups. [166] We may assume $G$ is positive dimensional, since otherwise $L^2(G)$ is finite-dimensional and the statement is obvious. In order to prove our statement for positive dimensional $G$ we note that for all $g \neq 1$ there exists a representation $V_g$ in which $T(g)$ is not the identity transformation. To prove the existence of such a $V_g$ note that if $T(g) = 1$ then $T(g^n) = 1$ but the infinite set $\{g^n\}_{n\in\mathbb{Z}}$ must have an accumulation point in $G$, since $G$ is compact. The accumulation point $g_*$ must in fact be $1_G$ since the distance in any natural $G$-invariant metric $d(g^n, g^{n+1}) = d(1_G, g) \to 0$. But the matrix elements $M^{\mu}_{ij}(g)$ are analytic functions of $g$. If they took the same value on a set with an accumulation point at $1_G$ they would be the matrix elements of the identity matrix in a neighborhood of $1_G$. This contradicts the statement that they form an ON basis of $L^2(G)$. Now consider any infinite sequence $g_1, g_2, \dots,$ which becomes dense in $G$. The kernels of the maps $G \to GL(V_{g_1} \oplus \cdots V_{g_n})$ form a nested strictly decreasing collection of closed subgroups of $G$ and must eventually become $\{1_G\}$ for some $n$.

♣This footnote should be moved earlier as it is an important remark. ♣

2. In fact, the statement that a compact Lie group is a matrix group is actually equivalent to the Peter-Weyl theorem. This is likewise not obvious. See chapter 9 of Graeme Segal's lectures in *Lectures on Lie Groups and Lie Algebras* for a nice discussion of various statements of the Peter-Weyl theorem.

3. The Peter-Weyl theorem has a remarkable generalization to the unitary (lowest weight) representation theory of the infinite-dimensional groups $LG = Map(S^1, G)$. For a positive dimensional group $G$, as we have remarked, the Peter-Weyl theorem is

[166] A nice example is given by considering the group $N/Z$ where $N$ is the group of $3\times 3$ upper triangular matrices with 1 on the diagonal. $Z$ is the subgroup of the form $1 + ne_{1,3}$ where $n \in \mathbb{Z}$. This is one model for the Heisenberg group of so much importance In quantum mechanics. It is not difficult to show (see Segal's lectures Theorem 6.5) that there is no injective homomorphism of $N/Z$ into $GL(N,\mathbb{C})$ for any finite $N$. That is, $N/Z$ is not a matrix group.

an infinite sum over irreducible representations. It turns out that the loop groups, or better, their central extensions, $\widehat{LG}_k$ are characterized by a "level $k$" and the natural analog of $L^2(\widehat{LG}_k)$ is defined by the conformal field theory of the WZW model. The decomposition

$$L^2(G) \cong \oplus_\mu (V^{(\mu)})^\vee \otimes V^{(\mu)} \tag{11.269}$$

for compact groups $G$ has an analogous form for $L^2(\widehat{LG}_k)$, but with only a finite set of irreducible representations. (The number is determined by the level $k$.) One might have thought that the infinite-dimensional group $\widehat{LG}_k$ would have many more representations than $G$, but the opposite turns out to be the case. All this is discussed in the beautiful book *Loop Groups* by A. Pressley and G. Segal.

4. This leaves the separate question of actually *constructing* the representations of the finite group $G$. For one description of how to do this in complete generality see: Vahid Dabbaghian-Abdoly, "An Algorithm for Constructing Representations of Finite Groups." Journal of Symbolic Computation Volume 39, Issue 6, June 2005, Pages 671-688

♣keep this last remark? ♣

---

**Exercise** *Due Diligence*

Check the orthogonality relations for the other matrix elements $M_{ij}^{(2)}$ on $S_3$.

---

---

**Exercise** *How Many One Dimensional Represenations Does An Arbitrary Finite Group Have?*

Show that the number of distinct one-dimensional representations of a finite group $G$ is the same as the index of the commutator subgroup $[G, G]$ in $G$. [167]

---

---

**Exercise** *Formula For The Dimension Of An Irrep V*

Show that the formula (11.238) we derived for the dimension of an irrep when proving irreps have finite dimension is actually a special case of the orthogonality relations.

---

[167] Hint: A one-dimensional representation is trivial on $[G, G]$ and hence descends to a representation of the abelianization of $G$, namely the quotient group $G/[G, G]$.

### 11.11 Explicit Decomposition Of A Representation

Let $(T, V)$ be any representation of a compact group $G$. Consider the linear operators on $V$:

$$P_{ij}^{(\mu)} := n_\mu \int_G (M_{ij}^{(\mu)}(g))^* T(g) dg \in \mathrm{End}(V) \tag{11.270}$$

where it is important that the matrix elements $M_{ij}^{(\mu)}(g)$ are taken wrt an ON basis for $V^\mu$ so that they are matrix elements of unitary matrices.

Using the orthogonality relations of the matrix elements it is now straightforward to show that:

$$P_{ij}^{(\mu)} P_{kl}^{(\nu)} = \delta_{\mu\nu}\delta_{j,k} P_{i\ell}^{(\nu)} \tag{11.271}$$

This operator algebra will be very useful, and we will use it to construct a system of orthogonal projection operators onto the isotypical components of $V$ when we come to characters below.

It also follows from the definition of $P_{ij}^{\mu}$ that, for any vector $\psi \in V$, we have

$$T(h) P_{ij}^{\mu} \psi = \sum_{k=1}^{n_\mu} M_{ki}^{\mu}(h) P_{kj}^{\mu} \psi \tag{11.272}$$

Therefore, if the $P_{ij}^{\mu}\psi \neq 0$ then, for fixed $\mu, j$ the span

$$Span\{P_{ij}^{\mu}\psi | i = 1, \dots, n_\mu\} \tag{11.273}$$

is a subspace of $V$ transforming in the representation $(V^\mu, T^\mu)$. Using these operators we can in principle decompose a representation into its irreps.

---

**Exercise** *Projector To The Isotypical Component For The Trivial Representation*
Show that

$$P = \int_G T(g) dg \tag{11.274}$$

is a projection operator onto the isotypical subspace of $(T, V)$ corresponding to the trivial representation.

---

---

**Exercise** *Isotypical Decomposition Of The Permutation Representation Of $S_3$ On $\mathbb{R}^3$*

Consider the example of the 3-dimensional reducible representation of $S_3$ acting on $\mathbb{R}^3$ in section *** above.

Referring to the three irreducible representations of $S_3$, called $1_+, 1_-, 2$ in section **** show that:

$$
\begin{aligned}
P_{1_+} &= \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad P_{1_+}^2 = P_{1_+} \\
P_{1_-} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
P_{\mathbf{2}}^{1,1} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1/2 & -1/2 \\ 0 & -1/2 & 1/2 \end{pmatrix} \qquad (P_{\mathbf{2}}^{1,1})^2 = P_{\mathbf{2}}^{1,1} \\
P_{\mathbf{2}}^{2,1} &= \begin{pmatrix} 0 & 1/\sqrt{3} & -1/\sqrt{3} \\ 0 & -1/2\sqrt{3} & 1/2\sqrt{3} \\ 0 & -1/2\sqrt{3} & 1/2\sqrt{3} \end{pmatrix}
\end{aligned} \tag{11.275}
$$

Therefore,

$$
P_{\mathbf{1}+}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (\frac{x+y+z}{3})\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{11.276}
$$

so:

$$
V_1 = Span\left\{\vec{V}_1 = (1,1,1)\right\} = \{(x,x,x) : x \in \mathbb{R}\} \tag{11.277}
$$

Similarly $P_{\mathbf{2}}$ projects onto the orthogonal subspace

$$
P_{\mathbf{2}}^{1,1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{2}(y-z)\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \qquad P_{\mathbf{2}}^{2,1}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{\sqrt{3}}{2}(y-z)\begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \tag{11.278}
$$

So:

$$
V_2 = Span\left\{\vec{V}_2 = (0,1,-1)\vec{V}_3 = (2,-1,-1)\right\} \tag{11.279}
$$

are the two invariant subspaces. Check

$$
S = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 2 & -1 & -1 \end{pmatrix} \tag{11.280}
$$

Conjugates the rep into block form. Thus, the decomposition into irreps is:

$$
\mathbb{R}^3 \cong V_{\mathbf{1}+} \oplus V_{\mathbf{2}} \tag{11.281}
$$

---

### 11.11.1 Block diagonalization of Hermitian operators

Suppose $H$ is an $N \times N$ Hermitian operator acting on the inner product space $\mathbb{C}^N$ with standard innder product. Suppose we have a group representation $T$ of $G$ with $\mathbb{C}^N$ as carrier space and that moreover, $G$ is a symmetry of $H$ in the sense that

$$\forall g \in G \qquad [T(g), H] = 0 \tag{11.282}$$

Then

$$[P^{\mu}_{ij}, H] = 0 \tag{11.283}$$

Then, the projection operators $P_\lambda$ onto the different eigenspaces of $H$ commute with the different projection operators $P^{\mu}_{ij}$.

*Therefore, in this situation, by reducing $\mathbb{C}^N$ to the irreps of $G$, we have partially block-diagonalized $H$.*

Of course, this is very useful in quantum mechanics where one wants to diagonalize Hermitian operators acting on wavefunctions.

### 11.11.2 Projecting quantum wavefunctions

Suppose $G$ acts on $X$ as a tranformation group, and $\{\psi_a\}$ is a collection of functions on $X$ transforming according to some representation $T$ of $G$. (For example, energy eigenfunctions in a Schrödinger problem.)
Then:

$$\psi^{ij}_{a,\mu}(x) \equiv \frac{n_\mu}{\mid G \mid} \sum_{h \epsilon G} (T^{\mu}_{ij}(h))^* \psi_a(h^{-1} \cdot x) \tag{11.284}$$

for fixed $\mu, j, a$, letting $i = 1, \dots, n_\mu$ be a collection of functions transforming according to the irrep $T^\mu$: This is a special case of the general statement we made above. Here, again, is the explicit calculation:

$$\begin{aligned}
\psi^{ij}_a(g^{-1} \cdot x) &= \frac{n_\mu}{\mid G \mid} \sum_{h \epsilon G} (T^{\mu}_{ij}(h))^* \psi_a(h^{-1} \cdot g^{-1} \cdot x) \\
&= \frac{n_\mu}{\mid G \mid} \sum_{h \epsilon G} (T^{\mu}_{ij}(g^{-1}h))^* \psi_a(h^{-1} \cdot x) \\
&= \sum_{s=1}^{n_\mu} (T^{\mu}_{is}(g^{-1}))^* \psi^{sj}_a(x) \\
&= \sum_{s=1}^{n_\mu} T^{\mu}_{si}(g) \psi^{sj}_a(x) \qquad \text{if } T^\mu \text{ is unitary.}
\end{aligned} \tag{11.285}$$

**Example 2** As a somewhat trivial special case of the above consider $\mathbb{Z}$ acting on $\mathbb{R} : x \to$

$-x$ therefore acts on $Fun(\mathbb{R})$. Decompose into irreps:

$\psi =$ function on $\mathbb{R}$

$$\begin{aligned}\psi^1(x) &= \frac{1}{2}(\psi(x)+\psi(-x)) \\ \psi^2(x) &= \frac{1}{2}(\psi(x)-\psi(-x))\end{aligned} \tag{11.286}$$

transform acccording to the two irreps of $\mathbb{Z}_2$.

### 11.11.3 Finding normal modes in classical mechanics

Consider a classical mechanical system with degrees of freedom

$$\vec{q} = (q^1, \dots, q^n) \tag{11.287}$$

Suppose we have a generalized harmonic oscillator so that the kinetic and potential energies are:

$$T = \frac{1}{2}m_{ij}\dot{q}^i\dot{q}^j \qquad\qquad V = \frac{1}{2}U_{ij}q^iq^j \tag{11.288}$$

where $U_{ij}$ is independent of $q^i$. We can obtain solutions of the classical equations of motion by taking

$$q^i(t) = \mathrm{Re}(\gamma^i e^{i\omega t}) \tag{11.289}$$

where

$$(-\omega^2 m_{ij} + U_{ij})\gamma^j = 0 \tag{11.290}$$

If a group $G$ acts as $(T(g)\cdot q)^i = T(g)^{ij}q^j$ and

$$[T(g), m] = [T(g), U] = 0 \tag{11.291}$$

then we can partially diagonalize $H = -\omega^2 m + U$ by choosing $\vec{\gamma}$ to be in irreps of $G$.

As a special case consider the following example. We have $n$ particles, of equal mass $m$, on the real line, at positions $q^i$ so that particle $i$ is connected to particle $i+1$ by a spring, and the last particle at $q^n$ is connected to the one at $q^1$ by a spring as well. Assume that all the spring constants are the same. The Lagrangian is

$$L = \frac{1}{2}m(\dot{q}^i)^2 - \frac{1}{2}k\sum_i (q^i - q^{i+1})^2 \tag{11.292}$$

where we understand the superscript on $q^i$ to modulo $n$:

$$q^{i+n} = q^i \tag{11.293}$$

The normal mode problem is

$$(-m\omega^2 + kA)\gamma = 0 \tag{11.294}$$

where

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & \cdots & -1 \\ -1 & 2 & -1 & 0 & 0 & 0 & \cdots \\ 0 & -1 & 2 & -1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & -1 \\ -1 & 0 & \cdots & 0 & 0 & -1 & 2 \end{pmatrix} \tag{11.295}$$

Thus, we reduce the problem of finding normal modes to the problem of diagonalizing this matrix.

The problem has an obvious $\mathbb{Z}_n$ symmetry where a generator $T(\omega)$ takes

$$T(\omega) : q^1 e_1 + \cdots q^n e_n \to q^1 e_2 + \cdots + q^n e_{n-1} + q^1 e_n \tag{11.296}$$

Here $e_i$ is the standard basis for $\mathbb{C}^n$ so $T(\omega^\ell) e_i = e_{i+\ell}$ and again the subscript is understood modulo $n$.

We know the irreps of $\mathbb{Z}_n$: $T^\mu(\omega^j) = \omega^{j\mu}$, $\mu = 0, \dots, n-1$. The projection operators are thus

$$\mathcal{P}^\mu = \frac{1}{n} \sum_{\ell=0}^{n-1} \omega^{-\mu\ell} T(\omega^\ell) \tag{11.297}$$

In terms of matrix units $e_{ij}$ we have

$$\mathcal{P}^\mu = \frac{1}{n} \sum_{i,\ell} \omega^{-\mu\ell} e_{i+\ell,i} \tag{11.298}$$

On the other hand,

$$A = 2 - \sum_i (e_{i,i+1} + e_{i+1,i}) \tag{11.299}$$

Now we carry out the multiplication:

$$A\mathcal{P}^\mu = 2\mathcal{P}^\mu - \sum_{i,\ell} \omega^{-\mu\ell} (e_{i+\ell,i+1} + e_{i+\ell,i-1}) \tag{11.300}$$

After some relabeling of the subscripts we easily find

$$\begin{aligned} A\mathcal{P}^\mu &= (2 - \omega^{-\mu} - \omega^\mu)\mathcal{P}^\mu \\ &= (2 \sin \frac{\pi\mu}{n})^2 \mathcal{P}^\mu \end{aligned} \tag{11.301}$$

Similarly,

$$\mathcal{P}^\mu \vec{q} = \frac{1}{n} \sum_{\ell,s} \omega^{-\mu\ell} q^s e_{s+\ell} = \frac{1}{n} \sum_s \omega^{\mu s} q^s \left( \sum_j \omega^{-\mu j} e_j \right) \tag{11.302}$$

is a linear combination of the vectors

$$\gamma^{(\mu)} = \sum_j \omega^{-\mu j} e_j \tag{11.303}$$

These are the normal mode vectors. The normal mode frequencies are:

$$\omega_\mu^2 = \frac{k}{m} (2 \sin \frac{\pi\mu}{n})^2 \tag{11.304}$$

and the normal mode motions are:

$$\vec{q}_{(\mu)} = \mathrm{Re}(\alpha_\mu e^{i\omega_\mu t} \gamma^{(\mu)}) \tag{11.305}$$

where $\alpha_\mu$ is an arbitrary nonzero complex number.

**Example 1**: For the trivial representation $\mu = 0$ we have $\omega_\mu = 0$ and $\gamma^{(0)}$ is the all ones vector: This is the "motion" where all the particles sit at the same point and are motionless.

**Example 2**: Consider $n = 2$, so we have two particles connected by a spring. Then we have a linear combination of the constant $c_0\gamma^{(0)}$, giving the center of mass and the other mode is

$$\vec{q}^{(1)}(t) = \alpha e^{i2\omega_0 t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{11.306}$$

and describes oscillation of the two particles around their center of mass.

**Remark**: Note that there is a *degeneracy* in the mode frequencies

$$\omega_\mu = \omega_{n-\mu}. \tag{11.307}$$

which is not explained by the cyclic group symmetry we have thus far exploited. *It is a general principle that when you have unexplained degeneracies you should search for a further symmetry in the problem.* Indeed, in this case, we have overlooked a further symmetry:

$$q^i \to q^{n-i} \tag{11.308}$$

As we will see when we discuss semidirect products adding this symmetry operation generalizes the $\mathbb{Z}_n$ symmetry to a group twice as larger often denoted as $D_n$. The larger $D_n$ symmetry of the problem explains the degeneracy (11.307).

---

**Exercise** *Normal Motions*

Describe the three distinct normal modes for $n = 3$ and draw the motions.

---

---

**Exercise**

Using the character tables above for $D_n$ construct the normal modes which are in representations of $D_n$.

♣This exercise needs to be moved below to the section where we introduce $D_n$. ♣

---

## 11.12 Orthogonality Relations For Characters And Character Tables

Recall that a class function on $G$ is a function $f : G \to \mathbb{C}$ whose value only depends on the conjugacy class: $f(g) = f(hgh^{-1})$ for all $g, h \in G$. The class functions form a subspace

$$L^2(G)^{class} \subset L^2(G) \ . \tag{11.309}$$

The characters of the irreps are class functions, and in fact they provide an ON basis for $L^2(G)^{class}$:

**Theorem**. $\{\chi^\mu\}$ is an ON basis for the vector space of class functions $L^2(G)^{class}$.

*Proof*: The proof that $\{\chi^\mu\}$ is an ON set of functions on $L^2(G)$ is an easy consequence of the orthogonality relations (11.242) or (11.247) above. If we take traces by setting $i = j$ and $k = \ell$ and summing on $i$ and $k$ in equation (11.247) we get the orthogonality relations of characters:

$$\int_G (\chi^{(\mu)}(g))^* \chi^{(\nu)}(g) dg = \delta_{\mu,\nu} \tag{11.310}$$

Now we show that the set of characters of the irreps $\{\chi^{(\mu)}\}$ is also a spanning set for the subspace of $L^2(G)$ of class functions. By the Peter-Weyl theorem any function $f \in L^2(G)$ be expanded

$$f(g) = \sum_{\mu,i,j} \hat{f}^\mu_{ij} M^\mu_{ij}(g) \tag{11.311}$$

Consider the averaged function defined by the integral:

$$\tilde{f}(g) := \int_G f(hgh^{-1}) dh \tag{11.312}$$

If $f$ is a class function we get $\tilde{f} = f$, provided the Haar measure is normalized so that $\mathrm{vol}\,(G) = 1$, as we are assuming.

On the other hand, substituting (11.311) into the integration we have:

$$\begin{aligned} \tilde{f}(g) &= \int_G f(hgh^{-1}) dh \\ &= \sum_{\mu,i,j} \hat{f}^\mu_{ij} \int_G M^\mu_{ij}(hgh^{-1}) dh \\ &= \sum_{\mu,i,j} \hat{f}^\mu_{ij} M^\mu_{kl}(g) \int_G M^\mu_{ik}(h) M^\mu_{lj}(h^{-1}) dh \\ &= \sum_{\mu,i} \frac{\hat{f}^\mu_{ii}}{n_\mu} \chi^\mu(g) \end{aligned} \tag{11.313}$$

But $\tilde{f} = f$ for a class function, so $f$ has been expanded in terms of the characters of irreps. ♠

The orthogonality of the characters is extremely powerful. For example, as we have seen, any representation $(T,V)$ of a compact group $G$ is completely reducible so

$$V \cong \oplus_\mu a_\mu V^{(\mu)} \tag{11.314}$$

with $a_\mu \in \mathbb{Z}_+$. But then we can determine these degeneracies by

$$a_\mu = \int_G (\chi^{(\mu)}(g))^* \chi_V(g) dg = \langle \chi^{(\mu)}, \chi_V \rangle \tag{11.315}$$

Therefore, we conclude the extremely important fact:

*The isomorphism class of a representation of a compact group is completely determined by its character function!*

Moreover, by computing the overlap integral of characters we can easily determine the decomposition into isotypical components.

**Remark** The characters satisfy many other interesting relations. For example, for a finite group one can show that if $z \in G$

$$\frac{1}{|G|}\#\{(a_i,b_i) \in G | [a_1,b_1]\cdots[a_g,b_g] = z\} = \sum_{\chi \in Irrep(G)} \left(\frac{|G|}{\chi(1)}\right)^{2g-2} \frac{\chi(z)}{\chi(1)} \tag{11.316}$$

This fact goes back to Frobenius 1896 and Hurwitz 1902, and has a very elegant interpretation today in terms of moduli spaces of flat $G$-bundles over a compact surface of genus $g$, and is related to topological quantum field theory.

---

**Exercise**

Using the orthogonality relations on matrix elements, derive the more general relation on characters:

$$\int_G \chi^\mu(g)\chi^\nu(g^{-1}h)dg = \frac{\delta_{\mu\nu}}{n_\mu}\chi^\nu(h) \tag{11.317}$$

We will interpret this more conceptually later.

---

### 11.12.1 Explicit Projection Onto Isotypical Subspaces

Suppose $V$ is completely reducible with isotypical decomposition:

$$V \cong \oplus_\mu \mathcal{H}^{(\mu)} \tag{11.318}$$

where $\mathcal{H}^{(\mu)} \cong D_\mu \otimes V^{(\mu)}$ and $D_\mu \cong \mathrm{Hom}_G(V^{(\mu)}, V)$ is the degeneracy space. We now construct explicit projection operators onto the isotypical subspaces of $V$ using the operators

$$P^\mu_{ij} := n_\mu \int_G (M^{(\mu)}_{ij}(g))^* T(g) dg \in \mathrm{End}(V) \tag{11.319}$$

defined above in (11.270). Recall that these satisfy the operator algebra

$$P^{\mu}_{ij} P^{\nu}_{kl} = \delta_{\mu\nu} \delta_{j,k} P^{\nu}_{i\ell} \tag{11.320}$$

Therefore, if we define

$$P^{\mu} := \sum_{i=1}^{n_\mu} P^{\mu}_{ii} \tag{11.321}$$

then this set of operators satisfies a much simpler algebra:

$$P^{\mu} P^{\nu} = \delta_{\mu,\nu} P^{\nu} \tag{11.322}$$

In particular, the $P^{\mu}$ are orthogonal projection operators. They are also much easier to compute than the $P^{\mu}_{ij}$ because it follows from the definition that

$$P^{\mu} = n_\mu \int_G (\chi^{\mu}(g))^* T(g) dg \tag{11.323}$$

Finally, note that if $T$ is a unitary representation then

$$(P^{\mu})^{\dagger} = P^{\mu} \tag{11.324}$$

so the $P^{\mu}$ are Hermitian.

Now we claim that *the $P^{\mu}$ are orthogonal projection operators onto the isotypical components of $V$.*

To see this first note that from the results above, any vector $\psi \in V$, the vector $P^{\mu}_{ij}\psi$ transforms like

$$T(h) P^{\mu} \psi = \sum_{i,k=1}^{n_\mu} M^{\mu}_{ki}(h) P^{\mu}_{ki} \psi \tag{11.325}$$

So $P^{\mu}\psi$ transforms within the space spanned by $\{P^{\mu}_{ki}\psi\}$ and we have already seen that by fixing $i$ and letting $k$ vary we get a set of vectors transforming in a representation isomorphic to $V^{\mu}$. So $P^{\mu}\psi \in \mathcal{H}^{\mu}$. Moreover

$$\mathrm{Tr}(P^{\mu}) = n_\mu \int_G (\chi^{\mu}(g))^* \chi(g) dg = n_\mu a_\mu \tag{11.326}$$

is precisely the dimension of the isotypical subspace.

---

**Exercise** *Due Diligence*

Show that if $T : G \to U(V)$ is a unitary representation then $P^{\mu}$ are Hermitian operators.

---

---

**Exercise** *Explicit Example*

Consider the permutation representation of $S_3$ on $\mathbb{R}^3$. Show that the projector for the two-dimensional irrep is:

$$\mathcal{P}_{\mathbf{2}} = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \tag{11.327}$$

---

### 11.12.2 Finite Groups And The Character Table

♣We have some notational inconsistencies $\chi_\mu$ vs. $\chi^\mu$. Irreps of $S_2, S_3$ are given different names in different places. ♣

In the case of a finite group there is another obvious basis of the space of class functions: Denote the distinct conjugacy classes by $C_i, i = 1, \dots r$. For each $C_i$ we can define a class function $\delta_{C_i}$ in $L^2(G)$ to be the characteristic function of $C_i$. In equations:

$$\delta_{C_i}(g) := \begin{cases} 1 & g \in C_i \\ 0 & else \end{cases} \tag{11.328}$$

Since any class function takes the same value for all group elements $g$ in a fixed conjugacy class it follows that the functions $\delta_{C_i}$ form a basis for the space of class functions. Any two bases for a given vector space have the same cardinality so we conclude:

**Theorem** The number of conjugacy classes of $G$ is the same as the number of irreducible representations of $G$.

Since the number of representations and conjugacy classes are the same we can define a square $r \times r$ array known as a *character table*: The rows are labeled by the irreps $V^\mu$ (almost always with the trivial representation taken as $V^1$) and the columns are labeled by conjugacy classes $C_i$ (almost always with the conjugacy class of the identity taken as $C_1$) and the matrix elements are the common value $\chi^\mu(C_i)$ that $\chi^\mu$ takes on all the elements of the conjugacy class $C_i$:

$$\begin{array}{cccccc} & m_1 C_1 & m_2 C_2 & \cdots & \cdots & m_r C_r \\ V^1 & \chi_1(C_1) & \cdots & \cdots & \cdots & \chi_1(C_r) \\ V^2 & \chi_2(C_1) & \cdots & \cdots & \cdots & \chi_2(C_r) \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ V^r & \chi_r(C_1) & \cdots & \cdots & \cdots & \chi_r(C_r) \end{array} \tag{11.329}$$

Here $m_i$ denotes the order of $C_i$. It is useful information to encode in the top row of the table when computing inner products of characters.

**Remark**: Generally speaking, most of what you typically want to know about a finite group is contained in its character table, even if you do not know how to compute irreps explicitly. Online you can find many lists of the character tables of commonly used small order groups. A particularly useful resource is the computer program GAP which stands for Groups, Algorithms, Programming. It is a system for computational discrete algebra and will give the character tables of almost any finite group you are likely to encounter.

For a finite group we can rewrite the orthogonality relations of the characters, equation (11.310), more explicitly as:

$$\frac{1}{|G|}\sum_{C_i\in\mathcal{C}} m_i\chi_\mu(C_i)\chi_\nu(C_i)^* = \delta_{\mu\nu} \tag{11.330}$$

where $m_i = |C_i|$ is the order of the conjugacy class $C_i$.

We claim that there is a "dual" orthogonality relation where we sum over irreps rather than conjugacy classes

$$\sum_\mu \chi_\mu(C_i)^*\chi_\mu(C_j) = \frac{|G|}{m_i}\delta_{ij} \tag{11.331}$$

The proof of (11.331) is very elegant: Note that equation (11.330) can be interpreted as the statement that the $r\times r$ matrix

$$S_{\mu i} := \sqrt{\frac{m_i}{|G|}}\chi_\mu(C_i) \qquad \mu = 1,\dots,r \quad i = 1,\dots,r \tag{11.332}$$

satisfies

$$\sum_{i=1}^{r} S_{\mu i}S^*_{\nu i} = \delta_{\mu\nu} \tag{11.333}$$

Therefore, $S_{\mu i}$ is a unitary matrix. The left-inverse is the same as the right-inverse, and hence we obtain (11.331).

**Example 1**: For the simplest group there are two conjugacy classes $\mathcal{C}_1 = [1]$ and $\mathcal{C}_2 = [(12)]$. They both have cardinality $m_i = 1$. There are two irreps $\mathbf{1}^+$ and the sign representation $\mathbf{1}^-$. We thus have the character table

| | [1] | [(12)] |
|---|---|---|
| $1^+$ | 1 | 1 |
| $1^-$ | 1 | −1 |

**Example 2**: Consider the cyclic group $\mathbb{Z}/n\mathbb{Z}$. We have seen that the irreps are themselves labeled by elements of the cyclic group $\rho_m$, with $m \in \mathbb{Z}/n\mathbb{Z}$. The irreps are all one

dimensional. The conjugacy classes all have order 1 since $\mathbb{Z}/n\mathbb{Z}$ is Abelian. The character table is an $n \times n$ matrix. The elements of the character table are thus: For row $\rho_m$ and column $\bar{j} \in \mathbb{Z}/n\mathbb{Z}$ we have $\rho_m(\bar{j}) = \omega^{mj}$ where $\omega = \exp[2\pi \mathrm{i}/n]$.

So, the character table for $\mathbb{Z}/3\mathbb{Z}$ is

| | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ |
|---|---|---|---|
| $\rho_0$ | 1 | 1 | 1 |
| $\rho_1$ | 1 | $\omega$ | $\omega^2$ |
| $\rho_2$ | 1 | $\omega^2$ | $\omega$ |

where $\omega = e^{2\pi \mathrm{i}/3}$.

Similarly, the character table for $\mathbb{Z}/4\mathbb{Z}$ is

| | $\bar{0}$ | $\bar{1}$ | $\bar{2}$ | $\bar{3}$ |
|---|---|---|---|---|
| $\rho_0$ | 1 | 1 | 1 | 1 |
| $\rho_1$ | 1 | i | $-1$ | $-$i |
| $\rho_2$ | 1 | $-1$ | 1 | $-1$ |
| $\rho_3$ | 1 | $-$i | $-1$ | i |

Note that the matrix is symmetric. This is related to the Pontryagin self-duality of $\mathbb{Z}/n\mathbb{Z}$. There are other symmetries of the character table. These arise from the automorphisms of the cyclic group. For the general case the matrix $S$ is:

$$S_{mj} = \sqrt{\frac{1}{n}}\rho_m(\omega^j) = \sqrt{\frac{1}{n}}\exp\left(2\pi \mathrm{i}\frac{mj}{n}\right) \tag{11.334}$$

This is the matrix of a finite Fourier transform.

**Example 3**: We listed the irreps of $S_3$ above and it is straightforward to compute the character table:

| | [1] | 3[(12)] | 2[(123)] |
|---|---|---|---|
| $\mathbf{1}^+$ | 1 | 1 | 1 |
| $\mathbf{1}^-$ | 1 | −1 | 1 |
| $\mathbf{2}$ | 2 | 0 | −1 |

Now let us see how we can use the orthogonality relations on characters to find the decomposition of a reducible representation.

**Example 1** Consider the $3 \times 3$ rep generated by the ♮ action of the permutation group $S_3$ on $\mathbb{R}^3$. We'll compute the characters by choosing one representative from each conjugacy class:

$$1 \to \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (12) \to \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (132) \to \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \tag{11.335}$$

From these representatives the character of $V = \mathbb{R}^3$ is easily calculated:

$$\chi_V(1) = 3 \qquad \chi_V([(12)]) = 1 \qquad \chi_V([(132)] = 0 \tag{11.336}$$

Using the orthogonality relations we compute

$$a_{\mathbf{1}_+} = (\chi_{\mathbf{1}_+}, \chi) = \frac{1}{6}3 + \frac{3}{6}1 + \frac{2}{6}0 = 1 \tag{11.337}$$

$$a_{\mathbf{1}_-} = (\chi_{\mathbf{1}_-}, \chi) = \frac{1}{6}3 + \frac{3}{6}(-1) \cdot 1 + \frac{2}{6}0 = 0 \tag{11.338}$$

$$a_{\mathbf{2}} = (\chi_{\mathbf{2}}, \chi) = \frac{1}{6}3 \cdot 2 + \frac{3}{6}0 \cdot 1 + \frac{2}{6}(-1) \cdot 0 = 1 \tag{11.339}$$

Therefore:

$$\chi_V = \chi_{\mathbf{1}_+} + \chi_{\mathbf{2}} \tag{11.340}$$

showing the decomposition of $\mathbb{R}^3$ into irreps, and confirming what we showed above.

**Example 4**: Let $V$ be any vector space. Consider the natural permutation on $V\otimes V$ by $S_2$. where $V$ is a finite dimensional vector space (over any field) of dimension $d$. Let us write out the isotypical decomposition under the $S_2$ action. The character is easily computed:

$$\begin{aligned}\chi_{V^{\otimes 2}}(1) &= d^2 \\ \chi_{V^{\otimes 2}}((12)) &= d\end{aligned} \tag{11.341}$$

To check the second line we choose a basis $\{v_i\}$ for $V$ so that $\{v_i\otimes v_j\}$ is a basis for $V\otimes V$. In this basis we have $(12)\cdot v_i\otimes v_j = v_j\otimes v_i$ and hence only the basis elements $v_i\otimes v_i$ contribute to the trace. Therefore, we compute the degeneracies of the trivial and sign representation from:

$$\begin{aligned}a_{1^+} &= \langle\chi^+,\chi_{V^{\otimes 2}}\rangle = \frac{1}{2}(\chi_{V^{\otimes 2}}(1)+\chi_{V^{\otimes 2}}((12))) \\ &= \frac{1}{2}d(d+1) \\ a_{1^-} &= \langle\chi^-,\chi_{V^{\otimes 2}}\rangle = \frac{1}{2}(\chi_{V^{\otimes 2}}(1)-\chi_{V^{\otimes 2}}((12))) \\ &= \frac{1}{2}d(d-1)\end{aligned} \tag{11.342}$$

so we have isotypical decomposition:

$$V^{\otimes 2} = \frac{1}{2}d(d+1)\mathbf{1}^+ \oplus \frac{1}{2}d(d-1)\mathbf{1}^- \tag{11.343}$$

General elements of $V^{\otimes 2}$ are of the form $T_{ij}v_i\otimes v_j$ and the components are often referred to as 2-index covariant tensors. The above decomposition is a decomposition into *symmetry types* of tensors: We can choose a basis of symmetric tensors:

$$\frac{1}{2}(e_i\otimes e_j + e_j\otimes e_i) \qquad 1\le i\le j\le d \tag{11.344}$$

and anti-symmetric tensors

$$\frac{1}{2}(e_i\otimes e_j - e_j\otimes e_i) \qquad 1\le i< j\le d \tag{11.345}$$

**Example 3**: The previous example is the beginning of a very beautiful story called Schur-Weyl duality. Let us consider the next case: Consider $S_3$ acting by permuting the various factors in the tensor space $V\otimes V\otimes V$ for any vector space $V$. Now, if $\dim V = d$ then we have

$$\begin{aligned}\chi([1]) &= d^3 \\ \chi([(ab)]) &= d^2 \\ \chi[(abc)] &= d\end{aligned} \tag{11.346}$$

as is easily computed by considering the action on the basis $\{v_i\otimes v_j\otimes v_k\}$.

So we can compute

$$a_{\mathbf{1}_+} = (\chi_{\mathbf{1}_+}, \chi) = \frac{1}{6}d^3 + \frac{3}{6}d^2 + \frac{2}{6}d = \frac{1}{6}d(d+1)(d+2) \tag{11.347}$$

$$a_{\mathbf{1}_-} = (\chi_{\mathbf{1}_-}, \chi) = \frac{1}{6}d^3 + \frac{3}{6}(-1)\cdot d^2 + \frac{2}{6}d = \frac{1}{6}d(d-1)(d-2) \tag{11.348}$$

$$a_{\mathbf{2}} = (\chi_{\mathbf{2}}, \chi) = \frac{1}{6}2d^3 + \frac{3}{6}0\cdot d^2 + \frac{2}{6}(-1)\cdot d = \frac{1}{3}d(d^2-1) \tag{11.349}$$

Thus, *as a representation of* $S_3$, we have

$$V^{\otimes 3} \cong \frac{d(d+1)(d+2)}{6}\mathbf{1}_+ \oplus \frac{d(d-1)(d-2)}{6}\mathbf{1}_- \oplus \frac{d(d+1)(d-1)}{3}\mathbf{2} \tag{11.350}$$

Note that the first two dimensions are those of $S^3V$ and $\Lambda^3 V$, respectively, and that the dimensions add up correctly. It is not supposed to be obvious, but it turns out that the last summand corresponds to tensor of mixed symmetry type that satisfy

$$T_{ijk} + T_{jki} + T_{kij} = 0 \tag{11.351}$$

together with

$$T_{ijk} = -T_{kji} \tag{11.352}$$

Why this is so, and the generalization to the $S_n$ action on $V^{\otimes n}$ is best discussed in the context of Young diagrams, representation theory of $S_n$ and Schur-Weyl duality discussed in subsection 11.16 below.

---

**Exercise**

Show that right-multiplication of the character table by a diagonal matrix produces a unitary matrix.

---

---

**Exercise** *Another proof of* (11.331)

Consider the operator $L(g_1) \otimes R(g_2)$ acting on the regular representation. We will compute

$$\mathrm{Tr}_{R_G}[L(g_1) \otimes R(g_2)] \tag{11.353}$$

in two bases.

First consider the basis $\phi^{\mu}_{ij}$ of matrix elements. We have:

$$L(g_1) \otimes R(g_2) \cdot \phi^{\mu}_{ij} = \sum_{i',j'} T^{\mu}_{j'j}(g_2) T^{\mu}_{i'i}(g_1^{-1}) \phi^{\mu}_{i'j'} \tag{11.354}$$

so the trace in this basis is just:

$$\mathrm{Tr}_{R_G}[L(g_1) \otimes R(g_2)] = \sum_{\mu} \chi_{\mu}(g_1)^* \chi_{\mu}(g_2) \tag{11.355}$$

On the other hand, we can use the delta-function basis: $\delta_g$. Note that

$$L(g_1) \otimes R(g_2) \cdot \delta_g = \delta_{g_1 g g_2^{-1}} \tag{11.356}$$

So in the delta function basis we get a contribution of +1 to the trace iff $g = g_1 g g_2^{-1}$, that is iff $g_2 = g^{-1} g_1 g$, that is iff $g_1$ and $g_2$ are conjugate, otherwise we get zero.

Now when $g_1$ and $g_2$ are conjugate we might as well take them to be equal. Then the functions $\delta_g$ contributing to the trace are precisely those with $g \in Z(g_1)$. But then

$$|Z(g_1)| = \frac{|G|}{|C(g_1)|} = \frac{|G|}{m_1} \tag{11.357}$$

completing the proof of (11.331).

---

---

**Exercise** *Average Number Of Fixed Points - Again*

Recall that in a previous exercise you showed that if a finite group $G$ acts on a finite set $X$ then

$$\frac{1}{|G|} \sum_g |X^g| = |\{orbits\}| \tag{11.358}$$

Prove this again by viewing $L^2(X)$ as a $G$-representation and using the orthgonality relations on characters. [168]

---

---

**Exercise**

Write out the unitary matrix $S_{\mu i}$ for $G = S_3$. [169]

---

---

**Exercise**

a.) Suppose we tried to define a representation of $S_3$ by taking $(12) \to 1$ and $(23) \to -1$. What goes wrong?

♣this is really about generators and relations and should be moved up. ♣

[168] *Answer*: Note that one basis for $L^2(X)$ is given by $\delta_x$. Here $g \cdot \delta_x = \delta_{g\cdot x}$ so $|X^g|$ is just the character of $g$ in this representation. Therefore the average number of fixed points is the degeneracy of the trivial representation in the isotypical decomposition. On the other hand, if $f : X \to \mathbb{C}$ is invariant under the $g$ action then $f$ takes the same value on any two $x_1, x_2$ in the same $G$-orbit. Therefore, the subspace corresponding to the trivial representation in the isotypical decomposition has a basis given by the characteristic function for each orbit.

[169] *Answer*: The unitary matrix $S_{\mu i}$ for $S_3$ is:

$$S_{\mu i} = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \end{pmatrix} \tag{11.359}$$

b.) Show that for any $n$ there are only two one-dimensional representations of $S_n$

---

---

**Exercise** *Universal Properties Of Character Tables*

Show that if we partially order the conjugacy classes and the representations so that the class of the identity and the trivial representation come first then the first row of the character table is all 1's and the first column of the character table gives the dimensions of the irreducible representations.

---

## 11.13 Decomposition Of Tensor Products Of Representations And Fusion Coefficients

A frequently asked question is the following: Suppose we know the reduction of two representations of $G$, say, $V_1, V_2$ into irreps. What is the decomposition of the tensor product $V_1 \otimes V_2$ into irreps? For compact groups this is nicely answered with characters.

Suppose we have two representations $(T_1, V_1)$ and $(T_2, V_2)$ and we know the isotypical decompositions:

$$V_1 = \oplus a_\mu V^\mu \qquad V_2 = \oplus b_\nu V^\nu \tag{11.360}$$

then

$$V_1 \otimes V_2 = \oplus_{\mu,\nu} a_\mu b_\nu V^\mu \otimes V^\nu \tag{11.361}$$

so the isotypical decomposition of $V_1 \otimes V_2$ follows immediately if we can find the isotypical decomposition of the tensor product of two irreps.

On general grounds we know that ♣$\kappa = \mathbb{C}$ here? ♣

$$V^\mu \otimes V^\nu \cong \oplus_\lambda \mathrm{Hom}_G(V^\lambda, V^\mu \otimes V^\nu) \otimes V^\lambda \tag{11.362}$$

The dimensions

$$N_{\mu\nu}^{\lambda} := \dim_\kappa \mathrm{Hom}_G(V^\lambda, V^\mu \otimes V^\nu) \tag{11.363}$$

are known as the *fusion coefficients* and they give the isotypical decomposition:

$$\begin{aligned} V^{(\mu)} \otimes V^{(\nu)} &= \oplus_\lambda \Bigl( \overbrace{V^{(\lambda)} \oplus \cdots \oplus V^{(\lambda)}}^{N_{\mu\nu}^{\lambda}\ times} \Bigr) \\ &= \oplus_\lambda N_{\mu\nu}^{\lambda} V^{(\lambda)} \end{aligned} \tag{11.364}$$

Recall that

$$\chi_{V_1 \otimes V_2}(g) = \chi_{V_1}(g) \chi_{V_2}(g) \tag{11.365}$$

Therefore, taking the trace of (11.364) we get a formula for the fusion coefficients:

$$\chi_\mu(g)\chi_\nu(g) = \sum_\lambda N_{\mu\nu}^{\lambda} \chi_\lambda(g) \tag{11.366}$$

Now, taking the inner product we get:

$$N_{\mu\nu}^{\lambda} = \langle \chi_\lambda, \chi_\mu \chi_\nu \rangle \tag{11.367}$$

There is a very beautiful explicit formula for the fusion coefficients in the case of finite groups. In this case we can write:

$$N_{\mu\nu}^{\lambda} = \frac{1}{|G|} \sum_{g \in G} \chi_\mu(g) \chi_\nu(g) \chi_\lambda(g^{-1}) \tag{11.368}$$

This can be written in a different way. For simplicity choose unitary irreps, and recall that the orthogonality relations on characters is equivalent to the statement that

$$S_{\mu i} := \sqrt{\frac{m_i}{|G|}} \chi_\mu(C_i) \qquad \mu = 1, \dots, r \quad i = 1, \dots, r \tag{11.369}$$

is a unitary matrix. Let 1 denote the trivial representation or the trivial conjugacy class. Then $S_{1i} = \sqrt{\frac{m_i}{|G|}}$. Thus we can write

$$N_{\mu\nu}^{\lambda} = \sum_i \frac{S_{\mu i} S_{\nu i} S_{\lambda i}^*}{S_{1i}} \tag{11.370}$$

This is a prototype of a celebrated result in conformal field theory known as the "Verlinde formula."

Equations (11.367)(11.368)(11.370) give a very handy way to get the numbers $N_{\mu\nu}^{\lambda}$. Note that, by their very definition the coefficients $N_{\mu\nu}^{\lambda}$ are nonnegative integers, although this is hardly obvious from, say, (11.370).

**Example 1**: Consider the irreps $\rho_m$ of $\mathbb{Z}/N\mathbb{Z}$. We then have the fusion rules:

$$\rho_m \otimes \rho_n \cong \rho_{m+n} \tag{11.371}$$

Recall that $\rho_m$ only depends on the integer $m \mathrm{mod} N$. The $S$-matrix above is the finite Fourier transform matrix:

$$S_{\mu i} = \sqrt{\frac{1}{|G|}} e^{2\pi \mathrm{i} \mu i / N} \tag{11.372}$$

from which one easily verifies (11.370).

**Remarks**

1. Denote the intertwiners from $V^\lambda$ to $V^\mu \otimes V^\nu$ by:

$$\mathcal{D}_{\mu\nu}^{\lambda} := \mathrm{Hom}_G(V^\lambda, V^\mu \otimes V^\nu) \tag{11.373}$$

Note that we can decompose the triple product $V^\mu \otimes V^\nu \otimes V^\lambda$ in two natural ways leading to the isomorphism

$$\oplus_\rho \mathcal{D}_{\mu\nu}^{\rho} \otimes \mathcal{D}_{\rho\lambda}^{\kappa} \cong \oplus_\rho \mathcal{D}_{\mu\rho}^{\kappa} \otimes \mathcal{D}_{\nu\lambda}^{\rho} \tag{11.374}$$

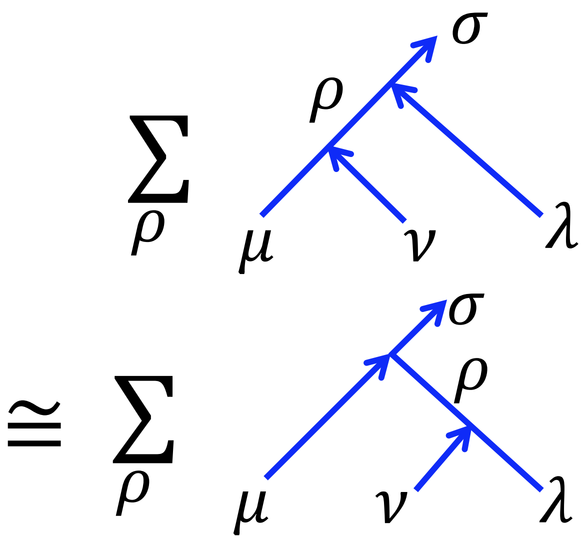



**Figure 32:** An important identity on the degeneracies in tensor products of irreducible representations, denoted pictorially. The trivalent vertex corresponds to the degeneracy $\mathcal{D}^{\lambda}_{\mu\nu}$. This identity is a prototype for an important identity in conformal field theory known as the "bootstrap equation." The choice of the isomorphism between the two lines itself satisfies an important relation known as the "pentagon identity."

This identity can be pictured as in Figure 32. If we choose bases for the spaces of intertwiners then the components of the isomorphism relative to these bases are called *fusion matrices* and they satisfy some nice identities.

****************************

EXPLAIN THE PENTAGON IDENTITY

*****************************

2. The considerations of this section lead rather naturally to the beautiful subject of Frobenius algebras and 2d topological quantum field theory. See Chapter 4 section **** for more about this.

3. The various identities satisfied by the fusion matrix and the $S$-matrix have remarkable analogues in the subject of 2d rational conformal field theory. For more about this see:

   a.) G. Moore and N. Seiberg, "Classical and Quantum Conformal Field Theory," Commun. Math. Phys. **123**(1989)177

   b.) G. Moore and N. Seiberg, "Lectures on Rational Conformal Field Theory," in *Strings '89*,Proceedings of the Trieste Spring School on Superstrings, 3-14 April 1989, M. Green, et. al. Eds. World Scientific, 1990. Available on G. Moore's home page.

   c.) Philippe Di Francesco, Pierre Mathieu, David Senechal, *Conformal Field Theory*, Springer

   d.) Jurgen Fuchs, *Affine Lie Algebras And Quantum Groups*

♣Need to add other references ♣

---

**Exercise** *Fusion Algebra For $S_3$*
Let $V^+, V^-, V^2$ denote the irreps of $S_3$ of dimensions $1, 1, 2$. Show that

$$\begin{aligned} V^+ \otimes V^\mu &\cong V^\mu \\ V^- \otimes V^- &\cong V^+ \\ V^- \otimes V^2 &\cong V^2 \\ V^2 \otimes V^2 &\cong V^+ \oplus V^- \oplus V^2 \end{aligned} \tag{11.375}$$

---

---

**Exercise** *Identities For Fusion Coefficients*
a.) Show that $N_{\mu\nu}^\lambda = N_{\nu\mu}^\lambda$
b.) Show that $N_{\mu 1}^\lambda = \delta_\mu^\lambda$
c.) For fixed $\mu$ regard $N_{\mu\nu}^\lambda$ as a matrix. It gives the matrix for tensor product with $V^\mu$. Show that this matrix is diagonalized by $S_{\mu i}$:

$$\sum_\lambda N_{\mu\nu}^\lambda S_{\lambda j} = \frac{S_{\mu j} S_{\nu j}}{S_{1j}} \tag{11.376}$$

$$\sum_{\nu,\lambda} S_{\nu i}^* N_{\mu\nu}^\lambda S_{\lambda j} = \delta_{ij} \left( \frac{S_{\mu j}}{S_{1j}} \right) \tag{11.377}$$

d.) Show that

$$\frac{S_{\mu 1}}{S_{11}} = \mathrm{dim} V^\mu \tag{11.378}$$

---

### 11.14 An Algebraic Viewpoint: The Group Algebra

There are some very interesting algebraic structures that are illustrated by the above results. In this subsection we explore them a little bit.

We first note that the regular representation is an algebra. Indeed, the space of complex-valued functions $Fun(X)$ on *any* space $X$ is an algebra under pointwise multiplication:

$$(f_1 *_1 f_2)(x) := f_1(x) f_2(x) \tag{11.379}$$

Note that this is a commutative algebra. In particular, we can apply this remark to $X = G$. For infinite groups we do not need to worry about the $L^2$ condition. For infinite groups we need to be more careful about the conditions on the functions. Many possibilities exist. For example, one can consider the functions with compact support. We will not explore this systematically.

When the domain of the function space is a group there is another kind of product we can define, the convolution product. If $f_1, f_2$ are functions on a group with a Haar measure then we can define

$$\begin{aligned}(f_1 *_2 f_2)(g) &:= \int_{G\times G} f_1(h_1)f_2(h_2)\delta_G(h_1h_2,g)dh_1dh_2 \\ &= \int_G f_1(h)f_2(h^{-1}g)dh\end{aligned} \tag{11.380}$$

Note that while the product $*_1$ is commutative, the convolution product $*_2$ is not commutative if the group $G$ is nonabelian. The proper domain for the convolution product is the space of $L^1$ functions. Again, we will not go into this.

We now specialize to a finite group. There is another nice way to think about the algebraic structure in terms of the *group ring*, usually denoted $\mathbb{Z}[G]$. As an Abelian group $\mathbb{Z}[G]$ is the free abelian group generated by $G$. That is, it is the space of formal linear combinations

$$\mathbb{Z}[G] := \{x = \sum_{g\in G} x(g)g : x(g) \in \mathbb{Z}\} \tag{11.381}$$

The ring structure is defined by taking

$$\begin{aligned}\left(\sum_{g\in G} x(g)g\right) *_2 \left(\sum_{g\in G} y(g)g\right) &:= \sum_{g,h\in G} x(g)y(h) \qquad \overbrace{g\cdot h}^{\text{group multiplication}} \\ &= \sum_{k\in G}\left[\sum_{g,h\in G:gh=k} x(g)y(h)\right]k\end{aligned} \tag{11.382}$$

Replacing $\mathbb{Z} \to \mathbb{C}$ (more properly, taking the tensor product with $\mathbb{C}$) gives the group algebra $\mathbb{C}[G]$.

As a vector space $\mathbb{C}[G]$ is naturally dual to the regular representation: $R_G \cong \mathbb{C}[G]^\vee$. The natural pairing is

$$\langle f, x\rangle := \sum_{g\in G} x(g)f(g) \tag{11.383}$$

Of course, there is a natural basis for $\mathbb{C}[G]$, namely the elements $g$ themselves, and the dual basis is the basis of "delta functions"

$$\delta_g(h) := \begin{cases} 1 & h = g \\ 0 & else \end{cases} \tag{11.384}$$

Having chosen a basis $\mathbb{C}[G]$ is isomorphic to $L^2(G)$. The isomorphism takes $f : G \to \mathbb{C}$ to $\sum_g f(g)g$.

As we have seen, there are two natural algebra structures on $L^2(G)$ and hence, there are two natural algebra structures on the group algebra $\mathbb{C}[G]$. Written out in the basis $g$ they are

$$g_1 *_1 g_2 = \begin{cases} g_1 & g_1 = g_2 \\ 0 & g_1 \neq g_2 \end{cases} \tag{11.385}$$

$$g_1 *_2 g_2 = g_1 g_2 \tag{11.386}$$

and we can then extend by linearity.

**Example** $\mathbb{Z}_2 = \{1, \sigma\}$. The group algebra is

$$\mathbb{C}[\mathbb{Z}_2] = \{a \cdot 1 + b \cdot \sigma : a, b \in \mathbb{C}\} \tag{11.387}$$

Now the algebra structure is:

$$(a \cdot 1 + b \cdot \sigma) *_2 (a' \cdot 1 + b' \cdot \sigma) = (aa' + bb')1 + (ba' + ab')\sigma \tag{11.388}$$

**Remark**: The above algebra is sometimes referred to as the "double-numbers," a lesser known cousin of the complex numbers. Note that the algebra of complex numbers is isomorphic, as a real algebra to the algebra of pairs $(a, b)$ of real numbers with multiplication law:

$$(a, b) \cdot (a', b') = \big((aa' - bb'), (ba' + ab')\big) \tag{11.389}$$

---

**Exercise**
Which elements of $\mathbb{C}[\mathbb{Z}_2]$ are invertible?

---

### 11.14.1 Explicit Decomposition Of The Regular Representation

For this subsubsection assume that $G$ is a finite group.

The group algebra is of course a representation of $G_{\rm left} \times G_{\rm right}$ acting as a left action:

$$T^{\rm reg.rep.}(g_1, g_2) \cdot x = g_1 x g_2^{-1} \tag{11.390}$$

for $x \in \mathbb{C}[G]$.

We know from the Peter-Weyl theorem that this representation is highly reducible. We can now construct a very explicit isomorphism of $L^2(G)$ with $\oplus_\mu {\rm End}(V^\mu)$.

Choose explicit bases $\{w_i^{(\mu)}\}$ for the distinct irreps $V^{(\mu)}$ so that we have matrix elements $M_{ij}^\mu(g)$ associated to the group elements $g \in G$. Let us define:

$$\mathcal{P}_{ij}^\mu \in \mathbb{C}[G] \qquad 1 \le i, j \le n_\mu \tag{11.391}$$

by

$$\mathcal{P}_{ij}^\mu := \frac{n_\mu}{|G|} \sum_{g \in G} \hat{M}_{ij}^{(\mu)}(g) g \tag{11.392}$$

Here

$$\hat{M}_{ij}^{(\mu)}(g) := M_{ji}^{(\mu)}(g^{-1}) \tag{11.393}$$

is the matrix element in the dual representation $(V^{(\mu)})^\vee$. If we choose an ON basis so that the matrices are unitary then $\hat{M}_{ij}^{(\mu)}(g) = (\hat{M}_{ij}^{(\mu)}(g))^*$.

The elements $\mathcal{P}^{\mu}_{ij}$ satisfy the simple product law:

$$\mathcal{P}^{\mu}_{ij}\mathcal{P}^{\mu'}_{i'j'} = \delta_{\mu\mu'}\delta_{ji'}\mathcal{P}^{\mu}_{ij'} \tag{11.394}$$

Moreover, one can show that:

$$g \cdot \mathcal{P}^{\mu}_{ij} = \sum_{s} M^{(\mu)}_{si}(g)\mathcal{P}^{\mu}_{sj} \tag{11.395}$$

$$\mathcal{P}^{\mu}_{ij} \cdot g^{-1} = \sum_{s} \hat{M}^{(\mu)}_{sj}(g)\mathcal{P}^{\mu}_{is} \tag{11.396}$$

and thus the $\mathcal{P}^{\mu}_{ij}$ explicitly transform like the matrix units $(w^{\mu}_i)^{\vee} \otimes w^{\mu}_j$ in $\mathrm{End}(V^{\mu})$. On the other hand they also form a basis for $\mathbb{C}[G] \cong L^2(G)$.

The reader will no doubt have recognized the close affinity to the operators (11.270) and their algebra (11.271). Indeed, given any representation $(T, V)$ of $G$ we can apply

$$T(\mathcal{P}^{\mu}_{ij}) := \frac{n_{\mu}}{|G|}\sum_{g\in G} \hat{M}^{(\mu)}_{ij}(g)T(g) = P^{\mu}_{ij} \in \mathrm{End}(V) \tag{11.397}$$

to recover the operators (11.270). So the above elements of the group algebra apply universally to all representations.

---

**Exercise**

Prove (11.394)(11.395) and (11.396) using the orthogonality of matrix elements in irreps.

---

### 11.14.2 Class Functions And The Center Of The Group Algebra

As we have noted, the group algebra $\mathbb{C}[G]$ with the convolution product is a noncommutative algebra if $G$ is nonabelian.

A little thought shows that the *center* of the group algebra $Z[\mathbb{C}[G]]$ is spanned by the elements

$$c_i = \sum_{g\in\mathcal{C}_i} g \tag{11.398}$$

where we sum over the conjugacy class $\mathcal{C}_i$ of $g$. This simply follows from

$$gc_ig^{-1} = \sum_{h\in\mathcal{C}_i} ghg^{-1} = c_i \tag{11.399}$$

Correspondingly, the center of $R_G$ under the convolution product are the class functions.

It is interesting to write out the two products in these two natural bases for the center of $\mathbb{C}[G]$. In the pointwise product we have:

$$\chi_{\mu} *_1 \chi_{\nu} = \sum_{\lambda} N^{\lambda}_{\mu\nu}\chi_{\lambda} \tag{11.400}$$

while

$$\delta_{\mathcal{C}_i} *_1 \delta_{\mathcal{C}_j} = \delta_{ij}\delta_{\mathcal{C}_i} \tag{11.401}$$

So the $\delta_{\mathcal{C}_i}$ are idempotents, but the $\chi_\mu$ have a more complicated multiplication law.

On the other hand, we can interpret the exercise with equation (11.317) above as saying that the characters satisfy:

$$\chi_\mu *_2 \chi_\nu = \frac{\delta_{\mu\nu}}{n_\mu}\chi_\mu \tag{11.402}$$

So the characters are idempotents with this product. On the other hand, the convolution product is nondiagonal in the conjugacy class basis $\delta_{C_i}$ that corresponds to the elements $c_i$ in the group algebra:

$$\delta_{C_i} *_2 \delta_{C_j} = \sum_k M_{ij}^k \delta_{C_k} \tag{11.403}$$

where one can express the matrix $M_{ij}^k$ by a formula similar to the Verlinde formula below.

So the unitary matrix $S_{\mu i}$ is a kind of Fourier transform which exchanges the product for which the characters of representations or characteristic functions of conjugacy classes is diagonalized:

$$\chi_\mu = \sum_i \sqrt{\frac{|G|}{m_i}} S_{i\mu}\delta_{C_i} \tag{11.404}$$

---

**Exercise**

Show that the structure constants in the basis $\delta_{C_i}$ in the convolution product can be written as:

$$M_{ij}^k = \frac{S_{0i}S_{0j}}{S_{0k}} \sum_\mu \frac{S_{i\mu}^* S_{j\mu}^* S_{k\mu}}{n_\mu} \tag{11.405}$$

---

---

**Exercise**

An algebra is always canonically a representation of itself in the left-regular representation. The representation matrices are given by the structure constants, thus

$$L(\chi_\mu)_\nu{}^\lambda = N_{\mu\nu}^\lambda \tag{11.406}$$

shows that the fusion coefficients are the structure constants of the pointwise multiplication algebra of class functions.

Note that since the algebra of class functions is commutative the matrices $L(\chi_\mu)$ can be simultaneously diagonalized.

Show that the unitary matrix $S_{i\mu}$ is in fact the matrix which diagonalizes the fusion coefficients.

---

### 11.14.3 Hopf Algebras

The above algebraic structures are closely related to the important topic of Hopf algebras. We motivate the general definition by considering the algebra of (suitable) functions on a group.

Let $G$ be a topological group and $A = Fun(G)$ be some suitable algebra of $\kappa$-valued functions, where $\kappa$ is a field. First of all, let us note that $A$ is an algebra by pointwise multiplication. We will denote it by $\mu : A \otimes A \to A$ and of course it is associative:

$$\begin{array}{ccccc} & & A \otimes A & & \\ & \overset{\mu\otimes Id}{\nearrow} & & \overset{\mu}{\searrow} & \\ A \otimes A \otimes A & & & & A \\ & \underset{Id\otimes\mu}{\searrow} & & \underset{\mu}{\nearrow} & \\ & & A \otimes A & & \end{array} \tag{11.407}$$

But now there will be extra structure on $A$ arising from the fact that the group $G$ has extra structure. In particular there is a multiplication

$$m : G \times G \to G \tag{11.408}$$

on the group $G$. This induces a *comultiplication* on $A$:

$$\Delta : A \to A \otimes A \tag{11.409}$$

To see why, after a little thought one can prove that:

$$A \otimes A = Fun(G) \otimes Fun(G) \cong Fun(G \times G) \tag{11.410}$$

Now, suppose $f \in A$. We then need to define $\Delta(f)$ as a function on $G \times G$. The multiplication on $G$ gives us a natural way to do this:

$$\Delta(f)(g_1, g_2) := f(g_1 g_2) \ . \tag{11.411}$$

Now the associativity of the group multiplication on $G$ is equivalent to the commutativity of the diagram:

$$\begin{array}{ccccc} & & G \times G & & \\ & \overset{m\times Id}{\nearrow} & & \overset{m}{\searrow} & \\ G \times G \times G & & & & G \\ & \underset{Id\times m}{\searrow} & & \underset{m}{\nearrow} & \\ & & G \times G & & \end{array} \tag{11.412}$$

The induced diagram on $A$ reverses all arrows and is the property of *coassociativity* of $\Delta$:

$$\begin{array}{ccccc} & & A \otimes A & & \\ & \overset{\Delta}{\nearrow} & & \overset{\Delta\otimes Id}{\searrow} & \\ A & & & & A \otimes A \otimes A \\ & \underset{\Delta}{\searrow} & & \underset{Id\otimes\Delta}{\nearrow} & \\ & & A \otimes A & & \end{array} \tag{11.413}$$

This makes $A$ a *coassociative coalgebra*. The coassociativity of $\Delta$ on $A$ is completely equivalent to the associativity of the group multiplication on $G$.

Let us now translate the other group axioms into statements about $A$.

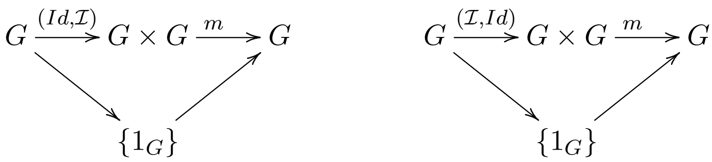
Now, the next group axiom postulates a unit $1_G$. The dual of this is the *counit* $\varepsilon : A \to \kappa$ (where $\kappa$ is the ground field). For $A = Fun(G)$ we would define $\varepsilon(f) := f(1_G)$. We have two diagrams expressing the properties of the unit:


(11.414)

The dual diagrams give the property of the counit:

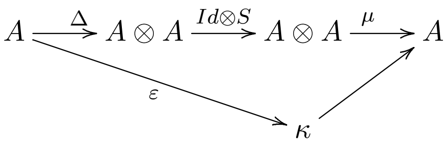

(11.415)

The final group axiom states that every group element has an inverse. If we say that $\mathcal{I} : G \to G$ is the map $g \mapsto g^{-1}$ then we can define a dual operation $S : A \to A$ by $S(f) = f \circ \mathcal{I}$. The linear operator $S$ is known as the *antipode*. Now the group axiom is the pair of diagrams:




(11.416)

Dually we get




(11.417)

and a second diagram with $S \otimes Id$.

This motivates the general definition:

**Definition** A unital algebra $A$ over $\kappa$ equipped with multiplication $\mu$, comultiplication $\Delta$, counit $\varepsilon$, and antipode $S$ satisfying equations (11.407),(11.413), (11.415),(11.417), is called a *Hopf algebra*.

We stress that this is a general concept. The algebra of functions on a group can be given a Hopf algebra structure, but, as we shall soon see, this is not the general Hopf algebra.

**Remark:** Now let us note that, quite generally, if $A$ is a Hopf algebra then the vector space dual $A^\vee := \mathrm{Hom}_\kappa(A,\kappa)$ is also a Hopf algebra. First, let us define the product

$$\mu^\vee : A^\vee \otimes A^\vee \to A^\vee \tag{11.418}$$

If $\ell_1, \ell_2$ are two linear functionals then we define their product $\mu^\vee(\ell_1 \otimes \ell_2)$ by declaring that the value on $a$ is obtained from forming $\ell_1 \otimes \ell_2(\Delta(a))$ and then using the multiplication $\kappa \otimes \kappa \to \kappa$. In more detail, suppose $a_i$ is a linear basis for $A$, and suppose

$$\Delta(a_i) = \sum_{j,k} \Delta_i^{jk} a_j \otimes a_k \tag{11.419}$$

where $\Delta_i^{jk} \in \kappa$ Then $\mu^\vee(\ell_1 \otimes \ell_2) \in A^\vee$ is defined by

$$\mu^\vee(\ell_1 \otimes \ell_2)(a_i) := \sum_{j,k} \Delta_i^{jk} \ell_1(a_j)\ell_2(a_k). \tag{11.420}$$

Similarly, the dual comultiplication on $A^\vee$ is defined by

$$\Delta^\vee(\ell)(a_1 \otimes a_2) := \ell(\mu(a_1 \otimes a_2)) \tag{11.421}$$

The dual counit is

$$\varepsilon_{A^\vee}(\ell) := \ell(1_A) \tag{11.422}$$

and the dual antipode is simply

$$S_{A^\vee}(\ell)(a) := \ell(S(a)) \tag{11.423}$$

We leave it to the reader to check that $(\mu^\vee, \Delta^\vee, \varepsilon_{A^\vee}, S_{A^\vee})$ in fact define a Hopf algebra structure on $A^\vee$.

Applying the above remark to our example of $A = Fun(G)$ we obtain the group algebra $A^\vee = \kappa[G]$. At least formally, this can be viewed as the linear span of $\mathbf{ev}_g : A \to \kappa$ given by $\mathbf{ev}_g(f) = f(g)$. Now the multiplication on $\kappa[G]$ is:

$$\mu^\vee(\mathbf{ev}_{g_1} \otimes \mathbf{ev}_{g_2}) = \mathbf{ev}_{g_1 g_2} \tag{11.424}$$

while the comultiplication is:

$$\Delta^\vee(\mathbf{ev}_g)(f_1 \otimes f_2) = f_1(g) f_2(g) \tag{11.425}$$

and hence

$$\Delta^\vee(\mathbf{ev}_g) = \mathbf{ev}_g \otimes \mathbf{ev}_g \tag{11.426}$$

The counit is

$$\varepsilon^\vee(\mathbf{ev}_g) = 1 \qquad \forall g \in G \tag{11.427}$$

and the antipode is

$$S(\mathbf{ev}_g) = \mathbf{ev}_{g^{-1}} \tag{11.428}$$

Now, rather confusingly, $A^\vee$ as a vector space can also be identified with an algebra of $\kappa$-valued functions on the group $G$ since we can write the general element as

$$\sum_{g\in G} f_g \mathbf{ev}_g \tag{11.429}$$

and $g \mapsto f_g$ is a function on the group. [170] However, viewed this way, the product $\mu^\vee$ is the convolution product:

$$\mu^\vee(f_1 \otimes f_2)(g) = \int_G f_1(h) f_2(gh^{-1}) dh \tag{11.430}$$

where $dh$ is a Haar measure of volume one, while $\Delta^\vee$ takes $g \mapsto f_g$ to a function on $G \times G$ given by

$$\Delta^\vee(f)(g_1, g_2) = f_{g_1} \delta_{g_1,g_2} \tag{11.431}$$

In general a Hopf algebra $B$ is said to be "cocommutative" if $\sigma \circ \Delta = \Delta$ where $\sigma : B \otimes B \to B \otimes B$ is the permutation operator.

♣Write commutative diagram ♣

The above two examples $A = Fun(G)$ with the pointwise product and $A^\vee = \kappa[G]$ with the convolution product have one property that does not hold for general Hopf algebras: $A$ is commutative and $A^\vee$ is cocommutative.

Of course, while $A = Fun(G)$ is commutative, it is not co-commutative when $G$ is noncommutative. Dually, $\kappa[G]$ is not commutative, when $G$ is noncommutative, but it is always cocommutative.

There are other examples of natural Hopf algebras associated to Lie algebras and groups. Hopf algebras have played an important role in topology. In the past few decades, starting with an important address, "Quantum Groups," by V. Drinfeld at the 1986 International Conference of Mathematicians, there has been a great deal of attention focused on them because of their relations to integrable systems and quantum invariants of manifolds.

#### 11.14.4 Frobenius Algebras

### 11.15 Representations Of The Symmetric Group

#### 11.15.1 Standard Tableaux And Their Projection Operators

In subsubsection 7.5.2 we introduced the idea of a Young diagram and its relation to partitions. They were used to give a pictorial representation of conjugacy classes in the symmetric group.

In this section we explain that Young diagrams can also be used to define projection operators onto the irreducible representations of $S_n$ in the group algebra $\mathbb{C}[S_n]$. We only give the barest minimal description here. See the references at the end of this section for a full account.

[170] We will not be careful here about the precise class of functions. For example, if we consider finite sums and $G$ is a continuous group then the relevant functions would be discontinuous, and would only be nonzero at a finite set of points.

Define a *Young tableau* to be a Young diagram in which each of the integers $1, 2, \ldots, n$ has been placed into a box. These integers can be placed in any order. For any fixed Young diagram there are $n!$ different Young tableau.

A *standard tableau* is one in which the integers are increasing in each row, left to right, and in each column, top to bottom. In general there are many fewer standard tableau than $n!$. For example, a Young diagram that is a horizontal or vertical row of boxes has exactly one standard tableau. Indeed, the number of standard tableau will give dimensions of irreps of $S_n$.

Let $T$ denote a Young tableau. Then we define two subgroups of $S_n$, $\mathcal{R}(T)$ and $\mathcal{C}(T)$. $\mathcal{R}(T)$ consists of permutations in $S_n$ which only permute numbers within each row of $T$. Similarly, $\mathcal{C}(T)$ consists of permutations which only permute numbers within each column of $T$.

For a Young tableau $T$ define the following elements in the group ring $\mathbb{Z}[S_n]$:

$$R(T) := \sum_{g \in R(T)} g \tag{11.432}$$

$$C(T) := \sum_{g \in C(T)} \epsilon(g) g \tag{11.433}$$

$$P(T) := R(T)C(T) = \sum_{p \in R(T), q \in C(T)} \epsilon(q) pq \tag{11.434}$$

Note that we have the following easy fact: If $T, T'$ are two Young tableau based on the same diagram then there is a permutation $\sigma \in S_n$ so that $P(T') = \sigma P(T) \sigma^{-1}$.

Then we have the following rather nontrivial statements:

1. $P(T)^2 = c_T P(T)$ where $c_T$ is an integer.
2. $P(T)P(T') = 0$ if $T$ and $T'$ correspond to different partitions.
3. If $T, T'$ are two different standard tableaux corresponding to the same partition then $P(T)P(T') = 0$.
4. $P(T)$ projects onto an irreducible representation of $S_n$. That is $\mathbb{C}[S_n] \cdot P(T)$ transforms as an irreducible representation $\mathcal{R}(T)$ of $S_n$ in the left-regular representation.
5. Moreover, all of the irreducible representations of $S_n$ are equivalent to one representations of the form $\mathbb{C}[S_n] \cdot P(T)$ for some $T$. Representations $\mathcal{R}(T)$ corresponding to different Young diagrams are inequivalent. As we noted above for two tableaux $T, T'$ based on the same Young diagram there is a $\sigma \in S_n$ so that $P(T') = \sigma P(T) \sigma^{-1}$. It follows that $\mathbb{C}[S_n] \cdot P(T')$ and $\mathbb{C}[S_n] \cdot P(T)$ are equivalent representations.

   *Thus the (isomorphism classes of) irreducible representations of $S_n$ are in 1-1 correspondence with the Young diagrams with $n$ boxes.*

   **Remark**: Recall that we also noted earler that Young diagrams are also in 1-1 correspondence with conjugacy classes in $S_n$. In general, although the number of irreducible representations and conjugacy classes of a finite group are the same, there is no natural 1-1 correspondence between these two sets. However, thanks to Pontryagin self-duality, there is such a correspondence for cyclic groups, and, because of the above facts, there is also such a correspondence for the symmetric groups.

6. The integer $c_T$ in $P(T)^2 = c_T P(T)$ is given by

$$c_T = \frac{n!}{\dim R(T)} \tag{11.435}$$

$\dim R(T)$ can also be characterized as the number of standard tableau corresponding to the underlying partition.

7. Another formula for $\dim R(T)$ is the *hook length formula.* For a box in a Young diagram define its hook length to be the number of squares to the right in its row, and underneath in its column, counting that box only once. Then

$$\dim R(T) = \frac{n!}{\prod \text{hooklengths}} \tag{11.436}$$

Proofs of these rather nontrivial statements can be found in several textbooks on group representation theory:

1. For further discussion of the above material see the books by Miller, Hammermesh, Curtis+Reiner, Fulton+Harris *Representation Theory.*

2. There is a second, elegant method for constructing the irreps of the symmetric group using induced representations. See the book by Sternberg for an account.

We will content ourselves with illustrating these statements for $S_3$ and $S_4$.

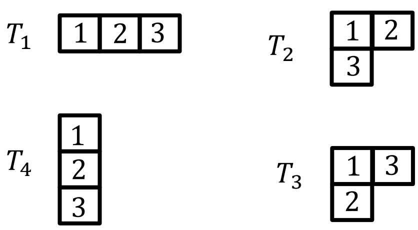



**Figure 33:** Four standard Young tableaux for $S_3$.

Example 1: $G = S_3$

There are 3 partitions, 3 Young diagrams, and 4 standard tableaux labeled $T_i$ and shown in 33.

Clearly $P(T_1) = \sum_{p \in S_3} p$ and $P(T_1)^2 = 6P(T_1)$. Moreover $\mathbb{Z}[S_3] \cdot P(T_1)$ is one-dimensional. This is the trivial representation.

Similarly, $P(T_4) = \sum \epsilon(q)q$ spans the one-dimensional sign representation.

There are two standard tableaux corresponding to $\lambda = (2, 1)$ with

$$P(T_2) = (1 + (12))(1 - (13)) = 1 + (12) - (13) - (132) \tag{11.437}$$

$$P(T_3) = (1 + (13))(1 - (12)) = 1 - (12) + (13) - (123) \tag{11.438}$$

One checks that $P(T_2)^2 = 3P(T_2)$ and $P(T_3)^2 = 3P(T_3)$ and $P(T_2)P(T_3) = 0$. These are consistent with the above statements about $c_T$ and the hooklength formula.

Now consider $\mathbb{Z}[S_3] \cdot P(T_2)$. We compute

$$\begin{aligned} 1 \cdot P(T_2) &:= v_1 \\ (12) \cdot P(T_2) &= v_1 \\ (13) \cdot P(T_2) &= -1 + (13) - (23) + (123) := v_2 \\ (23) \cdot P(T_2) &= -(12) + (23) - (123) + (132) = -v_1 - v_2 \\ (123) \cdot P(T_2) &= -1 + (13) - (23) + (123) = v_2 \\ (132) \cdot P(T_2) &= -(12) + (23) - (123) + (132) = -v_1 - v_2 \end{aligned} \tag{11.439}$$

Thus the space is two-dimensional.

From this it is easy to compute: $(12) \cdot v_2 = -v_1 - v_2$, $(13)v_2 = v_1$ and $(23)v_2 = (123)P(T_2) = v_2$ and so in this basis we have a representation generated by

$$\rho(12) = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} \qquad \rho(13) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \rho(23) = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \tag{11.440}$$

One easily checks that the character is the same as that we computed above for the **2**.

Of course, similar statements hold for $P(T_3)$, which the reader should check as an exercise.

Example 2: $G = S_4$

Similarly, the standard tableaux for $S_4$ are shown in Figure 34. The reader should verify these, construct the irreducible subrepresentations of $\mathbb{C}[S_4]$, and write out the character table for $S_4$.

### 11.16 Schur-Weyl Duality And The Irreps Of $GL(d, \kappa)$

Consider a vector space $V$ of dimension $d$. In previous sections we have seen that

$$V^{\otimes n} \equiv V \otimes V \otimes \cdots \otimes V \tag{11.441}$$

is a representation of $S_n$ with the group action defined by permuting the factors:

$$\sigma \cdot \ (w_1 \otimes w_2 \otimes \cdots \otimes w_n) := w_{\sigma(1)} \otimes w_{\sigma(2)} \otimes \cdots \otimes w_{\sigma(n)} \tag{11.442}$$

Now if $(T, V)$ is a representation of a group $G$ then $V^{\otimes n}$ is also a representation of $G$ via

$$T^{\otimes n}(g)(w_1 \otimes w_2 \otimes \cdots \otimes w_n) := T(g)(w_1) \otimes T(g)(w_2) \otimes \cdots \otimes T(g)(w_n) \tag{11.443}$$

**Figure 34:** 10 standard Young tableaux for $S_4$.

Note that the group actions by $\sigma \in S_n$ and $g \in G$ commute so that $V^{\otimes n}$ is naturally a representation of the direct product $G \times S_n$.

Considered as a representation of $S_n$, we have complete reducibility (because $S_n$ is compact):

$$V^{\otimes n} \cong \oplus_\lambda \mathcal{V}_\lambda \otimes R_\lambda \tag{11.444}$$

where $R_\lambda$ is the irrep of $S_n$ corresponding to the partition $\lambda$. By Schur's lemma we can characterize the multiplicity space $\mathcal{V}_\lambda$ as

$$\mathcal{V}_\lambda = \mathrm{Hom}_{S_n}(R_\lambda, V^{\otimes n}). \tag{11.445}$$

This is the vector space of $S_n$-equivariant maps from $R_\lambda$ to $V^{\otimes n}$. Because the $G$ and $S_n$ actions commute we know that $\mathcal{V}_\lambda$ will also be a representation of $G$.

Let us apply the above setup to the case that $G = GL(d, \kappa)$ with $\kappa = \mathbb{R}$ or $\mathbb{C}$ where $V = \kappa^d$ is the defining representation. Notice that $(T, V)$ is an irreducible representation. The beautiful result of Schur and Weyl – known as the Schur-Weyl duality theorem - states that in this case the $\mathcal{V}_\lambda$ are irreducible representations, and we can construct all the irreducible representations of $G$ in this way:

**Theorem** The representations $\mathcal{V}_\lambda$ are irreducible representations of $GL(d, \kappa)$ and, up to a power of the determinant representation, all irreducible representations may be obtained in this way. Moreover, these statements hold for $U(d) \subset GL(d, \mathbb{C})$.

*Proof* For complete proofs see

R. Carter, G. Segal, I. MacDonald, *Letures on Lie Groups and Lie Algebras*, Lie Groups, chapter 13.

W. Fulton and J. Harris, *Representation Theory*, chapter 6

We can see fairly easily why, if we consider all possible values of $n$, all the representations of $U(d)$ must occur, provided we are allowed to multiply by arbitrary powers of the determinant representation. The matrix elements of the representations $(\det V)^{\otimes \ell} \otimes V^{\otimes n}$ where we consider all $\ell \in \mathbb{Z}$ and all $n \geq 0$ will generate an algebra $\mathbb{C}[a_{ij}, (\det A)^{-1}]$ on $U(d)$. Note that this is the same as the algebra $\mathbb{C}[a_{ij}, (a_{ij})^*]$ since, by unitarity $(a_{ij})^*$ is a polynomial in $a_{ij}$ divided by $\det A$. But $\mathbb{C}[a_{ij}, (a_{ij})^*]$ is dense in $L^2(U(d))$, and, by the Peter-Weyl theorem, any representations not accounted for by $(\det V)^{\otimes \ell} \otimes V^{\otimes n}$ would have to have matrix elements orthogonal to the closure of this subspace. Therefore, all irreducible representations of $U(d)$ are obtained as subrepresentations of $(\det V)^{\otimes \ell} \otimes V^{\otimes n}$.

Now let us prove that $\mathcal{V}_\lambda$ is irreducible. Since $R_\lambda$ is an irreducible representation of $S_n$, Schur's lemma says that the algebra of operators on $V^{\otimes n}$ that commutes with $S_n$ is:

$$\mathrm{End}_{S_n}(V^{\otimes n}) \cong \oplus_\lambda (\mathrm{End}(\mathcal{V}_\lambda)) \otimes 1_{R_\lambda} \tag{11.446}$$

where $\mathrm{End}(\mathcal{V}_\lambda)$ is the space of all linear transformations on $\mathcal{V}_\lambda$. Now, consider the algebra generated just by the operators $T(g)$. If, for some $\lambda$, the $\mathcal{V}_\lambda$ were reducible, say

$$\mathcal{V}_\lambda \cong \mathcal{V}_\lambda^1 \oplus \mathcal{V}_\lambda^2 \tag{11.447}$$

for two nontrivial representations $\mathcal{V}_\lambda^i$, $i = 1, 2$, then the algebra of operators generated by $T(g)$ would, in this summand, only generate matrices of block diagonal form. In particular, they would not generate the whole algebra $\mathrm{End}(\mathcal{V}_\lambda)$.

Now we use some linear algebra. We have a linear isomorphism: $\psi$

$$\psi : \mathrm{End}(V^{\otimes n}) \to (\mathrm{End}(V)))^{\otimes n} \tag{11.448}$$

because

$$\mathrm{End}(V^{\otimes n}) \cong (V^{\otimes n})^\vee \otimes V^{\otimes n} \cong (V^\vee \otimes V)^{\otimes n} \cong (\mathrm{End}(V))^{\otimes n} \tag{11.449}$$

Now $\mathrm{End}_{S_n}(V^{\otimes n})$, the subalgebra of operators that commute with the $S_n$ action under this isomorphism are mapped to the symmetric product

$$\psi : \mathrm{End}_{S_n}(V^{\otimes n}) \to \left((\mathrm{End}(V))^{\otimes n}\right)^{S_n} \cong S^n(\mathrm{End}(V)) \tag{11.450}$$

where $S^n W$ of a vector space $W$ is the $n^{th}$ symmetric power.

Next, note that, for any vector space $W$ we can map

$$\theta : W \to S^n W \tag{11.451}$$

by the map

$$\theta : w \mapsto w \otimes w \otimes \cdots \otimes w \tag{11.452}$$

We stress that this is <u>not</u> a linear map! In general the dimension of $W$ will be much smaller than the dimension of $S^n W$. Nevertheless, it is clear that the linear span of the set of elements $\theta(w)$ for $w \in W$ will in fact be all of $S^n W$.

♣Explain this last sentence in more detail. ♣

So, now consider the algebra generated by the operators $T^{\otimes n}(g)$ for all $g \in GL(V)$. (In our application $T(g) = g$ is the defining representation.) Under the isomorphism $\psi$ these operators are the image of the map $\theta$ applied to the operators $T(g) \in \text{End}(V)$. But $GL(V)$ is dense in $\text{End}(V)$. So the algebra generated by $T(g)$ is all of $\text{End}(V)$. Therefore, the algebra generated by $T^{\otimes n}(V)$ is all of $S^n(\text{End}(V)$. (Apply the previous paragraph with $W = \text{End}(V)$.) Now, by the linear isomorphism $\psi$ we learn that the algebra generated by the operators $T^{\otimes n}(g)$ for all $g \in GL(V)$ is all of $\text{End}_{S_n}(V^{\otimes n})$, and not just a proper subalgebra. This is a key step.

Since the algebra generated by $T^{\otimes n}(g)$ is all of $\text{End}_{S_n}(V^{\otimes n})$ it follows from our remark above when that algebra is projected to $\text{End}(\mathcal{V}_\lambda)$ must get the full matrix algebra $\text{End}(\mathcal{V}_\lambda)$, and therefore $\mathcal{V}_\lambda$ must be an irreducible representation. Put differently, the operators that commute with the algebra generated by $T^{\otimes n}(g)$ must be scalar operators on $\mathcal{V}_\lambda$ and therefore $\mathcal{V}_\lambda$ must be an irreducible representation of $GL(V)$. ♠

**Remark**: As a biproduct of this proof we have a nice description of some important subalgebras of $\text{End}((V^{\otimes n})$ and their commutants. The algebra of operators $\mathcal{A}$ generated by linear combinations of $\sigma$ acting on $V^{\otimes n}$ commutes with the algebra of operators $\mathcal{B}$ generated by $g \in GL(d, k)$ acting on $V^{\otimes n}$. What we have shown above is that these algebras are commutants of each other: The algebra of operators in $\text{End}(V^{\otimes n})$ commuting with $\mathcal{A}$ is precisely $\mathcal{B}$, and the algebra of operators commuting with $\mathcal{B}$ is precisely $\mathcal{A}$.

We can construct $\mathcal{V}_\lambda$ explicitly by taking a Young diagram and its corresponding symmetrizer $P(T)$ (for some tableau $T$) and taking the image of $P(T)$ acting on $V^{\otimes n}$. As representations of $GL(d, k)$ we have

$$P(T)V^{\otimes n} \cong P(T')V^{\otimes n} \tag{11.453}$$

if $T$ and $T'$ correspond to the same partition, i.e. the same Young diagram.

With some work (see e.g. the Fulton-Harris book) it can be shown that - as representations of $GL(d, \kappa)$ we have an orthogonal decomposition

$$V^{\otimes n} = \oplus_T P(T)V^{\otimes n} \tag{11.454}$$

where $T$ runs over the *standard* tableaux of $S_n$. In particular, this recovers fact that the dimension of $R_\lambda$ is the number of standard tableaux. Of course, tableaux with columns of length $\geq d$ will project to the zero vector space and can be omitted.

Tensors Of Definite Symmetry Type:

Finally, let us choose a basis $\{v_i\}$ for $V = \kappa^d$, then a typical element can be expanded in the basis as:

$$v = \sum_{i_1, i_2, \dots, i_n} t^{i_1, i_2, \dots, i_n} v_{i_1} \otimes \cdots \otimes v_{i_n} \tag{11.455}$$

We will often assume the summation convention where repeated indices are automatically summed. Under the action of $GL(d,k)$:

$$g \cdot v_i = g_{ji} v_j \tag{11.456}$$

we therefore have:

$$(g \cdot t)^{i_1 \cdots i_n} = t^{j_1 \cdots j_n} g_{j_1 i_1} \cdots g_{j_n i_n} \tag{11.457}$$

that is, elements in $V^{\otimes n}$ transform as tensors and $V^{\otimes n}$ is *also* a representation of $GL(d,\mathbb{R})$: On the other hand

$$(\sigma \cdot t)^{i_1 \ldots i_n} = t^{i_{\sigma^{-1}(1)} \cdots i_{\sigma^{-1}(n)}} \tag{11.458}$$

Thus the Young tableau operators $P(T)$ can be understood as projecting to *tensors of a definite symmetry type*, where a symmetry type is labeled by a representation of $S_n$. If $T$ is a Young tableau associated to partition $\lambda$ then

$$P(T) = \sum_{\sigma \in S_n} a_\sigma \sigma \tag{11.459}$$

the tensors of the symmetry type $\lambda$ are of the form

$$\tilde{t}^{i_1 \ldots i_n} = \sum_{\sigma \in S_n} a_\sigma t^{i_{\sigma^{-1}(1)} \cdots i_{\sigma^{-1}(n)}} \tag{11.460}$$

Schur-Weyl duality assures us that tensors of a definite symmetry type will transform in irreducible representations of $GL(d,\kappa)$.

**Examples**

1. Returning to equation (11.343) we saw that the isotypical decomposition of $V^{\otimes 2}$ as a representation of $S_2$ is:

$$V^{\otimes 2} = S^2 V \otimes \mathbf{1}^+ \oplus \Lambda^2 V \otimes \mathbf{1}^- \tag{11.461}$$

   Schur-Weyl duality says that for $V$ the defining representation of $GL(d,\kappa)$ or $U(d)$ the symmetric and antisymmetric powers are irreps.

2. We return to the discussion around equations (11.480) and (**??**) and discuss tensors which are neither totally symmetric nor totally antisymmetric but still have a definite symmetry type. Recall that if $V$ has dimension $d$ then $V^{\otimes 3}$ as a representation of $S_3$ has isotypical decomposition:

$$V^{\otimes 3} \cong \frac{d(d+1)(d+2)}{6} \mathbf{1}_+ \oplus \frac{d(d-1)(d-2)}{6} \mathbf{1}_- \oplus \frac{d(d+1)(d-1)}{3} \mathbf{2} \tag{11.462}$$

   Clearly, the first and second summands correspond to the totally symmetric and totally antisymmetric tensors, respectively.

Consider the Young tableau associated with the partition $(2, 1)$ denoted $T_3$ above. We have:

$$P(T_3)v_i \otimes v_j \otimes v_k = v_i \otimes v_j \otimes v_k - v_j \otimes v_i \otimes v_k + v_k \otimes v_j \otimes v_i - v_j \otimes v_k \otimes v_i \quad (11.463)$$

and therefore the components are of the form:

$$\tilde{t}^{ijk} = t^{ijk} - t^{jik} + t^{kji} - t^{kij} \quad (11.464)$$

for arbitrary $t^{ijk}$. That is

$$P(T_3)(t^{ijk} v_i \otimes v_j \otimes v_k) = \tilde{t}^{ijk} v_i \otimes v_j \otimes v_k \quad (11.465)$$

Now, the set of tensors of the form $\tilde{t}^{ijk}$ is the same as the set of tensors $\tilde{t}^{ijk}$ which satisfy the identities

$$\tilde{t}^{ijk} + \tilde{t}^{jki} + \tilde{t}^{kij} = 0 \quad (11.466)$$

$$\tilde{t}^{ijk} = -\tilde{t}^{jik} \quad (11.467)$$

The first set of equations (11.480) cuts out a space of dimension $\frac{2}{3}d(d^2 - 1)$. The next set of equations (11.481) cuts it down by half so we get $\frac{1}{3}d(d^2 - 1)$ as the dimension of this space of tensors. Indeed we can identify

$$P(T_3)V^{\otimes 3} = \ker(\Lambda^2 V \otimes V \to \Lambda^3 V) \quad (11.468)$$

3. If we take the partition $\lambda = (n, 0, 0, \dots)$ then $P(\lambda)$ projects to totally symmetric tensors and we get $S^n(V)$. Weyl's theorem tells us that if $V = \kappa^d$ is the fundamental representation of $GL(d, k)$ then $S^n(V)$ is an irreducible representation.

   Incidentally, the dimension of this representation is

$$\dim S^n(V) = \binom{n + d - 1}{n} \quad (11.469)$$

   as one easily proves by noting that the partition function for a collection of bosonic oscillators $\{a_i, a_i^\dagger\}_{i=1}^d$ with Hamiltonian $H = \sum_j a_j^\dagger a_j$ is

$$\sum_{n=0}^{\infty} q^n \dim S^n(V) = \frac{1}{(1 - q)^d} \quad (11.470)$$

   Note in particular that we get infinitely many irreducible representations of arbitrarily large dimension.

4. If we take the partition $\lambda = (1, 1, \dots, 1)$ then $P(\lambda)V^{\otimes n} = \Lambda^n V$ is the subspace of totally antisymmetric tensors of dimension

$$\dim \Lambda^n(V) = \binom{d}{n} \quad (11.471)$$

Note that this subspace vanishes unless $\dim V = d > n$.

If we consider the partition function for a collection of fermionic oscillators $\{a_i, a_i^\dagger\}_{i=1}^d$ with Hamiltonian $H = \sum_j a_j^\dagger a_j$ we get

$$\sum_{n=0}^{\infty} q^n \dim\Lambda^n(V) = (1+q)^d \tag{11.472}$$

5. Consider $G = U(2)$. We need only consider the Young diagrams with one or two rows. The projectors $P(T)$ associated with three or more rows will always vanish, since the indices $t^{i_1,\dots,i_n}$ now have $i$ running over only two values. In fact, the anti-symmetrization for diagrams with $d=2$ always produces copies of the determinant representation. That is, if $v_1, v_2 \in \mathbb{C}^2$ and $u \in U(2)$

$$u \cdot (v_1 \otimes v_2 - v_2 \otimes v_1) = (\det u)(v_1 \otimes v_2 - v_2 \otimes v_1) \tag{11.473}$$

For our current purposes define $v_1 \wedge v_2 := (v_1 \otimes v_2 - v_2 \otimes v_1)$. [171] Suppose $T$ is a Young tableau with $\ell$ columns of two boxes and $n$ columns of one box and we choose the tableau where the first column has labels $1,2$, the next $3,4$ and so on up to $(2\ell-1, 2\ell)$. Then

$$\begin{aligned} &T^{\otimes(2\ell+n)}(u)P(T)v_1 \otimes \cdots v_{2\ell} \otimes v_{2\ell+1} \otimes \cdots \otimes v_{2\ell+n} = \\ &\quad (\det u)^\ell (v_1 \wedge v_2) \otimes (v_3 \wedge v_4) \otimes \cdots (v_{2\ell-1} \wedge v_{2\ell}) \otimes \otimes T^{\otimes n}(u) \left(P(T')(v_{2\ell+1} \otimes \cdots \otimes v_{2\ell+n})\right) \end{aligned} \tag{11.474}$$

where $T'$ is a tableau consisting of a single row of $n$ boxes. So, restricting to $G = SU(2)$, *the only relevant Young diagrams are given by a single row of $n$ boxes, corresponding to the totally symmetric power of the fundamental representation:*

*That is, the irreducible representations of $SU(2)$ are in one-one correspondence with the positive integers and are all obtained by taking the symmetric powers of the fundamental representation.*

In the section on induced representations below we are going to understand this beautiful and significant fact in much more depth.

**Remarks**:

1. There are some very beautiful relations of the above mathematics to the quantum mechanics of a collection of noninteracting free fermions on a circle. See

   (a) M. Douglas, "Conformal field theory techniques for large N group theory," hep-th/9303159.

[171] The standard definition of exterior produce would divide this by 2!, but that would introduce many irrelevant factors of two in the next few lines.

(b) S. Cordes, G. Moore, and S. Ramgoolam, "Lectures on 2D Yang-Mills Theory, Equivariant Cohomology and Topological Field Theories," Nucl. Phys. B (Proc. Suppl 41) (1995) 184, section 4. Also available at hep-th/9411210.

and references therein for details.

2. In QM a system of $n$ particles has - a priori - a Hilbert space $\mathcal{H}^{\otimes n}$ where $\mathcal{H}$ is the single-particle Hilbert space. For a system of $n$ identical particles the true Hilbert space is the subspace $S^n\mathcal{H}$ for bosons and $\Lambda^n\mathcal{H}$ for fermions. We now recognize these as the isotypical subspaces for the one-dimensional representations of $S_n$. An old idea is that there should be a generalization of statistics - "parastatistics" - where the system of $n$ particles transforms in some other isotypical component corresponding to higher dimensional representations of the symmetric group. This is ruled out in relativistic field theory in spacetime dimensions above 3. In two dimensions the idea of anyons and, in particular, nonabelions is similar because there is a surjective homomorphism $\mathcal{B}_n \to S_n$. But the lift to $\mathcal{B}_n$ seems to give a much more interesting version of the idea.

3. Consider the integrals over $U(d)$ of the form:

$$\int_{U(d)} U_{i_1,j_1} \cdots U_{i_n,j_n} U^*_{i'_1,j'_1} \cdots U^*_{i'_m,j'_m} \mu(U(d)) \tag{11.475}$$

where we use the unit norm Haar measure $\mu(U(d))$. We have $\mu(U(d)) \propto \mu(U(1))\mu(SU(d))$ and the integral over the overall $U(1)$ phase projects to $n = m$. Then one can prove, using Schur-Weyl duality [172] that the integral is given by

$$\sum_{\sigma,\tau\in S_n} \delta_{i_1,i'_{\sigma(1)}} \cdots \delta_{i_n,i'_{\sigma(n)}} \delta_{j_1,j'_{\tau(1)}} \cdots \delta_{j_n,j'_{\tau(n)}} Wg(\sigma\tau^{-1}, d, n) \tag{11.476}$$

where $Wg(\sigma, d, n)$ is the *Weingarten function*

$$Wg(\sigma, d, n) = \frac{1}{(n!)^2} \sum_{\lambda} \frac{(\mathrm{dim} R_\lambda)^2}{\mathrm{dim}\mathcal{V}_\lambda} \chi^\lambda(\sigma) \tag{11.477}$$

where we sum over all partitions of $n$, or equivalently over all irreps of $S_n$, and $\mathcal{V}_\lambda$ is the corresponding irrep of $U(d)$ These integrals are of great use in the strong coupling expansion of lattice gauge theory.

♣GIVE SOME EXAMPLES ♣

---

**Exercise** *Mixed Symmetry Type*

Repeat the example above of mixed symmetry type rank 3 tensors for $T_2$.

a.) Apply the projector (11.437) to a tensor $t^{ijk} v_i \otimes v_j \otimes v_k$ produces a tensor with mixed symmetry:

$$P(T_2) v_i \otimes v_j \otimes v_k = v_i \otimes v_j \otimes v_k + v_j \otimes v_i \otimes v_k - v_k \otimes v_j \otimes v_i - v_k \otimes v_i \otimes v_j \tag{11.478}$$

[172] B. Collins, https://arxiv.org/pdf/math-ph/0205010.pdf

and therefore deduce that the components are of the form:

$$\tilde{t}^{ijk} = t^{ijk} + t^{jik} - t^{kji} - t^{jki} \tag{11.479}$$

for arbitrary $t^{ijk}$.

b.) Show that this tensor space is the space of tensors $\tilde{t}^{ijk}$ which satisfy the identities:

$$\tilde{t}^{ijk} + \tilde{t}^{jki} + \tilde{t}^{kij} = 0 \tag{11.480}$$

together with

$$\tilde{t}^{ijk} = -\tilde{t}^{kji} \tag{11.481}$$

c.) Compute the dimension of this space of tensors.

---

---

**Exercise** *Simplest Nontrivial Weingarten Function*
Using the above formulae compute

$$\int_{U(d)} U_{i_1,j_1} U^*_{i'_1,j'_1} \mu(U) \tag{11.482}$$

where $\mu(U)$ is the unit-volume Haar measure. [173]

---

### 11.17 Orthogonality Relations And Pontryagin Duality

It is interesting to think about the orthogonality of characters in the context of Pontryagin duality. Let $S$ be a locally compact Abelian group and $\widehat{S}$ its Pontryagin dual.

Compact groups have discrete sets of irreducible representations, but the representations of noncompact groups typically come in continuous families. Thus, for example, the Pontryagin dual of $\mathbb{Z}$ is the continuous group $U(1)$. The orthogonality relations for locally compact Abelian groups extend to this case, but the $\delta_{\mu,\nu}$ delta function on the representations becomes a delta function $\delta_{\hat{S}}(\chi_1 - \chi_2)$ on the characters, where the Dirac measure is relative to the Haar measure on $\hat{S}$. Thus we have:

$$\int_S \chi_1(s)^* \chi_2(s) ds = \delta_{\widehat{S}}(\chi_1 - \chi_2) \tag{11.484}$$

an similarly:

$$\int_{\widehat{S}} \chi(s_1)^* \chi(s_2) d\chi = \delta_S(s_1 - s_2) \tag{11.485}$$

---

[173] *Answer* We have

$$W(1, d, 2) = \frac{2}{d^2 - 1} \qquad W((12), d, 2) = -\frac{2}{d(d^2 - 1)} \tag{11.483}$$

and the integral is *******

Since $\chi(s) = \hat{s}(\chi)$ we see that

$$\delta_{\widehat{S}}(\hat{s}_1 - \hat{s}_2) = \delta_S(s_1 - s_2) \tag{11.486}$$

So Pontryagin duality extends to groups with measure. Given a function $\psi \in L^2(S)$, we can define its Fourier transform, quite generally, as

$$\hat{\psi}(\chi) := \int_S \chi(s)^* \psi(s) ds \tag{11.487}$$

So now we have the relation

$$\langle \psi_1, \psi_2 \rangle_{L^2(S)} = \langle \hat{\psi}_1, \hat{\psi}_2 \rangle_{L^2(\widehat{S})} \tag{11.488}$$

This is a result known as either the *Plancherel theorem* or the *Parseval theorem.* [174]

An important special case of the above is the case $S = \Gamma \subset \mathbb{R}^n$, an embedded lattice. Then we have:

$$\sum_{\gamma\in\Gamma} \chi_{\bar{k}_1}(\gamma)^* \chi_{\bar{k}_2}(\gamma) = \delta_{\hat{\Gamma}}(\bar{k}_1 - \bar{k}_2) \tag{11.489}$$

Here $\delta_{\hat{\Gamma}}(\bar{k})$ is the delta measure on the dual group

$$\hat{\Gamma} = \mathbb{R}^n / \Gamma^\vee \tag{11.490}$$

We can lift $\bar{k}$ to $k \in \mathbb{R}^n$, and then

$$\delta_{\hat{\Gamma}}(\bar{k}_1 - \bar{k}_2) = \sum_{\gamma^\vee \in \mathbb{R}^n} \delta_{\mathbb{R}^n}(k_1 - k_2 - \gamma^\vee) \tag{11.491}$$

On the other hand, using the explicit formula:

$$\chi_{\bar{k}}(\gamma) = e^{2\pi \mathrm{i} k\cdot\gamma} \tag{11.492}$$

we get the relation

$$\sum_{\gamma\in\Gamma} e^{2\pi\mathrm{i}(k_2-k_1)\cdot\gamma} = \sum_{\gamma^\vee\in\Gamma^\vee} \delta_{\mathbb{R}^n}(k_2 - k_1 - \gamma^\vee) \tag{11.493}$$

This is one version of the *Poisson summation formula.* Since the PSF is an important result we will unpack this a bit in the next section.

### 11.17.1 The Poisson Summation Formula

Let us begin with a standard derivation of the very useful Poisson summation formula:

Let $f : \mathbb{R} \to \mathbb{C}$ that decays fast enough that

$$F(x) := \sum_{n\in\mathbb{Z}} f(x+n) \tag{11.494}$$

[174] The original statements by Plancherel and Parseval concerned special cases and the two terms are not consistently used in the literature.

exists. Then $F(x)$ is clearly periodic of period one and therefore defines a function $F : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$. Since $\mathbb{R}/\mathbb{Z} \cong U(1)$ we can decompose in terms of irreducible representations:

$$F(x) = \sum_{\rho_m} \hat{F}(\rho_m) e^{2\pi i m x} \tag{11.495}$$

$$\hat{F}(\rho_m) = \int_0^1 e^{-2\pi i m t} F(t) dt \tag{11.496}$$

Now note that

$$\begin{aligned} \int_0^1 e^{-2\pi i m t} F(t) dt &= \int_0^1 e^{-2\pi i m t} \sum_{k\in\mathbb{Z}} f(k+t) dt \\ &= \int_{-\infty}^{+\infty} e^{-2\pi i m t} f(t) dt \end{aligned} \tag{11.497}$$

Putting $x = 0$ we learn that for suitably rapidly decaying functions:

$$\sum_{m\in\mathbb{Z}} f(m) = \sum_{w\in\mathbb{Z}} \widehat{f}(w) \tag{11.498}$$

where $\widehat{f}$ is the Fourier transform:

$$\widehat{f}(w) = \int_{\mathbb{R}} e^{-2\pi i t w} f(t) dt \tag{11.499}$$

This is valid for functions such that

1. $f$ decays rapidly enough so that the sum on the LHS converges.
2. The Fourier transform $\widehat{f}$ exists.
3. The Fourier transform $\widehat{f}$ decays rapidly enough so that the sum on the RHS converges.

Note that this can also be understood as a special case of the orthogonality relations for characters on $\mathbb{Z}$:

$$\boxed{\sum_{n\epsilon\mathbb{Z}} e^{2\pi i n t} e^{-2\pi i n t'} = \delta_{\mathbb{R}/\mathbb{Z}}(t-t') = \sum_{k\in\mathbb{Z}} \delta_{\mathbb{R}}(t-t'-k)} \tag{11.500}$$

where $\delta_{\mathbb{Z}}$ means the delta function on the Pontryagin dual group $\mathbb{Z}$.

The generalization of this statement is our result (11.493) above.

Put differently, we have:

$$\sum_{\vec{n}\in\mathbb{Z}^d} f(\vec{n}) = \sum_{\vec{m}\in\mathbb{Z}^d} \int_{\mathbb{R}^d} e^{-2\pi i \vec{m}\cdot\vec{t}} f(t) dt = \sum_{\vec{m}\in\mathbb{Z}^d} \int_{R^d} e^{+2\pi i \vec{m}\cdot\vec{t}} f(t) dt \tag{11.501}$$

That is:

$$\sum_{\vec{v}\in\Gamma} f(\vec{v}) = \sum_{\vec{l}\in\Gamma^\vee} \hat{f}(\vec{l}) \ . \tag{11.502}$$

**Remarks**

1. Since people have different conventions for the factors of $2\pi$ in Fourier transforms it is hard to remember the factors of $2\pi$ in the PSF. The equation (11.500) has no factors of $2\pi$. One easy way to see this is to integrate both sides from $t = -1/2$ to $t = +1/2$.

2. One application of this is the x-ray crystallography: The $LHS$ is the sum of scattered waves. The $RHS$ constitutes the bright peaks measured on a photographic plate.

3. Another application is to analytic number theory. If $\tau$ is in the upper half complex plane, and $\theta, \phi, z$ are complex numbers define the *Riemann theta function with characteristics* $\theta, \phi$

$$\vartheta[{\theta \atop \phi}](z|\tau) := \sum_{n\in\mathbb{Z}} e^{i\pi\tau(n+\theta)^2+2\pi i(n+\theta)(z+\phi)} \tag{11.503}$$

(usually, $\theta, \phi$ are taken to be real numbers, but $z$ is complex). This converges to an entire function of $z$ and is also holomorphic for $\mathrm{Im}\tau > 0$.

Using the Poisson summation formula one can show that it obeys the modular transformation law:

$$\vartheta[{\theta \atop \phi}](\frac{-z}{\tau}|\frac{-1}{\tau}) = (-i\tau)^{1/2} e^{2\pi i\theta\phi} e^{i\pi z^2/\tau} \vartheta[{-\phi \atop \theta}](z|\tau) \tag{11.504}$$

---

**Exercise**
a.) Show that

$$\sum_{n\in Z} e^{-\pi a n^2 + 2\pi i b n} = \sqrt{\frac{1}{a}} \sum_{m\in Z} e^{-\frac{\pi(m-b)^2}{a}} \tag{11.505}$$

b.) Check equation (11.504).

---

### 11.17.2 The Finite Heisenberg Group And The Quantum Mechanics Of A Particle On A Discrete Approximation To A Circle

♣There is too much redundancy in the following discussion with previous material. ♣

Now, as we have discussed at length $L^2(G)$ is naturally a unitary representation of $G\times G$ for any compact group $G$. When $G = S$ is Abelian the left and right $G$ actions are essentially the same, so there is no real loss in simply regarding $L^2(S)$ simply as an $S$-representation. However, thanks to the orthogonality relations and Pontryagin duality

$$L^2(S) \cong L^2(\widehat{S}) \tag{11.506}$$

so, actually, $L^2(S)$ is <u>also</u> a representation of the Pontryagin dual group $\widehat{S}$. However, $L^2(S)$ is <u>not</u> a representation of the direct product $S \times \widehat{S}$, but rather of a Heisenberg extension of $S \times \widehat{S}$. See Section ********* below for the general discussion. In this section we explain

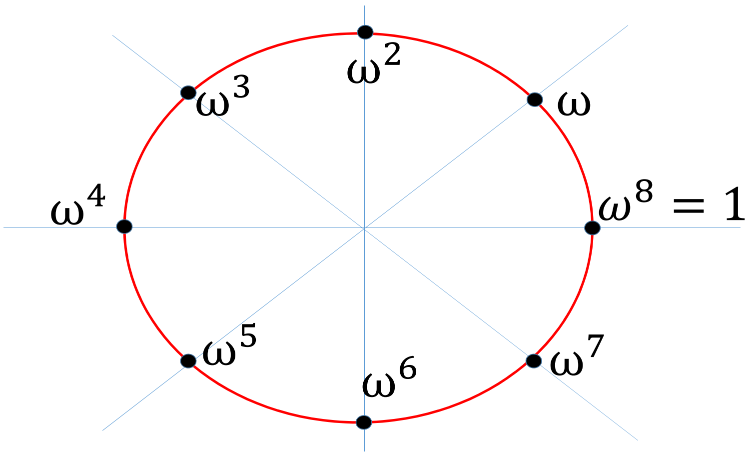



**Figure 35:** Roots of unity on the unit circle in the complex plane. Here $\omega = e^{2\pi \mathrm{i}/8}$ is a primitive eighth root of 1.

this very important fact in the special case where $S \cong \widehat{S}$ is a self dual group, in particular we explain it for $S = \mu_N$.

We can consider the subgroup $\mu_N \subset U(1)$ as a discrete approximation to a circle. The case $N = 8$ is illustrated in Figure 35. Thus, a more physical view of what we are discussing is the quantum mechanics of a particle constrained to live on the points of $\mu_N$ on the circle. We will see some nice physical manifestations of Pontryagin duality and also obtain some insights about the finite Heisenberg groups $\mathrm{Heis}_N$ introduced in equations (7.133) and (7.134) et. seq.

As a vector space $L^2(\mu_N)$ is isomorphic to $\mathbb{C}^N$. To specify a function is to specify the $N$ different complex values $\Psi(\omega^k)$ where $\omega$ is a primitive $N^{th}$ root of one, say, $\omega = \exp[2\pi \mathrm{i}/N]$ for definiteness, and $k = 0, \dots, N-1$. (We will not try to normalize our wavefunctions $\Psi$, but we could. It would make no difference to the present considerations.) Using the Haar measure we have a Hilbert space with inner product:

$$\langle \Psi_1, \Psi_2 \rangle := \frac{1}{|G|} \sum_{g \in G} \Psi_1^*(g) \Psi_2(g) \tag{11.507}$$

Now, recall the general definition from section 5.1. $G$ acts naturally on $X$ (which happens to be $G$ itself) by left-multiplication. The induced action of $G$ on the complex-valued functions on $G$ in this case is such that the generator $\omega$ of $\mathbb{Z}_N$ acts on the space of functions via:

$$\begin{aligned} \tilde{\phi}(\omega, \Psi)(\omega^k) &:= \Psi(\phi(\omega^{-1}, \omega^k)) \\ &= \Psi(\omega^{k-1}) \end{aligned} \tag{11.508}$$

So the generator $\omega$ of the group $\mathbb{Z}_N$ acts linearly on the functions $\mathcal{F}[\mathbb{Z}_N \to \mathbb{C}]$. We call this linear operator $P$. We can therefore rewrite (11.508) as

$$(P \cdot \Psi)(\omega^k) := \Psi(\omega^{k-1}) \tag{11.509}$$

Note that with respect to the inner product (11.507) $P$ is clearly a unitary operator.

The operator $P$ can be viewed as translation operator around the discrete circle by one step in the clockwise direction. Recall that in the quantum mechanics of a particle on the line translation by a distance $a$ is

$$(T(a) \cdot \Psi)(x) = \Psi(x + a) \tag{11.510}$$

This equation makes sense also for a particle on the circle, that is, with $x, a$ considered periodically. So our $P$ is $T(a)$ for translation by $2\pi/N$ times around the circle clockwise.

**Remark**: In the quantum mechanics of a particle on a line or circle we could also write

$$(T(a) \cdot \Psi)(x) = \Psi(x + a) = (\exp[ia\hat{p}])\Psi(x) = (\exp[a\frac{d}{dx}] \cdot \Psi)(x) \tag{11.511}$$

so the momentum operator $\hat{p}$ generates translations. In the finite Heisenberg group there is no analog of the infinitesimal translations generated by $\hat{p}$, but only of a finite set of discrete translations.

Now let $Q$ be the position operator:

$$(Q \cdot \Psi)(\omega^k) := \omega^k \Psi(\omega^k) \tag{11.512}$$

$Q$ is likewise a unitary operator.

Now note that

$$\begin{aligned}(P \circ Q \cdot \Psi)(\omega^k) &= (Q \cdot \Psi)(\omega^{k-1}) \\ &= \omega^{k-1}\Psi(\omega^{k-1})\end{aligned} \tag{11.513}$$

while

$$\begin{aligned}(Q \circ P \cdot \Psi)(\omega^k) &= \omega^k (P \cdot \Psi)(\omega^k) \\ &= \omega^k \Psi(\omega^{k-1})\end{aligned} \tag{11.514}$$

and therefore we conclude that we have the operator equation:

$$Q \circ P = \omega P \circ Q \tag{11.515}$$

which is a key defining relation when presenting the Heisenberg group in terms of generators and relations. (The other relations are $P^N = Q^N = 1$ and $\omega P = P\omega$ and $\omega Q = Q\omega$.)

A natural orthogonal basis of wave functions is the delta-function basis:

$$\delta_j(\omega^k) = \delta_{\bar{j},\bar{k}} \tag{11.516}$$

where $\bar{j}, \bar{k} \in \mathbb{Z}/N\mathbb{Z}$, viewed additively. In the ordered basis $\delta_0, \dots, \delta_{N-1}$ we easily compute

$$P \cdot \delta_j = \delta_{j+1} \tag{11.517}$$

and therefore the matrix for $P$ relative to the basis $\{\delta_j\}$, is the matrix with matrix elements

$$P_{i,j} = \delta_{i,j+1} \tag{11.518}$$

so, for $N = 3$ it is

$$P = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{11.519}$$

Similarly, in the basis $\delta_j$ we have

$$Q_{i,j} = \omega^j \delta_{i,j} \tag{11.520}$$

and since $j = 0, 1, \dots, N-1$ we have for $N = 3$:

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix} \tag{11.521}$$

Thus, we have recovered the $N \times N$ clock and shift matrices we discussed above. The group of unitary operators generated by discrete position and translation operators is the finite Heisenberg group that we studied in equations (7.133) and (7.134) et. seq.

It is interesting to study the operators $P$ and $Q$ in a different basis. We can introduce a "plane wave basis" of functions $\Psi_j \in \mathcal{H}$, with $j = 0, \dots, N-1$ defined by

$$\Psi_j(\omega^k) = \omega^{jk} \tag{11.522}$$

It is now easy to compute the action of $P$ and $Q$ on this basis:

$$\begin{aligned} P \cdot \Psi_j &= \omega^{-j} \Psi_j \\ Q \cdot \Psi_j &= \Psi_{j+1} \end{aligned} \tag{11.523}$$

The roles of $P$ and $Q$ have been exchanged! $P$ (as is $P^{-1}$) is now representated by "clock matrix" and $Q$ is represented by a "shift matrix". Indeed, what we have done is perform a transformation from a position representation to a momentum representation in the language of quantum mechanics.

The reader should recognize the planewave basis $\Psi_j$ as the basis of characters on $\mathbb{Z}_N$. The fact that they form an ON basis of $L^2(\mathbb{Z}_N)$ is, once again, a statement of the Peter-Weyl theorem we studied above.

Moreover, the exchange of clock and shift matrices by passing from the $\delta_j$ basis to the basis of characters is again an instance of Pontryagin duality. Recall that for reasonable [175] Abelian groups we have

$$\widehat{\widehat{G}} \cong G \tag{11.524}$$

If $\chi$ is a homomorphism we can define the Fourier transform

$$\widehat{\Psi}(\chi) := \frac{1}{\sqrt{|G|}} \sum_{g \in G} \chi(g)\Psi(g) \tag{11.525}$$

giving an <u>isometry</u> of $L^2(\widehat{G})$ with $L^2(G)$. Being an isometry means that

$$\langle \widehat{\Psi}_1, \widehat{\Psi}_2 \rangle_{L^2(\widehat{G})} = \langle \Psi_1, \Psi_2 \rangle_{L^2(G)} \tag{11.526}$$

which is equivalent to the character orthogonality relations for $G$ and $\widehat{G}$:

$$\begin{aligned} \frac{1}{|\widehat{G}|} \sum_{\chi \in \widehat{G}} \chi^*(g_1)\chi(g_2) &= \delta_{g_1,g_2} \\ \frac{1}{|G|} \sum_{g \in G} \chi_1^*(g)\chi_2(g) &= \delta_{\chi_1,\chi_2} \end{aligned} \tag{11.527}$$

Now, for $G = \mathbb{Z}_N$ we have $\widehat{G} \cong \mathbb{Z}_N$. The passage from the $\delta_j$ basis to the $\Psi_j$ basis, which diagonalizes $P$ is just the *finite Fourier transform.*

In more concrete terms: The trace of all the powers of $P$ less than $N$ is also obviously zero and $P^N = 1$ and no smaller power of $P$ is the identity. So $P$ must be unitarily equivalent to $Q$. Now we can easily check that

$$SPS^{-1} = Q \tag{11.528}$$

where $S$ is the finite Fourier transform matrix

$$S_{j,k} = \frac{1}{\sqrt{N}} e^{2\pi i \frac{jk}{N}} \tag{11.529}$$

One easy way to check this is to multiply the matrices $SP$ and $QS$ in the $\delta_j$ basis. The reader should check that $S$ is in fact a unitary matrix and that the matrix elements only depend on the projections $\bar{j}, \bar{k} \in \mathbb{Z}/N\mathbb{Z}$.

In any case, $S_{j,k}$ takes us from a position basis $\delta_j$ to a "momentum basis" where $P$ is diagonal, in beautiful analogy to how the Fourier transform converts a position basis to a momentum basis for a particle on the line.

**Remark**: Let us return to the use of characters to diagonalze the normal modes of a system of particles connected by springs in Section *** above. We recognize that

$$A = 2 - P - P^{-1} \tag{11.530}$$

We have just seen that the finite Fourier transform diagonalizes $P$, and therefore it diagonalizes $A$.

♣THE FOLLOWING EXERCISES ARE NOW REDUNDANT ♣

[175] e.g. locally compact Abelian groups. See the book by Kirillov on representation theory.

---

**Exercise** *The Pontryagin Dual Of $\mathbb{Z}_N$ Is Isomorphic To $\mathbb{Z}_N$*

Let $\chi : \mathbb{Z}_N \to U(1)$ be a homomorphism. Let $g$ be a generator of $\mathbb{Z}_N$. Show that $\chi(g)$ must be an $N^{th}$ root of unity, and choosing any $N^{th}$ root of unity defines the homomorphism. Conclude that $\widehat{\mathbb{Z}_N} \cong \mathbb{Z}_N$.

---

---

**Exercise** *Orthogonality For $\mathbb{Z}_N$*

Write out the equations (11.527) for the case $G = \mathbb{Z}_N$

---

### 11.18 Induced Representations

Let $G$ be a group and $H$ a subgroup. Suppose that $\rho : H \to GL(V)$ is a representation of the *subgroup* $H$. Using this data we are going to produce, canonically, a representation of the larger group $G$. The new representation of $G$ is known as an *induced representation.* Note that, in general, there is no way to extend $\rho : H \to GL(V)$ to a homomorphism $\rho : G \to GL(V)$. As a counterexample, we will see later that the only one-dimensional representation of $SU(2)$ is the trivial representation. However, there are nontrivial representations of the subgroup of diagonal matrices. So, in general, a representation $\rho : H \to GL(V)$ of a subgroup $H \subset G$ is not just the restriction of a representation of $G$ on $V$.

Then, as we have seen, $\mathrm{Map}(G, V)$ is canonically a $G \times H$-space. The left-action of $G \times H$ is defined by declaring that for $(g, h) \in G \times H$ and $\Psi \in \mathrm{Map}(G, V)$ the new function $\phi((g, h), \Psi) \in \mathrm{Map}(G, V)$ is the function $G \to V$ defined by:

$$\phi((g, h), \Psi)(g_0) := \rho(h) \cdot \Psi(g^{-1} g_0 h) \tag{11.531}$$

for all $g_0 \in G$. Or, in slightly lighter notation:

$$(g, h) \cdot \Psi(g_0) := \rho(h)\Psi(g^{-1} g_0 h) \tag{11.532}$$

Now, we can consider the subspace of functions *fixed by the action of* $1 \times H$. That is, we consider functions which satisfy

$$\boxed{\Psi(gh^{-1}) = \rho(h)\Psi(g)} \tag{11.533}$$

for every $g \in G$ and $h \in H$. Put differently: There are two natural left-actions by $H$ on $\mathrm{Map}(G, V)$ and we consider the subspace where they are equal. Such functions are said to be the *$H$-equivariant.* See the exercise below for the justification of this terminology. Note that the space of $H$-equivariant functions $G \to V$ is a linear subspace of $\mathrm{Map}(G, V)$. We will denote it by $\mathrm{Ind}_H^G(V)$:

$$\mathrm{Ind}_H^G(V) := \{\Psi : G \to V \ \mid \ \Psi(gh^{-1}) = \rho(h)\Psi(g) \qquad \forall g \in G, h \in H\} \tag{11.534}$$

Note that since we are taking the fixed points of the subgroup $\{1_G\} \times H$ of the $G \times H$ action on $\mathrm{Map}(G, V)$ there is still a $G$ action on the fixed point set. More explicitly, if $\Psi$ is an $H$-equivariant function satisfying (11.533) then $g \cdot \Psi$ with values

$$(g \cdot \Psi)(g_0) := \Psi(g^{-1} g_0) \tag{11.535}$$

is also an $H$-equivariant function. (Check this!) Thus $\mathrm{Ind}_H^G(V)$ is a representation space of $G$.

The subspace $\mathrm{Ind}_H^G(V) \subset \mathrm{Map}(G, V)$ of $H$-equivariant functions, i.e. functions satisfying (11.533) is called the *induced representation of $G$, induced by the representation $V$ of the subgroup $H$*. This is an important construction with a beautiful underlying geometrical interpretation. In a sense we will explain below all the representations of compact groups follow from this construction. One can also use it to construct representations of many important noncompact and infinite-dimensional groups. For this reason it appears in many places in physics.

Two examples of important applications in physics are:

1. The irreducible unitary representations of space groups in condensed matter physics.

2. The irreducible unitary representations of the Poincaré group in QFT.

**Example**: Let us take $V = \mathbb{C}$ with the trivial representation of $H$, i.e. $\rho(h) = 1$. Then the induced representation is the vector space of functions on $G$ which are invariant under right-multiplication by $H$. This is precisely the vector space of $\mathbb{C}$-valued functions on the homogeneous space $G/H$. Recall that for $G = SU(2)$ and $H = U(1)$ we have seen that $G/H = \mathbb{CP}^1 \cong S^2$ so we can get a nice basis of orthogonal functions on $S^2$ from the functions on $SU(2)$. These are called *spherical harmonics* and we will discuss them more below.

---

**Exercise** *Functoriality Of Induction*

a.) Show that

$$\mathrm{Ind}_H^G(V_1 \oplus V_2) \cong \mathrm{Ind}_H^G(V_1) \oplus \mathrm{Ind}_H^G(V_2) \tag{11.536}$$

b.) If $A \in \mathrm{Hom}_H(V_1, V_2)$ is an intertwiner between $H$-reps then there is an induced map $F(A) : \mathrm{Ind}_H^G(V_1) \to \mathrm{Ind}_H^G(V_2)$ which is an intertwiner of $G$-representations. [176]

---

---

**Exercise** *$H$-Equivariance*

[176] *Answer* $(F(A)(\Psi))(g) = A(\Psi(g))$.

Show that a function $\Psi \to V$ satisfying (11.533) fits in a commutative diagram:

$$\begin{array}{ccc} G & \xrightarrow{\Psi} & V \\ {\scriptstyle R(h)}\downarrow & & \downarrow{\scriptstyle \rho(h^{-1})} \\ G & \xrightarrow{\Psi} & V \end{array} \tag{11.537}$$

where $R(h) : g \mapsto gh$ is the right action of $H$ on $G$. Thus a function satisfying (11.533) is a morphism of $H$-spaces, justifying the term "equivariant."

---

### 11.18.1 The Geometrical Interpretation

Note that there is a right $H$-action on the set $G \times V$:

$$\phi_h : (g, v) \mapsto (gh, \rho(h^{-1})v) \tag{11.538}$$

we can therefore form the quotient space of orbits. In this case it is usually denoted $G \times_H V$, but it is just the set of equivalence classes under the above right $H$-action. There is a natural map

$$\pi : G \times_H V \to G/H \tag{11.539}$$

given by $\pi : [(g, v)] \mapsto gH$. Referring back to the discussion of equation (8.261) we see that this is the *associated bundle to the principal* $H$ *bundle* $\pi : G \to G/H$.

When $G, H$ are Lie groups and $\rho$ is a continous representation the map $\pi$ is continuous. Moreover, the fiber above any coset $gH$ is the vector space $V$. We therefore have an example of a vector bundle over $G/H$ with fiber $V$. The sections of the vector bundle are, by definition, continuous maps

$$s : G/H \to G \times_H V \tag{11.540}$$

that are a right-inverse to $\pi$, that is $\pi \circ s = Id_{G/H}$. To construct such a section we have to identify, for each coset $gH$ an equivalence class in $G \times V$ which projects back down to $gH$. If we represent $gH$ by $g$ then it must be an equivalence class of the form

$$s(gH) = [(g, v(g))] \tag{11.541}$$

for some vector $v(g)$ associated to $g$. But now $gH = \tilde{g}H$ when $\tilde{g} = gh$ for $h \in H$. So it must be that

$$[(g, v(g))] = s(gH) = s(\tilde{g}H) = [(\tilde{g}, v(\tilde{g})] = [(gh, v(\tilde{g})] = [(g, \rho(h)v(\tilde{g}))] \tag{11.542}$$

and hence

$$v(g) = \rho(h)v(\tilde{g}) = \rho(h)v(gh) \qquad \Rightarrow \qquad v(gh) = \rho(h^{-1})v(g) \tag{11.543}$$

This must hold for all $g \in G$, and all $h \in H$, and hence $g \mapsto v(g)$ is an equivariant function: Thus,

*the space of sections of the homogeneous vector bundle* $\pi : G \times_H V \to G/H$ *is canonically identified with the space of* $H$*-equivariant functions* $G \to V$ *satisfying* (11.533).

### 11.18.2 Explicit Characters For A Finite Group $G$

Suppose that $G$ is a finite group and $H$ is a subgroup. Then we can write the distinct left cosets, i.e. the orbits under the right $H$-action as

$$G/H = \amalg_{i=1}^{r} X_i \tag{11.544}$$

Geometrically, $G/H$ is just a disjoint union of points.

We are now going to describe how to write an explicit basis for an induced representation by a representation of $H$.

First, for each $i$ choose a $g_i \in G$ so that $X_i = g_i H$. Another choice would be related by $\tilde{g}_i = g_i h_i$ for some $h_i \in H$. The group $G$ acts on $G/H$ by a left action and we can say

$$g \cdot X_i = X_{g \cdot i} \tag{11.545}$$

so for each $g$ we get a permutation of $\{1, \dots, r\}$ denoted by $i \mapsto g \cdot i$. It follows that if $g \cdot i = j$ then we can write

$$g g_i = g_j h(g, i) \tag{11.546}$$

for some set of group elements $h(g, i) \in H$ parametrized by $G \times \{1, \dots, r\}$.

Now let $\rho : H \to GL(V)$ be a finite dimensional representation of $H$. We can write an explicit basis for $\mathrm{Ind}_H^G(V)$ as follows. First, choose a basis $\{w_a\}$ for $V$. Now we define $H$-equivariant functions $\Psi_{i,a}$ defined by the condition:

$$\Psi_{i,a}(g_j) = \delta_{i,j} w_a \tag{11.547}$$

This condition defines a unique equivariant function $\Psi_{i,a}$ because for any other group element $g \in G$ we know that $g \in X_j$ for some $j$. Then if $j \neq i$ $\Psi_{i,a}(g) = 0$. If $g \in X_i$ then $g = g_i h$ for some $h$ and then

$$\Psi_{i,a}(g) = \Psi_{i,a}(g_i h) = \rho(g^{-1}) w_a \tag{11.548}$$

so the value is determined. Moreover, it is easy to see that the $\Psi_{i,a}$ is a linearly independent spanning set. These equivariant functions give a basis for the induced representation.

In the geometrical interpretation the associated bundle is a vector bundle with a vector space $\cong V$ over each of the points of $G/H$, and the equivariant functions $\Psi_{i,a}$ correspond to sections with support only on the point $X_i$.

Now, let us work out the action of some group element $g \in G$ in this basis. By definition:

$$(g \cdot \Psi_{i,a})(g_0) = \Psi_{i,a}(g^{-1} g_0) \tag{11.549}$$

so this vanishes unless $g^{-1} g_0 \in X_i$, that is unless $g_0 \in g \cdot X_i = X_{g \cdot i}$. If we denote $j := g \cdot i$ then it means that $g \cdot \Psi_{i,a}$ has support on $X_j$, and is a linear combination of the functions $\Psi_{j,b}$. To find out what linear combination it is we compute:

$$(g \cdot \Psi_{i,a})(g_j) = \Psi_{i,a}(g^{-1} g_j) = \Psi_{i,a}(g_i h(g, i)^{-1}) \tag{11.550}$$

where we have made use of (11.546). But now using equivariance:

$$(g \cdot \Psi_{i,a})(g_j) = \Psi_{i,a}(g^{-1}g_j) = \rho(h(g,i))\Psi_{i,a}(g_i) = \rho(h(g,i))w_a \tag{11.551}$$

This completely determines the group action in this basis.

Note that the character of the representation can now be computed:

$$\chi_{\mathrm{Ind}_H^G(V)}(g) = \sum_{i:g\cdot i=i} \chi_V(h(g,i)) = \sum_{i:g\cdot i=i} \chi_V(g_i^{-1}gg_i) \tag{11.552}$$

where in the second equality we have again made use of (11.546). Note that in $\chi_V(g_i^{-1}gg_i)$ we cannot use cyclicity of the trace to drop $g_i$ because $g$ is not in the group $H$ and $\chi_V(g)$ does not make any sense.

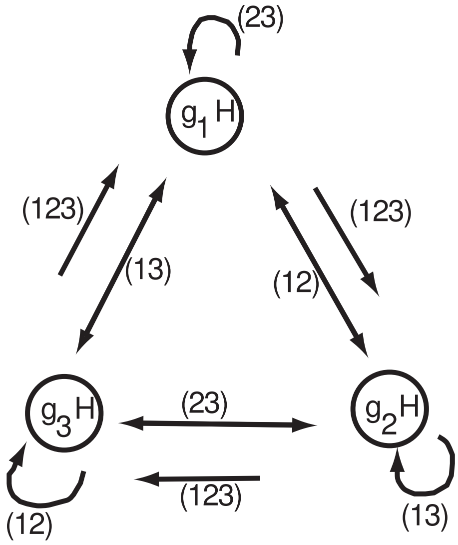



**Figure 36:** The left action of $G = S_3$ on $G/H$. In fact, this picture should be considered as a picture of a category, in this case, a groupoid.

**Example**: The simplest nontrivial example is given by taking $G = S_3$ to be the permutation group and $H = \{1, (12)\} \cong S_2$ be a subgroup isomorphic to $\mathbb{Z}_2$. The set of cosets $G/H$ consists of 3 points, but it is much more than that. It is a nice example of a groupoid, when one takes into account the left action of $G$ on these points. See figure (**??**). For more about the category associated to any group action on any set see Section 17 below.

♣The notation in the figure is not consistent with the notation in the text!!! ♣

Now let us consider the induced representations starting from the two irreps of $H$. There are two irreducible representations of $H$, the trivial and the sign representation Call them $V(\epsilon)$, with $\epsilon = \pm$. The carrier space is one-dimensional so $V(\epsilon) \cong \mathbb{C}$ as a vector space. Accordingly, we are looking at a line bundle over $G/H$ and the vector space of sections of $G \times_H V(\epsilon)$ is 3-dimensional.

To write an explicit basis as described below we make the choice of three coset representatives:

$$g_1 = (13), \qquad g_2 = (23), \qquad g_3 = (12) \tag{11.553}$$

So that:

$$
\begin{aligned}
(12)g_1 &= g_2(12) \\
(12)g_2 &= g_1(12) \\
(12)g_3 &= g_3(12) \\
(13)g_1 &= g_3(12) \\
(13)g_2 &= g_2(12) \\
(13)g_3 &= g_1(12) \\
(23)g_1 &= g_1(12) \\
(23)g_2 &= g_3(12) \\
(23)g_3 &= g_2(12) \\
(123)g_1 &= g_2 \\
(123)g_2 &= g_3 \\
(123)g_3 &= g_1
\end{aligned}
\tag{11.554}
$$

Now, since $V \cong \mathbb{C}$ and equivariant function $\Psi : G \to V$ is just a complex valued function on $G$. An equivariant function in $\mathrm{Ind}_{S_2}^{S_3}(V(\epsilon))$ will satisfy:

$$
\begin{aligned}
\Psi(12) &= \epsilon\Psi(1) \\
\Psi(123) &= \epsilon\Psi(13) \\
\Psi(132) &= \epsilon\Psi(23)
\end{aligned}
\tag{11.555}
$$

These conditions cut down the six-dimensional space of all complex-valued functions on $G$ to a three-dimensional space, in accord with the geometrical interpretation. Choosing the basis vector $w_a = 1$ in $\mathbb{C}$ our basis from above becomes:

$$\Psi_i(g_j) = \delta_{i,j} \tag{11.556}$$

Now, let us compute the representation matrices in this basis. It is easy to see that $(123) \cdot \Psi_1$ is supported on $g_2 H$ and that

$$((123) \cdot \Psi_1)(g_2) = \Psi_1((132)g_2) = \Psi_1((13)) = 1 \tag{11.557}$$

So $(123) \cdot \Psi_1 = \Psi_2$. Working out the action on $\Psi_2, \Psi_3$ (exercise!) gives

$$\rho_{ind}((123)) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{11.558}$$

Note that the character is $\chi_{ind}((123)) = 0$, in accord with our general formula because there are no fixed points of the action of $(123)$ on $G/H$.

Now let us work out the representation of (12). Again, it is easy to see that $(12) \cdot \Psi_1$ will be supported on $g_2H$ so it will be proportional to $\Psi_2$. We compute:

$$\begin{aligned} ((12)\cdot\Psi_1)(g_2) &= \Psi_1((12)(23)) \\ &= \Psi_1(123) \\ &= \Psi_1((13)(12)) \\ &= \rho_\epsilon(12)\Psi_1(13) \\ &= \epsilon \end{aligned} \tag{11.559}$$

and we conclude that $(12) \cdot \Psi_1 = \epsilon\Psi_2$. So:

$$\rho_{ind}(12) = \begin{pmatrix} 0 & \epsilon & 0 \\ \epsilon & 0 & 0 \\ 0 & 0 & \epsilon \end{pmatrix} \tag{11.560}$$

The character is clearly $\chi_{ind}((12)) = \epsilon$.

Now, using orthogonality of the characters we get the isotypical decomposition:

$$\mathrm{Ind}_{S_2}^{S_3}(V(\epsilon)) = W(\epsilon) \oplus W_2 \tag{11.561}$$

where $W(\epsilon)$ is the one-dimensional irrep of $S_3$.

********************

MORE MATERIAL IN GTLect4-IntroRepTheory-2020

*******************

***********************************************

### 11.18.3 Frobenius Reciprocity

The theory of induced representations is already interesting and nontrivial for $G, H$ finite groups. In this case $G \to G/H$ is a finite cover (by $H$) of a discrete set of points. Nevertheless, the general geometrical ideas apply.

Let $\mathrm{Rep}(G)$ denote the category of finite-dimensional representations of $G$. Morphisms between $W_1, W_2 \in \mathrm{Rep}(G)$ are linear transformations commuting with $G$, i.e. $G$-intertwiners, and the vector space of all morphisms is denoted $\mathrm{Hom}_G(W_1, W_2)$. The induced representation construction defines a functor

$$\mathrm{Ind} : \mathrm{Rep}(H) \to \mathrm{Rep}(G). \tag{11.562}$$

(We denoted this by $\mathrm{Ind}_H^G$ before but $H, G$ will be fixed in what follows so we simplify the notation.) On the other hand, there is an obvious functor going the other way, since any $G$-rep $W$ is *a foriori* an $H$-rep, by restriction. Let us denote this "restriction functor"

$$R : \mathrm{Rep}(G) \to \mathrm{Rep}(H) \tag{11.563}$$

How are these two maps related? The answer is that they are "adjoints" of each other! This is the statement of Frobenius reciprocity:

$$\text{Hom}_G(W, \text{Ind}(V)) = \text{Hom}_H(R(W), V) \tag{11.564}$$

We can restate the result in another way which is illuminating because it helps to answer the question: How is $\text{Ind}_H^G(V)$ decomposed in terms of irreducible representations of $G$? Let $W_\alpha$ denote the distinct irreps of $G$. Then Schur's lemma tells us that

$$\text{Ind}_H^G(V) \cong \oplus_\alpha W_\alpha \otimes \text{Hom}_G(W_\alpha, \text{Ind}_H^G(V)) \tag{11.565}$$

But now Frobenius reciprocity (11.564) allows us to rewrite this as

$$\text{Ind}_H^G(V) \cong \oplus_\alpha W_\alpha \otimes \text{Hom}_H(R(W_\alpha), V) \tag{11.566}$$

where the sum runs over the unitary irreps $W_\alpha$ of $G$, with multiplicity one.

The statement (11.566) can be a very useful simplification of (11.565) if $H$ is "much smaller" than $G$. For example, $G$ could be nonabelian, while $H$ is abelian. But the representation theory for abelian groups is much easier! Similarly, $G$ could be noncompact, while $H$ is compact. etc.

*Proof of Frobenius reciprocity*:

In order to prove (11.566) we note that it is equivalent (see the exercise below) to the statement that the character of $\text{Ind}_H^G(V)$ is given by

$$\chi(g) = \sum_{x \in G/H} \hat{\chi}(x^{-1} g x) \tag{11.567}$$

where $x$ runs over a set of representatives and $\hat{\chi}$ is the character $\chi_V$ for $H$ when the argument is in $H$ and zero otherwise.

On the other hand, (11.567) can be understood in a very geometrical way. Think of the homogeneous vector bundle $G \times_H V$ as a collection of points $g_j H$, $j = 1, \dots, n$ with a copy of $V$ sitting over each point. Now, choose a representative $g_j \in G$ for each coset. Having chosen representatives $g_j$ for the distinct cosets, we may write:

$$g \cdot g_j = g_{g \cdot j} h(g, j) \tag{11.568}$$

where $j \mapsto g \cdot j$ is just a permutation of the integers $1, \dots, n$, or more invariantly, a permutation of the points in $G/H$.

Now let us define a basis for the induced representation by introducing a basis $v_a$ for the $H$-rep $V$ and the equivariant functions determined by:

$$\psi_{i,a}(g_j) := v_a \delta_{i,j} \tag{11.569}$$

Geometrically, this is a section whose support is located at the point $g_i H$. The equivariant function is then given by

$$\psi_{i,a}(g_j h) := \rho(h^{-1}) v_a \delta_{i,j} \tag{11.570}$$

Now let us compute the action of $g \in G$ in this basis:

$$
\begin{aligned}
(g \cdot \psi_{i,a})(g_j) &= \psi_{i,a}(g^{-1} g_j) \\
&= \psi_{i,a}(g_{g^{-1}\cdot j} h(g^{-1}, j)) \\
&= \delta_{i,g^{-1}\cdot j} \rho(h(g^{-1}, j)^{-1}) \cdot v_a
\end{aligned}
\tag{11.571}
$$

Fortunately, we are only interested in the trace of this $G$-action. The first key point is that *only the fixed points of the $g$-action on $G/H$ contribute.* Note that the RHS above is supported at $j = g \cdot i$, but if we are taking the trace we must have $i = j$. But in this case $gg_i = g_i h(g,i)$ and hence $g^{-1} g_i = g_i h(g,i)^{-1}$ so for fixed points we can simplify $h(g^{-1}, i) = h(g,i)^{-1}$, and hence when we take the trace the contribution of a fixed point $gg_iH = g_iH$ is the trace in the $H$-rep of $h(g,i) = g_i^{-1} g g_i$, as was to be shown ♠

**Remark**: The *Ind* map does not extend to a ring homomorphism of representation rings.

**Remark**: Let us check the Frobenius reciprocity theorem in the explicit example $\mathrm{Ind}_{S_2}^{S_3}(V(\epsilon))$ we worked out above.

*****************************************

NEED TO CLEAN UP THIS TEXT:

Now let us look at Frobenius reciprocity. The irreducible representations of $G$ are $W(\epsilon)$ defined by $(ij) \to \epsilon$, and $W_2$ defined by the symmetries of the equilateral triangle, embedded into $O(2)$:

$$
\begin{aligned}
\rho_2(12) &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\
\rho_2(123) &= \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}
\end{aligned}
\tag{11.572}
$$

As $H = \mathbb{Z}_2$ representations, we have $W(\epsilon) \cong V(\epsilon)$ and

$$
W_2 \cong V(+1) \oplus V(-1) \tag{11.573}
$$

Therefore

$$
\begin{aligned}
\mathrm{Hom}_H(W(\epsilon), V(\epsilon')) &= \delta_{\epsilon,\epsilon'} \mathbb{R} \\
\mathrm{Hom}_H(W_2, V(\epsilon)) &= \mathbb{R}
\end{aligned}
\tag{11.574}
$$

By Frobenius reciprocity we have

*****************************************************

### 11.19 Representations Of $SU(2)$

#### 11.19.1 Homogeneous Polynomials

We now use the idea of induced representations to construct representations of $SU(2)$. In fact, we will construct all the irreducible representations.

♣Using $\rho$ rather than $T$ for representations here. Uniformize? ♣

We take $G = SU(2)$ and we take the subgroup $H \subset SU(2)$ to be a "maximal torus" namely the subgroup of diagonal $SU(2)$ matrices. We will denote it by $\mathbb{T}$. As a group $\mathbb{T} \cong U(1)$. For $V$ we choose an irreducible representation $\rho_k$ of $U(1)$ where $k \in \mathbb{Z}$ and the carrier space is $V \cong \mathbb{C}$. Thus we choose the one-dimensional representation of $\mathbb{T}$ defined by:

♣Cannot use $\mathcal{D}$ because that looks like Wigner functions. ♣

$$\rho_k \left( \begin{pmatrix} z & 0 \\ 0 & z^{-1} \end{pmatrix} \right) := z^k \tag{11.575}$$

where $z \in U(1)$ and $k \in \mathbb{Z}$. Then the induced representation

$$\mathrm{Ind}_{U(1)}^{SU(2)}(\rho_{-k}) \tag{11.576}$$

is, by definition the space of functions $F : SU(2) \to V = \mathbb{C}$ satisfying the equivariance condition

$$F(\begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix} \begin{pmatrix} e^{\mathrm{i}\theta} & \\ & e^{-\mathrm{i}\theta} \end{pmatrix}) = e^{\mathrm{i}k\theta} F(\begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix}) \tag{11.577}$$

We can abbreviate

$$F(\begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix}) \Rightarrow F(u, v) \tag{11.578}$$

so we just have

$$F(ue^{\mathrm{i}\theta}, ve^{\mathrm{i}\theta}) = e^{\mathrm{i}k\theta} F(u, v) \tag{11.579}$$

Of course $g \in SU(2)$ implies $|u|^2 + |v|^2 = 1$. So we can also view this as the space of equivariant functions $S^3 \to \mathbb{C}$ for the right-action of $U(1)$.

Actually, we don't want <u>all</u> functions - that space is too big. One useful subspace to restrict to is $L^2(S^3)$. Still the induced representation is infinite dimensional and very far from being irreducible. For example, for $k \geq 0$, the functions $u^k |u|^\ell$ for $\ell$ are all $L^2$ functions on $SU(2)$ and are independent vectors in $L^2(SU(2))$.

It turns out we can cut down the $\rho_k$-equivariant functions $L^2(SU(2))$ down to a finite dimensional irrep of $SU(2)$ by imposing one more condition. It is a condition of *holomorphy* as we now explain.

♣Later, when we have explained Casimirs we can also cut down the induced rep using the Casimirs. This is closer to what is done in writing the irreps of the Poincare group. ♣

One way to view this restriction is first to view $S^3 \subset \mathbb{C}^2$ and consider instead the space of all functions of two variables

$$\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{C}^2 \tag{11.580}$$

which are equivariant in the sense of equation (11.579). This larger space of functions is also a representation of $SU(2)$ and also we can restrict functions to give equivariant functions on $SU(2)$. But now we can look at the subspace of equivariant functions which are restrictions of <u>holomorphic functions</u> on $\mathbb{C}^2$ to the sphere. This will be a finite-dimensional space of functions.

Holomorphy requires that $k \geq 0$. Such homogeneous holomorphic functions must be polynomials in $u, v$. It is convenient to set $k = 2j \in \mathbb{Z}_+$. The restriction to the holomorphic equivariant functions gives the finite dimensional space $\mathcal{H}_{2j}$ of homogeneous polynomials in $u, v$ of degree $2j$. In physics this is called the *spin-$j$ representation* for reasons explained below. In general we will denote the isomorphism class of the spin-j representation by $V_j$.

There is a second, very useful, way to think about the restriction to holomorphic functions. We now pause to explain that other viewpoint in detail. Recall that:

$$SU(2)/U(1) \cong SL(2,\mathbb{C})/B \cong \mathbb{CP}^1 \tag{11.581}$$

where $B$ is the subgroup of upper triangular matrices.

$$B = \{ \begin{pmatrix} z & w \\ 0 & z^{-1} \end{pmatrix} | z \in \mathbb{C}^*, w \in \mathbb{C} \} \tag{11.582}$$

Both quotients are $\mathbb{CP}^1$. But the latter description is useful because we can introduce the idea of *holomorphy*. Note that the representation $\rho_k$ of $U(1)$ extends uniquely to a holomorphic representation of $B$ for all $k \in \mathbb{Z}$:

$$\rho_k( \begin{pmatrix} z & w \\ 0 & z^{-1} \end{pmatrix} ) = z^k \tag{11.583}$$

Now consider

$$\mathrm{Ind}_B^{SL(2,\mathbb{C})}(\rho_{-k}) \tag{11.584}$$

Once again we can consider the equivariant functions:

$$F(gb) = \rho_{-k}(b^{-1})F(g) \qquad\qquad b \in B \tag{11.585}$$

Now $g \in SL(2,\mathbb{C})$ looks like

$$g = \begin{pmatrix} u & s \\ v & t \end{pmatrix} \qquad\qquad ut - vs = 1 \tag{11.586}$$

and any three matrix elements will serve as holomorphic coordinates as long as we can solve for the fourth. However we have matrices

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \in B \tag{11.587}$$

and equivariance under these matrices implies that

$$F\left( \begin{pmatrix} u & s \\ v & t \end{pmatrix} \right) = F\left( \begin{pmatrix} u & s + ux \\ v & t + vx \end{pmatrix} \right) \tag{11.588}$$

and hence $B$-equivariant functions on $SL(2,\mathbb{C})$ are again functions of just the vector

$$\begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{C}^2 - \{0\} \ . \tag{11.589}$$

such that

$$F( \begin{pmatrix} u \\ v \end{pmatrix} z) = z^k F( \begin{pmatrix} u \\ v \end{pmatrix}) \tag{11.590}$$

for $z \in \mathbb{C}^*$. Such functions restrict to the functions on the 3-sphere $|u|^2 + |v|^2 = 1$ and a function equivariant for $B$ by $\rho_{-k}$ restricts to a function on $SU(2)$, equivariant for $\mathbb{T}$ by $\rho_{-k}$ since

$$B \cap SU(2) = \mathbb{T} \tag{11.591}$$

Conversely we can use the equivariance to extend the functions in $\mathrm{Ind}_{U(1)}^{SU(2)}(\rho_{-k})$ to functions in $\mathrm{Ind}_{B}^{SL(2,\mathbb{C})}(\rho_{-k})$ so they are really the same representation:

$$\mathrm{Ind}_{B}^{SL(2,\mathbb{C})}(\rho_{-k}) \cong \mathrm{Ind}_{U(1)}^{SU(2)}(\rho_{-k}) \tag{11.592}$$

The great advantage of the description $\mathrm{Ind}_{B}^{SL(2,\mathbb{C})}(\rho_{-k})$ is that in this description it is manifest that there is a subspace of <u>holomorphic equivariant functions</u>: Such functions extend uniquely to holomorphic functions on all of $\mathbb{C}^2$ (by a theorem known as Hartog's theorem) so we are simply considering holomorphic functions on $\mathbb{C}^2$. In this way we have found a nice finite dimensional subspace of $\mathrm{Ind}_{B}^{SL(2,\mathbb{C})}(\rho_{-k})$ of equivariant holomorphic functions.

We should check that the $SU(2)$ action preserves the holomorphy condition. This is easy: note that if $g \in SU(2)$, we can write:

$$g = \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} \qquad |\alpha|^2 + |\beta|^2 = 1 \tag{11.593}$$

and then

$$g \cdot \begin{pmatrix} u & -\bar{v} \\ v & \bar{u} \end{pmatrix} = \begin{pmatrix} \bar{\alpha}u + \bar{\beta}v & * \\ -\beta u + \alpha v & * \end{pmatrix} \tag{11.594}$$

so

$$(g \cdot F)(u, v) = F(\bar{\alpha}u + \bar{\beta}v, -\beta u + \alpha v) \tag{11.595}$$

so if $F(u, v)$ is a restriction to $S^3$ of a holomorphic function on $\mathbb{C}^2$ then so is $(g \cdot F)$.

Therefore, the space of homogeneous polynomials of degree $2j = k$ in $u, v$ forms a nice representation of $SU(2)$. (The space is empty if $k < 0$.) Denote this space by $\mathcal{H}_k$.

A basis for $\mathcal{H}_k$ is given by $u^k, u^{k-1}v, \dots, uv^{k-1}, v^k$. To make contact with physics we define $j := k/2$ and write this exact same basis as:

$$\tilde{f}_{j,m}(u, v) := u^{j+m} v^{j-m} \tag{11.596}$$

for $m = -j, -j+1, -j+2, \cdots, j-1, j$. Note that $m$ increases in steps of $+1$ and hence $j \pm m$ is always an integer even though $j, m$ might be half-integer. We will also denote this representation of $SU(2)$ as $V_j$. Thus $V_j \cong \mathcal{H}_{2j}$ is a complex vector space of dimension $2j + 1$. we compute the matrix elements for this representation of $SU(2)$ relative to this basis via:

$$\begin{aligned} (g \cdot \tilde{f}_{j,m})(u, v) &:= \tilde{f}_{j,m}(\bar{\alpha}u + \bar{\beta}v, -\beta u + \alpha v) \\ &= (\bar{\alpha}u + \bar{\beta}v)^{j+m}(-\beta u + \alpha v)^{j-m} \\ &:= \sum_{m'} \tilde{D}^j_{m'm}(g) \tilde{f}_{j,m'} \end{aligned} \tag{11.597}$$

Note that the basis $\tilde{f}_{j,m}$ diagonalizes the action of the diagonal $SU(2)$ matrices with $\beta = 0$:

$$g \cdot \tilde{f}_{j,m} = \bar{\alpha}^{j+m} \alpha^{j-m} \tilde{f}_{j,m} = \alpha^{-2m} \tilde{f}_{j,m} \tag{11.598}$$

That is,

$$\tilde{D}^j_{m_L,m_R}(g) = \delta_{m_L,m_R} \alpha^{-2m_R} \tag{11.599}$$

when $\beta = 0$.

More generally, we can derive an explicit formula for the matrix elements $\tilde{D}^j_{m'm}(g)$ as functions on $SU(2)$ by expanding out the two factors in (11.597) using the binomial theorem and collecting terms:

$$\tilde{D}^j_{m'm}(g) = \sum_{s+t=j+m'} \binom{j+m}{s} \binom{j-m}{t} \bar{\alpha}^s \alpha^{j-m-t} \bar{\beta}^{j+m-s} (-\beta)^t \tag{11.600}$$

Here the sum is over integers $s \geq 0$ and $t \geq 0$ and we recall that $\binom{a}{b} = 0$ when $b > a$ so this is a finite sum. We recover the functions $\tilde{f}_{j,m}$, up to scale from

$$\tilde{D}^j_{-m,-j}(g) = (-1)^{j-m} \binom{2j}{j-m} \alpha^{j+m} \beta^{j-m} \tag{11.601}$$

Note they are holomorphic in $\alpha, \beta$, while the general functions $\tilde{D}^j_{m_L,m_R}$ are not.

We claim that the representations $\mathcal{H}_{2j}$ are irreducible, and moreover give a full set of irreducible representations of $SU(2)$. We will prove this using characters in the next section.

**Remarks**

1. *Representations Of $SO(3)$ That Lift To Representations Of $SU(2)$*

   Since $SO(3) \cong SU(2)/Z$ where $Z \cong \mathbb{Z}_2$ is the central subgroup consisting of $SU(2)$ matrices proportional to the unit matrix: $Z = \{\pm 1_{2\times 2}\}$, we can easily determine the irreducible representations of $SO(3)$. On the spin $j$ representation $T_j$ the central element acts as $T_j(-1) = (-1)^{2j} 1_{V_j}$. Therefore: *the irreducible representations of $SO(3)$ are given by $V_j$ for $j \in \mathbb{Z}$, that is, the representations where $\dim_{\mathbb{C}} V_j$ is odd. The other irreducible representations of $SU(2)$ with $\dim_{\mathbb{C}} V_j$ even are not representations of $SO(3)$, although they are projective representations - see section **** below.*

2. Returning to the induced representation $\mathrm{Ind}_{U(1)}^{SU(2)}(\rho_{-k})$ of smooth (not holomorphic) equivariant functions on $SU(2)$ we see that the functions $\tilde{D}$ satisfy the equivariance condition

$$\tilde{D}^j_{m_L,m_R}(g \begin{pmatrix} e^{\mathrm{i}\theta} & \\ & e^{-\mathrm{i}\theta} \end{pmatrix}) = (e^{\mathrm{i}\theta})^{-2m_R} \tilde{D}^j_{m_L,m_R}(g) \tag{11.602}$$

   and hence we get the equivariance condition for $2m_R = -k$. Such values of $m_R$ can only appear for $j \geq |k|/2$ with $j = |k|/2 \mathrm{mod} 1$. So the induced representation is spanned by the functions $\tilde{D}^{j}_{m_L,-k/2}$ with $j \geq |k|/2$ and $j = |k|/2 \mathrm{mod} 1$ and $m \in \{-j, -j+1, \dots, j-1, j\}$. Thus as a representation of $SU(2)$ we have:

$$\mathrm{Ind}_{U(1)}^{SU(2)}(\rho_{-k}) \cong V_{|k|/2} \oplus V_{|k|/2+1} \oplus \cdots \tag{11.603}$$

When $k$ is nonnegative we can set $k = 2j_0$ and the span of the set of functions $D^{j_0}_{m_L,-j_0}$ where $m$ ranges over $-j_0, -j_0+1, \dots, j_0-1, j_0$ transforms under the left regular representation in a representation isomorphic to $V_j$. These are the functions with a holomorphic interpretation. When $k$ is negative there is no holomorphic interpretation.

3. It is instructive to evaluate the functions $\tilde{D}^{j}_{m'm}(g)$ where $g$ is parametrized by Euler angles. We find:

$$\tilde{D}^{j}_{m_L,m_R}(g) = e^{-\mathrm{i}\phi m_L} \tilde{P}^{j}_{m_L,m_R}(\theta) e^{-\mathrm{i}\psi m_R} \tag{11.604}$$

♣SAY WHICH PARAMETRIZATION. WE USED TWO DIFFERENT ONES ABOVE ♣

where $\tilde{P}^{j}_{m_L,m_R}(\theta)$ is a polynomial in $\cos(\theta/2)$ and $\sin(\theta/2)$:

$$\tilde{P}^{j}_{m_L,m_R}(\theta) = (-\mathrm{i})^{2j+m_L+m_R} \sum_{s+t=j+m_L} (-1)^s \binom{j+m_R}{s} \binom{j-m_R}{t} (\cos\theta/2)^{j-m_R+s-t} ((\sin\theta/2)^{j+m_R+t-s} \tag{11.605}$$

It is closely related to an *associated Legendre polynomial* and the functions $\tilde{D}^{j}_{m'm}$ are closely related to *Wigner functions*. We will explain more about this below.

4. Looking ahead to the relation to Lie algebras, diagonal $SU(2)$ matrices can be written as $g = \exp[-\mathrm{i}\sigma^3\phi]$ and then

$$g \cdot \tilde{f}_{j,m} = e^{\mathrm{i}2m\phi} \tilde{f}_{j,m} \tag{11.606}$$

so the basis $\tilde{f}_{j,m}$ is proportional to the physicist's basis $|j,m\rangle$. Note that

$$\begin{pmatrix} e^{\mathrm{i}\phi} & \\ & e^{-\mathrm{i}\phi} \end{pmatrix} = \exp[-2\phi\left(-\frac{\mathrm{i}}{2}\sigma^3\right)] \tag{11.607}$$

so this basis diagonalizes the generator $T^3 = -\frac{\mathrm{i}}{2}\sigma^3$, that is $\rho(T^3) \cdot \tilde{f}_{j,m} = -m\tilde{f}_{j,m}$.

---

**Exercise** *Explicit Matrices For Small j*

a.) Show that for $j = 1/2$, relative to the ordered basis $\{\tilde{f}_{+1/2}, \tilde{f}_{-1/2}\}$ we have

$$\tilde{D}^{j=1/2}(g) = \begin{pmatrix} \bar{\alpha} & \bar{\beta} \\ -\beta & \alpha \end{pmatrix} = g^* \tag{11.608}$$

b.) Show that for $j = 1$, relative to the ordered basis $\{\tilde{f}_1, \tilde{f}_0, \tilde{f}_{-1}\}$ we have

$$\tilde{D}^{j=1}(g) = \begin{pmatrix} \bar{\alpha}^2 & -\bar{\alpha}\beta & \beta^2 \\ 2\bar{\alpha}\bar{\beta} & |\alpha|^2 - |\beta|^2 & -2\alpha\beta \\ \bar{\beta}^2 & \alpha\bar{\beta} & \alpha^2 \end{pmatrix} \tag{11.609}$$

**Exercise** *Induced Representations From A Finite Subgroup Of* $SU(2)$

Consider the subgroup of $\mathbb{T}$ of phases which are in $\mu_N$. Using the reasoning that led to (11.603) describe the induced representation $\mathrm{Ind}_{\mu_N}^{SU(2)}(\rho)$ where $\rho$ is the defining representation of $\mu_N$.

♣Need to provide answer in a footnote. ♣

### 11.19.2 Chararacters Of The Representations $V_j$

Every $g \in SU(2)$ is diagonalizable so we can say

$$g \sim d(z) := \begin{pmatrix} z & \\ & z^{-1} \end{pmatrix} \tag{11.610}$$

where $|z| = 1$ and both $z$ and $z^{-1}$ define the same conjugacy class [177]

For the spin $j$ representation denote the character by $\chi_j$. In the basis described above $\tilde{D}^j(d(z))$ is diagonal and given by

$$\tilde{D}_j(d(z)) = \mathrm{Diag}\{z^{-2j}, z^{-2j+2}, \dots, z^{2j-2}, z^{2j}\} \tag{11.611}$$

Therefore

$$\chi_j(g) = z^{-2j} + z^{-2j+2} + \cdots + z^{2j-2} + z^{2j} = \frac{z^{2j+1} - z^{-2j-1}}{z - z^{-1}} \tag{11.612}$$

Now we write the orthogonality relations for the characters. We have written the Haar measure above in Euler angles, but this is not the most convenient form for integrating class functions. Given the Euler angles $(\phi, \theta, \psi)$ of $u \in SU(2)$ it is not so evident how to find the value of the angle $\xi$ such that

$$g = \cos\xi + \mathrm{i}\sin\xi \hat{n}\cdot\vec{\sigma} = \exp[\mathrm{i}\xi\hat{n}\cdot\vec{\sigma}] \tag{11.613}$$

The conjugacy class with angle $\xi$ is a two-dimensional sphere in $S^3$ with radius $\sin\xi$, and the range of $\xi$ is $[0, \pi]$.

Since the area of an $S^2$ with radius $\sin\xi$ goes like $\sin^2\xi$ we can say that the properly normalized measure on $SU(2)$ is such that on a class function $F$ we have:

♣Explain this better ♣

$$\int_{SU(2)} F(u)[du] = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\sin^2\theta d\theta = -\frac{1}{4\pi\mathrm{i}}\oint g(z)(z - z^{-1})^2\frac{dz}{z} \tag{11.614}$$

where we have changed $\xi$ to $\theta$ (not to be confused with an Euler angle) and

$$F(\begin{pmatrix} e^{\mathrm{i}\theta} & \\ & e^{-\mathrm{i}\theta} \end{pmatrix}) = f(\theta) = g(z) \tag{11.615}$$

[177] $d(z) = vd(z^{-1})v^{-1}$ where, for example, we can take $v = \mathrm{i}\sigma^1$.

with $z = e^{i\theta}$. Note that for a class function $f(\theta+2\pi) = f(-\theta) = f(\theta)$ allowing us to extend the integration range over $\theta$ from $[0, \pi]$ to $[0, 2\pi]$.

Now, the space of class functions $L^2(SU(2))^{class}$ can be identified with the (completion of) the space of Laurent polynomials in $z$ which is even under the involution $z \to 1/z$. This space is clearly spanned by $z^{2j} + z^{-2j}$, and that basis is related to the $\chi_j$ by an upper triangular matrix. Moreover, one easily confirms the general relation

$$\langle \chi_j, \chi_{j'} \rangle = \delta_{j,j'} \tag{11.616}$$

From this we learn that we have a complete set of representations. Moreover, now we can prove that the $V_j$ are irreducible. We do this by induction on $j$. If $V_j$ were a direct sum of smaller representations, then its character would be a linear combination of $\chi_{j'}$ with $j' < j$ and nonnegative integer coefficients. This is incompatible with the orthogonality of the functions $\chi_j$.

**Remarks**

1. If we write $k = 2j$ and $z = e^{i\theta}$ then

$$\chi_j(u) = \frac{\sin((k+1)\theta)}{\sin\theta} = U_{k+1}(\cos\theta) \tag{11.617}$$

   where $U_{k+1}(x)$ are known as the *Chebyshev polynomials of the second kind.*

2. *Weyl Density Formula.* For more details about the following see Chapter 5, Survey of Matrix Groups. The formula for integrating class functions with the Haar measure has a nice analog for all the classical groups. For example, if $F : SU(n) \to \mathbb{C}$ is a class function then $F(u)$ only depends on the eigenvalues of $u$. Write the unordered set of eigenvalues as $\{e^{i\theta_1}, \dots, e^{i\theta_n}\}$. Then define

$$F(u) = f(\theta_1, \dots, \theta_n) \tag{11.618}$$

   and then, we have

♣Should check this reduces to the above formula for $n = 2$. ♣

$$\int_{SU(n)} F(u)du = \frac{1}{n!} \int_0^{2\pi} f(\theta_1, \dots, \theta_n) \prod_{1 \le i < j \le n} |e^{i\theta_i} - e^{i\theta_j}|^2 \prod_i \frac{d\theta_i}{2\pi} \tag{11.619}$$

   and there are similar formulae for $SO(n)$ and $USp(2n)$. In $SO(n)$ we can conjugate any matrix to the form

$$Diag\{R(\theta_1), \dots, R(\theta_r)\} \tag{11.620}$$

   when $n = 2r$ is even and

$$Diag\{R(\theta_1), \dots, R(\theta_r), 1\} \tag{11.621}$$

   when $n = 2r + 1$ is odd. We then have

$$\int_{SO(2n)} f[dg] = \frac{2^{(n-1)^2}}{\pi^n n!} \int_{-\pi}^{\pi} \prod_{1 \le j < k \le n} (\cos\theta_j - \cos\theta_k)^2 f(\theta_1, \dots, \theta_n) \prod d\theta_i \tag{11.622}$$

$$\int_{SO(2n+1)} f[dg] = \frac{2^{n^2}}{\pi^n n!} \int_{-\pi}^{\pi} \prod_{1\le j<k\le n} (\cos\theta_j - \cos\theta_k)^2 \prod_{j=1}^{n} \sin^2(\theta_j/2) f(\theta_1,\dots,\theta_n) \prod d\theta_i \tag{11.623}$$

Where we have normalized the Haar measure to have volume one. Finally, for the unitary symplectic group $USp(2r)$, any matrix can be conjugated to

$$Diag\{e^{i\theta_1},\dots,e^{i\theta_r},e^{-i\theta_1},\dots,e^{-i\theta_r}\} \tag{11.624}$$

in terms of which the formula for integrating class functions is:

$$\int_{USp(2n)} f[dg] = \frac{2^{n^2}}{\pi^n n!} \int_{-\pi}^{\pi} \prod_{1\le j<k\le n} (\cos\theta_j - \cos\theta_k)^2 \prod_{j=1}^{n} \sin^2(\theta_j) f(\theta_1,\dots,\theta_n) \prod d\theta_i \tag{11.625}$$

These kinds of formulae appear very frequently in the subject of random matrix theory.

♣Need to give a reference here. ♣

3. *Weyl-Kac Character Formula* The character formula for $SU(2)$ can be written as

$$\chi_j = e^{-\rho}\frac{\sum_{w\in W}\epsilon(w)e^{w(\lambda+\rho)}}{\prod_{\alpha>0}(1-e^{-\alpha})} \tag{11.626}$$

where $e^\lambda$, $e^\alpha$ and $e^\rho$ are functions on the maximal torus such that

$$e^\lambda(d(z)) = z^{2j} \tag{11.627}$$

$$e^\alpha(d(z)) = z^2 \qquad e^\rho = e^{\alpha/2} \tag{11.628}$$

and the sum on $w$ is a sum over the Weyl group, with $\epsilon(w)\in\{\pm1\}$ a homomorphism to $\mu_2$.

### 11.19.3 Unitarization

As we have seen, the finite-dimensional representations of $SU(2)$ are unitarizable, and unitarity is very important in physics and mathematics.

One might first be tempted to define a unitary structure on $\mathcal{H}_{2j}$ by declaring the inner product of two homogeneous polynomials to be the integral of $\psi_1^*\psi_2$. But this will not converge. Instead we take

$$\langle\psi_1,\psi_2\rangle_{\mathcal{H}_{2j}} := \frac{1}{\pi(2j+1)!}\int_{\mathbb{C}^2}\psi_1(u,v)^*\psi_2(u,v)e^{-|u|^2-|v|^2}d^2u d^2v \tag{11.629}$$

This will give finite overlaps on $\mathcal{H}_{2j}$. Note that since $du\wedge dv$ and $|u|^2+|v|^2$ are $SU(2)$ invariant the inner product is $SU(2)$-invariant, i.e. it is unitary:

$$\langle g\cdot\psi_1, g\cdot\psi_2\rangle_{\mathcal{H}_{2j}} = \langle\psi_1,\psi_2\rangle_{\mathcal{H}_{2j}} \tag{11.630}$$

The reason for the funny normalization constant will be clearer below.

---

**Exercise** *ON Basis*
Show that an ON basis is given by

$$f_{j,m} = \frac{1}{\sqrt{\pi}} \sqrt{\frac{(2j+1)!}{(j+m)!(j-m)!}} u^{j+m} v^{j-m} \tag{11.631}$$

---

### 11.19.4 Inhomogeneous Polynomials And Mobius Transformations On $\mathbb{CP}^1$

We saw in section 8.4 above that there is a close relation between $SU(2)/U(1)$ and $SL(2,\mathbb{C})/B$ and $\mathbb{CP}^1$. Indeed, $SL(2,\mathbb{C})$ acts naturally on $\mathbb{CP}^1$:

$$\begin{pmatrix} \alpha & s \\ \beta & t \end{pmatrix} \cdot [z_1 : z_2] \to [\alpha z_1 + s z_2 : \beta z_1 + t z_2] \tag{11.632}$$

On the other hand we can identify $\mathbb{CP}^1$ with $S^2$ and then we can stereographically project $\mathbb{CP}^1$ to $\mathbb{C} \cup \{\infty\}$. See the exercise below for the precise formulae. One projection from $\mathbb{CP}^1$ to the extended complex plane is just:

$$[z_1 : z_2] \mapsto \begin{cases} z := z_1/z_2 & z_2 \neq 0 \\ \infty & z_2 = 0 \end{cases} \tag{11.633}$$

(Of course, there are other projections. For example we could map to $z_2/z_1$. We will come back to this point.) It follows that $SL(2,\mathbb{C})$ acts on the Riemann sphere by Mobius transformations:

$$\begin{pmatrix} \alpha & s \\ \beta & t \end{pmatrix} \cdot z \to \frac{\alpha z + s}{\beta z + t} \tag{11.634}$$

and restricted to $SU(2)$ this becomes:

$$\begin{pmatrix} \alpha & -\beta^* \\ \beta & \alpha^* \end{pmatrix} \cdot z \to \frac{\alpha z - \beta^*}{\beta z + \alpha^*} \tag{11.635}$$

Since the $SU(2)$ action on $S^2$ factors through the usual $SO(3)$ rotations it is not surprising that the round volume form is invariant under $SU(2)$. We can verify this with the volume form

$$\omega := \frac{1}{2\pi} \frac{\mathrm{i} dz \wedge d\bar{z}}{(1+|z|^2)^2} \tag{11.636}$$

directly since if $\tilde{z} = g \cdot z$ then one easily checks:

$$d\tilde{z} = \frac{dz}{(\beta z + \alpha^*)^2} \tag{11.637}$$

$$\frac{1}{(1+|\tilde{z}|^2)^2} = \frac{|\beta z + \alpha^*|^2}{(1+|z|^2)^2} \tag{11.638}$$

and therefore the measure $\omega$ is invariant. Indeed, as you show in an exercise below, it is just the usual round measure on $S^2$ evaluated in stereographic coordinates.

Now let us return to the space $\mathcal{H}_{2j}$ of homogeneous polynomials in $z_1, z_2$. This vector space is canonically isomorphic to the vector space of polynomials in $z = z_1/z_2$ of degree $\leq 2j$ since we can identify a homogenous polynomial $F$ with a polynomial $p$ by

$$p(z) := z_2^{-2j} F(z_1, z_2) \tag{11.639}$$

The map $F \to p$ defines a map $\mathcal{H}_{2j} \to \mathcal{P}_{2j}$ and we can make this an $SU(2)$-equivariant map if we declare the action of $SU(2)$ on polynomials $p$ to be [178]

$$(g \cdot p)(z) := (-\beta z + \alpha)^{2j} p\left(\frac{\alpha^* z + \beta^*}{-\beta z + \alpha}\right) \tag{11.640}$$

The model for the spin $j$ representation $V_j$ as the space of polynomials of degree $\leq 2j$ can be turned into a Hilbert space by defining

$$h(\psi_1, \psi_2) := (1 + |z|^2)^{-2j} \psi_1^*(z) \psi_2(z) \tag{11.641}$$

One checks that this is invariant under the $SU(2)$ action so we integrate:

$$\langle \psi_1, \psi_2 \rangle_{\mathcal{P}_{2j}} = \int_{\mathbb{C}} h(\psi_1, \psi_2) \omega \tag{11.642}$$

and then

$$\langle g \cdot \psi_1, g \cdot \psi_2 \rangle = \langle \psi_1, \psi_2 \rangle \tag{11.643}$$

so $g$ acts as a unitary operator on $\mathcal{P}_{2j}$ in this inner product.

Using the integrals

$$\int_0^\infty \frac{x^{n_1}}{(1+x)^{n_2}} dx = \frac{\Gamma(n_1+1)\Gamma(n_2 - n_1 - 1)}{\Gamma(n_2)} \qquad -1 < n_1 \quad \& \quad 1 < n_2 - n_1 \tag{11.644}$$

we check that an ON basis for $\mathcal{P}_{2j}$ is given by:

$$\psi_\ell := N_\ell z^\ell \qquad \ell = 0, \ldots, 2j \tag{11.645}$$

$$N_\ell := \sqrt{(2j+1)\binom{2j}{\ell}} \tag{11.646}$$

In physics this is usually written as:

$$\psi_{j,m} = \sqrt{\frac{(2j+1)!}{(j+m)!(j-m)!}} z^{j+m} = N_{j,m} z^{j+m} \tag{11.647}$$

Now the Wigner functions are defined by the matrix elements relative to this ON basis:

$$g \cdot \psi_{j,m_R} = \sum_{m_L} D^j_{m_L, m_R}(g) \psi_{j,m_L} \tag{11.648}$$

In physics notation $\psi_{j,m}$ would usually be written as the ket vector $|j, m\rangle$, the spin $j$ representation might be written as operators $T^j(g)$ and we would have:

$$D^j_{m'm}(g) := \langle j, m' | T^j(g) | j, m \rangle \tag{11.649}$$

**Remarks**:

[178] One needs to be careful to remember that the left $g$-action on functions takes the argument $z$ of the function to $g^{-1} z$. Note it is $g^{-1}$ and not $g$.

1. There is an isometry of Hilbert spaces between $\mathcal{H}_{2j}$ with inner product $\langle\cdot,\cdot\rangle_{\mathcal{H}_{2j}}$ defined in equation (11.629) above and $\langle\cdot,\cdot\rangle_{\mathcal{P}_{2j}}$. Indeed, in the integral (11.629) we can change coordinates to $z = u/v$ and $v$. Then $du \wedge dv = v dz \wedge dv$ and one can do the $v$ integral explicitly to show the isometry.

2. To relate the Wigner function $D^j_{m'm}$ to our functions $\tilde{D}^j_{m'm}$ defined above we compute

$$\begin{aligned} D^j_{m',m}(g) &:= \langle \psi_{j,m'}, g\cdot\psi_{j,m}\rangle \\ &= N_{j,m'}N_{j,m}\int_{\mathbb{C}} (z^{j+m'})^*(-\beta z+\alpha)^{j-m}(\alpha^* z+\beta^*)^{j+m}\frac{d\phi}{2\pi}\frac{d(r^2)}{(1+r^2)^{2j+2}} \\ &= N_{j,m'}N_{j,m}\sum_{s+t=j+m'}\binom{j+m}{s}\binom{j-m}{t}\bar{\alpha}^s\alpha^{j-m-t}\bar{\beta}^{j+m-s}(-\beta)^t\int_0^\infty r^{2j+2m'}\frac{d(r^2)}{(1+r^2)^{2j+2}} \\ &= \sqrt{\frac{(j+m')!(j-m')!}{(j+m)!(j-m)!}}\tilde{D}^j_{m',m}(g) \end{aligned} \tag{11.650}$$

3. As we noted before, if we induce from the trivial representation then we get functions on $G/H$ which in this case is $G/H = S^2$. The trivial representation of $U(1)$ is $m_R = 0$. In this case $j = \ell$ must be an integer. The functions

$$D^\ell_{m_L,0}(g) = e^{-\mathrm{i}m_L\phi}P^\ell_{m_L}(\theta) = Y_{\ell,m}(\theta,\phi) \tag{11.651}$$

are known as *spherical harmonics*. From what we have said they form a complete orthonormal set of functions on $S^2$ and are widely used in electromagnetism and quantum mechanics.

4. If we specialize further to $m_L = 0$ we get the famous Legendre polynomials

$$P_\ell(\cos\theta) = 2^{-\ell}\sum_{s=0}^{\ell}\binom{\ell}{s}^2(-1)^{\ell-s}(1+\cos\theta)^s(1-\cos\theta)^{\ell-s} \tag{11.652}$$

whereas the $P_{\ell,m}(\theta)$ are known as associated Legendre functions. Now some addition theorems become transparent from group theory. For example, if we consider $g_1 = e^{\mathrm{i}\theta_1\sigma^1/2}$ and $g_2 = e^{\mathrm{i}\theta_2\sigma^1/2}$ then

$$D^\ell_{00}(g_1g_2^{-1}) = \sum_m D^\ell_{0m}(g_1)D^\ell_{m0}(g_2^{-1}) \tag{11.653}$$

becomes one of the identities for Legendre polynomials known as the "addition theorem." See, for example, J.D. Jackson, *Classical Electrodynamics*, section 3.6. [179]

5. The Wigner functions have many applications in physics. As we will see, from group theory, they satisfy some nice differential equations and hence appear in the wavefunctions of atoms. They also appear in the study of the quantum Hall effect on $S^2$

[179] We have specialized Jackson's equation 3.62. For the more general statement let $g_1 = e^{\mathrm{i}\theta_1\sigma^1/2}e^{\mathrm{i}\phi_1\sigma^3/2}$ and similarly for $g_2$.

and in the study of wavefunctions of electrons in the field of a magnetic monopole of magnetic charge $2j$. And, of course, the special case of spherical harmonics have a vast number of applications in many diverse fields.

---

**Exercise** *Relations Between Coordinates On $S^2$ And Stereographic Projection*

a.) Let $\hat{n} = (\hat{x}^1, \hat{x}^2, \hat{x}^3) = (\sin\theta\cos\phi, \sin\theta, \sin\phi, \cos\theta)$ parametrize the unit sphere $S^2$. Verify the following relations: Now we have the relation between stereographic and angular coordinates where $\hat{x}^3 = 1$ corresponds to $z = \infty$ and $\hat{x}^3 = -1$ corresponds to $z = 0$:

$$\begin{aligned} \frac{z}{1+|z|^2} &= \frac{1}{2}(\hat{x}^1 + \mathrm{i}\hat{x}^2) = \frac{1}{2}e^{\mathrm{i}\phi}\sin\theta \\ \frac{|z|^2}{1+|z|^2} &= \frac{1}{2}(1+\hat{x}^3) = \frac{1+\cos\theta}{2} \\ \frac{1}{1+|z|^2} &= \frac{1}{2}(1-\hat{x}^3) = \frac{1-\cos\theta}{2} \\ \frac{|z|^2-1}{|z|^2+1} &= \hat{x}^3 \\ z &= \frac{x_1 + \mathrm{i}x_2}{1-x_3} = e^{\mathrm{i}\phi}\cot(\frac{\theta}{2}) \end{aligned} \tag{11.654}$$

b.) Also show that the standard round metric of the unit sphere can be written as

$$ds^2 = \frac{1}{\pi}\frac{|dz|^2}{(1+|z|^2)^2} \tag{11.655}$$

and the volume form is:

$$\begin{aligned} \omega :=& \frac{1}{2\pi}\frac{\mathrm{i}dz\wedge d\bar{z}}{(1+|z|^2)^2} \\ =& \frac{1}{(1+|z|^2)^2}\frac{d\phi}{2\pi}d(|z|^2) \\ =& \frac{1}{4\pi}\sin\theta d\phi\wedge d\theta \end{aligned} \tag{11.656}$$

is the spherically symmetric unit volume measure on the sphere.

c.) Show that

$$\frac{1}{2}(1+\hat{x}\cdot\vec{\sigma}) = \frac{1}{1+|z|^2}\begin{pmatrix} |z|^2 & \bar{z} \\ z & 1 \end{pmatrix} \tag{11.657}$$

is a $2\times 2$ projection matrix that depends smoothly on $\hat{x}\in S^2$. It plays many important roles and is known as the *Bott projector*.

---

---

**Exercise** *Spherical Harmonics And Polynomials*
Show that the (rescaled) spherical harmonics can be written as

$$\tilde{D}^{\ell}_{m_L,0} = |\beta|^{2\ell} \sum_{s+t=\ell+m_L} \binom{\ell}{s}\binom{\ell}{t}(-1)^t \bar{z}^s z^{\ell-t} \tag{11.658}$$

where $z = \alpha/\beta$.

---

### 11.19.5 The Geometrical Interpretation Of $\mathcal{P}_{2j}$

We discussed in section 11.18.1 above that equivariant functions can be viewed as sections of an associated bundle. In this section we unpack that idea a bit for the case of $SU(2)$.

In the case of $SU(2)/U(1)$ we have the line bundle

$$\mathcal{L}_k = (SU(2) \times V_{-k})/U(1) \tag{11.659}$$

where $V_{-k}$ is the carrier space of the representation $\rho_{-k}$ of $U(1)$. As a vector space $V_{-k} \cong \mathbb{C}$.

Abstractly, a section is a map $s : SU(2)/U(1) \to \mathcal{L}_k$ so that $\pi \circ s = Id$. This can be made much more explicit using *local trivializations*.

To motivate the description of sections in terms of local trivializations consider the transition we made above from homogenous functions of the $(z_1, z_2)$ to polynomials in $z = z_1/z_2$.

♣SOME REPETITION OF THE FOLLOWING PARAGRAPH IN THE SECTION ON STABILIZER ORBIT ♣

If we think of $\mathbb{CP}^1$ as the space of equivalence classes $[z_1 : z_2]$ then on the "patch" $\mathcal{U}_N$ where $z_2 \neq 0$ we can define a map

$$\phi_N : \mathcal{U}_N \to \mathbb{C} \qquad \phi_N([z_1 : z_2]) := z_N := z_1/z_2 \tag{11.660}$$

Note that the point $[1 : 0]$ corresponds to a "point at infinity" in this mapping. In terms of the identification $S^2 \cong \mathbb{CP}^1$ this corresponds to stereographic projection from the north pole:

$$z_N = \frac{\hat{x}^1 + \mathrm{i}\hat{x}^2}{1 - \hat{x}^3} \tag{11.661}$$

In this language we identified a homogeneous polynomial of $z_1, z_2$ of degree $2j$ with a polynomial $p_N(z_N)$ by

$$p_N(z_N) = z_2^{-2j} F(z_1, z_2) \tag{11.662}$$

We could clearly have given an analogous discussion based on stereographic projection from the south pole:

$$z_S = \frac{\hat{x}^1 - \mathrm{i}\hat{x}^2}{1 + \hat{x}^3} \tag{11.663}$$

Note that

$$z_S z_N = 1 \tag{11.664}$$

In terms of $\mathbb{CP}^1$, we could introduce a 'patch" $\mathcal{U}_N$ where $z_1 \neq 0$ we can define a map

$$\phi_S : \mathcal{U}_S \to \mathbb{C} \qquad \phi_S([z_1 : z_2]) := z_S := z_2/z_1 \tag{11.665}$$

Note that the point $[0:1]$ corresponds to a "point at infinity" in this mapping. Again it is manifest that $z_N z_S = 1$ on the patch overlap $\mathcal{U}_S \cap \mathcal{U}_N$ where both functions are defined.

Now, the same homogeneous function $F$ could be used to define a different polynomial

$$p_S(z_S) = z_1^{-2j} F(z_1, z_2) \tag{11.666}$$

Comparing the equations we see that on patch overlaps we can say:

$$p_S(z_S) = z_N^{-2j} p_N(z_N) \tag{11.667}$$

or equally well

$$p_N(z_N) = z_S^{-2j} p_S(z_S) \tag{11.668}$$

Equations (11.667) and (11.668) are known as gluing conditions and $z_S^{-2j}$ and $z_N^{-2j}$ are known as transition functions. We can glue together the total space of the line bundle $\mathcal{L}_k$ by taking two trivial line bundles

$$\mathcal{U}_S \times \mathbb{C} \tag{11.669}$$

$$\mathcal{U}_N \times \mathbb{C} \tag{11.670}$$

and identifying $(z_S, v_S)$ with $(1/z_N, z_N^{-2j} v_N)$ on the patch overlaps.

Note the the function $h$ is invariant under this description:

$$h(p_S, p_S) = h(p_N, p_N) \tag{11.671}$$

because, on patch overlaps we have:

$$(1 + |z_S|^2)^{-2j} |p_S(z_S)|^2 = (1 + |z_N|^2)^{-2j} |p_N(z_N)|^2 \tag{11.672}$$

as the reader should carefully check. Thus this is a globally defined notion of the length-square of a section. It is called a *Hermitian metric on the line bundle.*

**Remarks**:

1. Recalling that $2j = -k$ the transition rules for all sections of the line bundle $\mathcal{L}_k$ can be written

$$\psi_N(z_N, \bar{z}_N) = z_S^k \psi_S(z_S, \bar{z}_S) \tag{11.673}$$

where there is now no condition of holomorphy. The space of all sections can be identified with the space of all functions in $\mathrm{Ind}_{U(1)}^{SU(2)}(\rho_k)$. This formulation makes sense for any integer $k$. When $k > 0$ there are <u>no</u> holomorphic sections, but there are plenty of $\mathcal{C}^\infty$ sections.

2. *The Borel-Weil-Bott Theorem*: Here is a rough statement: If $G$ is a compact group and $T$ a maximal subgroup then one can give $G/T$ a complex structure so that $G/T \cong G_{\mathbb{C}}/B$. Irreps of $T$ induce reps of $B$ and then the holomorphic sections of a line bundle of $G_{\mathbb{C}}/B$, corresponding to the equivalent holomorphic induced representation $Ind_B^{G_{\mathbb{C}}}(V)$ is an irreducible representation of $G$, and all such arise in this way. $G_c/B$ for other finite dimensional simple Lie groups. ♣Improve this statement. ♣

3. We will see later that $SL(2,\mathbb{C})$ is isomorphic to the Lorentz group in 3+1 dimensions. The Lorentz group acts on the set of light rays through the origin of Minkowski space, and we can identify a light ray with its point on the celestial sphere. Under this identification, the action of the Lorentz group on the set of light rays is just the Mobius action on the sphere.

4. *Representations Of Lorentz, Poincaré, and Affine Euclidean Groups.* Also, the construction of the irreducible unitary representations of Lorentz groups, and affine Euclidean and Poincaré groups proceeds using this method. (That observation goes back to Wigner and Bargmann.) Briefly, for a representation of the Poincaré group one induces from a representation of the translation group (or the translation group semidirect product with a compact rotation group). For a representation of the Lorentz group one considers the homogeneous spaces from orbits of $SO(1,d)$ in momentum space. For example, the mass shell $p^2 = m^2$ with $p^0 > 0$ can be identified with $SO(1,d)/SO(d)$. Then one induces from a representation of $SO(d)$ to produce a unitary representation of $SO(1,d)$. One source that approaches the subject from the present viewpoint is M. Carmeli, *Group Theory And General Relativity.*

♣Should expand this into a section. Irreps of Lorentz. Then irreps of semidirect products. ♣

5. Loop Groups: See the book of Pressley and Segal

6. Diffeomorphism Groups: Nice paper of A. Alekseev and S. Shatashvili

7. *Geometric Quantization.* What we have just investigated is also a basic example in the quantization of classical phase spaces which are not just symplectic vector spaces. In this case, the $S^2$ can be considered to be a symplectic manifold with symplectic form proportional to $\omega$. If the proportionality constant is linear in $j$ then, since we expect one quantum space per unit area of phase space (we set $\hbar = 1$) we expect a finite dimensional Hilbert space of dimension growing linearly in $j$. This is indeed satisfied by the formula $\dim V_j = 2j+1$ where we can view the $+1$ as a quantum correction to the semiclassical expectation. There is a general theory of quantization of certain compact phase spaces known as the theory of geometric quantization. EXPLAIN MORE.

---

**Exercise** *Tangent Bundle Of* $S^2$

The tangent bundle of $S^2$ can be complexified to be the sum of holomorphic and antihomolomorphic tangent bundles. The holomorphic tangent bundle is spanned by $\frac{\partial}{\partial z_N}$ on $\mathcal{U}_N$ and $\frac{\partial}{\partial z_S}$ on $\mathcal{U}_S$. Show that we can have

$$p_N(z_N)\frac{\partial}{\partial z_N} = p_S(Z_S)\frac{\partial}{\partial z_S} \tag{11.674}$$

and conclude that the holomorphic tangent bundle corresponds to the case $j=1$ above.

---

## 11.20 The Clebsch-Gordon Decomposition For $SU(2)$

In physics the spin $j$ representation $V_j$ shows up almost universally. Among other applications it appears in the quantum mechanical theory of spin. (See the relation to Lie algebras below for an explanation of this.)

The combination of two systems of spins naturally leads to the question of how to give an isotypical decomposition of $V_{j_1} \otimes V_{j_2}$. This is known as the Clebsch-Gordon decomposition. The general formula is:

$$V(j_1) \otimes V(j_2) \cong V(|j_1 - j_2|) \oplus V(|j_1 - j_2| + 1) \oplus \cdots\cdots \oplus V(j_1 + j_2) \tag{11.675}$$

The nontrivial element of the center, $-1 \in SU(2)$ must act in the same way on both LHS and RHS. On a spin $j$ representation $\rho_j(-1)$ is just $(-1)^{2j}$ times the identity operator. Thus, on the LHS the element $-1$ acts as $(-1)^{2(j_1+j_2)}$. It is easy to check that this is also the case for each of the summands on the RHS.

Let us give a proof of (15.589) using characters. Because a representation is uniquely determined by its character we can consider the character of $V(j_1)\otimes V(j_2)$. If we can write this as a linear combination of characters $\chi_j$ with nonnegative integer coefficients we can uniquely determine the decomposition into irreps.

The easiest thing to do is prove that

$$V_{\frac{1}{2}} \otimes V_j \cong V_{j+1/2} \oplus V_{j-1/2} \tag{11.676}$$

If $j = 0$ we interpret $V_{-1/2}$ as the zero vector space. Since a representation is uniquely determined by its character we can prove this by simply computing

$$\begin{aligned} \chi_{1/2}(z)\chi_j(z) &= (z + z^{-1})\frac{z^{2j+1} - z^{-2j-1}}{z - z^{-1}} \\ &= \frac{z^{2j+2} - z^{-2j-2}}{z - z^{-1}} + \frac{z^{2j} - z^{-2j}}{z - z^{-1}} \\ &= \chi_{j+1/2} + \chi_{j-1/2} \end{aligned} \tag{11.677}$$

The general result (15.589) now follows by induction.

Alternatively, we can write:

$$\begin{aligned} \chi_{j_1}(z)\chi_{j_2}(z) &= \left(\frac{z^{2j_1+1} - z^{-2j_1-1}}{z - z^{-1}}\right) \cdot \left(\frac{z^{2j_2+1} - z^{-2j_2-1}}{z - z^{-1}}\right) \\ &= \frac{1}{z - z^{-1}}\left(\frac{z^{2j_1+2j_2+2} - z^{2(j_1-j_2)}}{z - z^{-1}} + \frac{z^{-2j_1-2j_2-2} - z^{-2(j_1-j_2)}}{z - z^{-1}}\right) \end{aligned} \tag{11.678}$$

Now, WLOG assume that $j_1 \geq j_2$. Then we use the identity:

$$\frac{z^{a+2} - z^b}{z - z^{-1}} = z^{b+1}\frac{z^{a-b+2} - 1}{z^2 - 1} = z^{b+1} + z^{b+3} + \cdots + z^{a+1} \tag{11.679}$$

for each of the two terms in the sum above, then realize that the two terms are related by $z \to 1/z$ and we directly obtain:

$$\chi_{j_1}\chi_{j_2} = \chi_{j_1+j_2} + \chi_{j_1+j_2-1} + \cdots + \chi_{|j_1-j_2|} \tag{11.680}$$

It is also instructive to give a proof using the orthogonality of characters. One can check directly by contour integration that

$$-\frac{1}{4\pi \mathrm{i}}\oint \chi_{j_1}(z)\chi_{j_2}(z)\chi_j(z)(z-z^{-1})^2\frac{dz}{z} = \begin{cases} +1 & |j_1-j_2|\leq j\leq j_1+j_2 \qquad \& \qquad 2j=2j_1+2j_2 \mathrm{mod} 2 \\ 0 & \text{else} \end{cases} \tag{11.681}$$

♣Explain this better ♣

**Definition** Choose an ON basis $\psi_{j,m}$ of $V_j$. Let $P_j$ the the orthogonal projector onto the subspace of $V_{j_1}\otimes V_{j_2}$ transforming in the representation $V_j$. Then

$$\langle \psi_{j,m}, P_j(\psi_{j_1,m_1}\otimes\psi_{j_2,m_2})\rangle \tag{11.682}$$

is known as a *Clebsch-Gordon coefficient*. Many notations for these are used in the physics notation. We will just use the notation

$$\langle j,m|j_1,m_1;j_2,m_2\rangle \tag{11.683}$$

**Example**: Let us consider in detail the important case of

$$V_{1/2}\otimes V_{1/2}\cong V_0\oplus V_1 \tag{11.684}$$

The orthogonal projectors are obtained from

$$P^j_{m_L,m_R} = \int_{SU(2)} (D^j_{m_L,m_R}(g))^* T(g) dg \tag{11.685}$$

In particular, as noted above in equation (11.274) the projector to the isotypical component of the trivial representation is always:

$$P = \int_G T(g)dg \tag{11.686}$$

In our example,

$$T(g)|\beta\rangle\otimes|\delta\rangle = g_{\alpha\beta}g_{\gamma\delta}|\alpha\rangle\otimes|\gamma\rangle \tag{11.687}$$

Tradition demands that an ON basis for the fundamental representation be denoted $\{|+\rangle,|-\rangle\}$ with

$$d(z)\cdot|+\rangle = z|+\rangle \qquad\qquad d(z)\cdot = z^{-1}|-\rangle \tag{11.688}$$

We denote the basis vectors $|\alpha\rangle$ with $\alpha,....\in\{+,-\}$. If we look at the projector to the trivial representation then we need to evaluate

$$\int_{SU(2)} g_{\alpha\beta}g_{\gamma\delta}dg \tag{11.689}$$

Now recall that $g_{\alpha\beta} = \epsilon_{\alpha\alpha'}\epsilon_{\beta\beta'}(g^*)^{\alpha'\beta'}$ so we can use the orthogonality relations to say

$$\int_{SU(2)} g_{\alpha\beta}g_{\gamma\delta}dg = \frac{1}{2}\epsilon_{\alpha\gamma}\epsilon_{\beta\delta} \tag{11.690}$$

confirming the direct computation from the exercise (11.110).

It follows that an ON basis for the spin 0 ("singlet") isotypical component of $V_{1/2} \otimes V_{1/2}$ must be proportional to

$$\epsilon_{\alpha\gamma}\epsilon_{\beta\delta}|\alpha\rangle \otimes |\gamma\rangle = \epsilon_{\beta\delta}\left(|+\rangle \otimes |-\rangle - |-\rangle \otimes |+\rangle\right) \tag{11.691}$$

We can normalize the state as:

$$\psi_s := \frac{1}{\sqrt{2}}\left(|+\rangle \otimes |-\rangle - |-\rangle \otimes |+\rangle\right) \tag{11.692}$$

This state is also called a *Bell pair* in quantum information theory. It is the simplest example of a nontrivially entangled state and was used in the famous EPR paper - the first paper on quantum information theory.

The orthogonal complement to the line through $\psi_s$ must be the three-dimensional space of spin one states. One natural basis for this orthogonal complement that diagonalizes the action of the diagonal unitary matrices is

$$\begin{aligned} |1,1\rangle &= |+\rangle \otimes |+\rangle \\ |1,0\rangle &= \frac{1}{\sqrt{2}}\left(|+\rangle \otimes |-\rangle + |-\rangle \otimes |+\rangle\right) \\ |1,-1\rangle &= |-\rangle \otimes |-\rangle \end{aligned} \tag{11.693}$$

A very general statement of selection rules in quantum mechanics for $SU(2)$ symmetry is known in physics as the *Wigner-Eckart theorem*:

Suppose we have a quantum system with some form of $SU(2)$ symmetry. Therefore, the Hilbert space $\mathcal{H}$ is a (possibly infinite dimensional) representation of $SU(2)$. Accordingly, we can choose a basis of states $\psi^{\alpha}_{j,m}$ where $\alpha$ is some independent index (whose range might depend on $j$) such that for fixed $\alpha$ the states t transform like the $f_{j,m}$ basis of $SU(2)$ discussed above. Typically $\alpha$ will label eigenvalues of some other operators that commute with the $SU(2)$ operators (and they might also label degeneracies).

Similarly, in a quantum system it makes sense to decompose the set of all observables into representations of $SU(2)$. Note that the observables are the self-adjoint elements of $\mathrm{End}(\mathcal{H}) \cong \mathcal{H}^{\vee} \otimes \mathcal{H}$, so $\mathrm{End}(\mathcal{H})$ is also a representation of $SU(2)$. Therefore we can also choose a basis for all observables of the form $\mathcal{O}^{I}_{j,m}$ which, for fixed $I$ transform in the spin $j$ representation. [180] Then the most general matrix element we will encounter when computing quantum overlaps is of the form:

$$\langle \psi^{\alpha}_{j,m} | \mathcal{O}^{I}_{j_1,m_1} | \psi^{\beta}_{j_2,m_2} \rangle \tag{11.694}$$

The Wigner-Eckart theorem tells us that the dependence of these matrix elements on $m_1, m_2, m_3$ is completely determined by group theory. It is rather trivially true (using the diagonal subgroup of $SU(2)$) that the matrix element is proportional to $\delta_{m_1+m_2,m}$. The

[180] Since $SU(2)$ representations are pseudoreal and multiplication by complex numbers does not preserve the self-adjoint property it is best to work with all operators $\mathrm{End}(\mathcal{H})$.

nontrivial content of the theorem is that the dependence of the matrix element on $m_1, m_2$ is completely determined by group theory.

To prove the Wigner-Eckert theorem we use the very useful fact that the Wigner D-functions form an algebra:

$$D^{j_1}_{m_1',m_1}(g) D^{j_2}_{m_2',m_2}(g) = \sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{m,m'} \langle j, m' | j_1, m_1'; j_2, m_2' \rangle \langle j, m | j_1, m_1; j_2, m_2 \rangle D^{j}_{m',m}(g) \tag{11.695}$$

Then note that for any $g \in SU(2)$:

$$\begin{aligned} \langle \psi^{\alpha}_{j,m} | \mathcal{O}^{I}_{j_1,m_1} | \psi^{\beta}_{j_2,m_2} \rangle &= \langle \psi^{\alpha}_{j,m} | U(g)^{\dagger} U(g) \mathcal{O}^{I}_{j_1,m_1} | \psi^{\beta}_{j_2,m_2} \rangle \\ &= \sum_{m',m_1',m_2'} (D^{j}_{m',m}(g))^{*} D^{j_1}_{m_1',m_1}(g) D^{j_2}_{m_2',m_2}(g) \langle \psi^{\alpha}_{j,m'} | \mathcal{O}^{I}_{j_1,m_1'} | \psi^{\beta}_{j_2,m_2'} \rangle \end{aligned} \tag{11.696}$$

Now average this equation over $g \in SU(2)$ and use the algebra relation on the D-functions to learn that

$$\begin{aligned} \langle \psi^{\alpha}_{j,m} | \mathcal{O}^{I}_{j_1,m_1} | \psi^{\beta}_{j_2,m_2} \rangle &= \langle j, m | j_1, m_1; j_2, m_2 \rangle \\ &\times \left( \sum_{m',m_1',m_2'} \langle j, m' | j_1, m_1'; j_2, m_2' \rangle \langle \psi^{\alpha}_{j,m'} | \mathcal{O}^{I}_{j_1,m_1'} | \psi^{\beta}_{j_2,m_2'} \rangle \right) \end{aligned} \tag{11.697}$$

The term in large parentheses does not depend on $m, m_1, m_2$. It is therefore a function $F(j_1, j_2, j; \alpha, \beta, I)$ that only depends on $j_1, j_2, j; \alpha, \beta, I$ and not $m, m_1, m_2$. It is sometimes called the "reduced matrix element." It must be computed with other techniques that do not rely on the representation of $SU(2)$ on $\mathcal{H}$.

---

**Exercise** *Recovering Dimension*

As a nice check of the expression on the far RHS in (11.612) show that the limit $u \to 1$ reproduces the dimension of $V_j$.

---

## 11.21 Lie Groups And Lie Algebras And Lie Algebra Representations

### 11.21.1 Some Useful Formulae For Working With Exponentials Of Operators

In this section we will discuss some formulae that are very useful for working with exponentials of matrices (and linear operators). In particular we will derive the Baker-Campbell-Hausdorff formula.

Let us recall that if $A$ is a matrix or an operator then $e^{A}$ is the matrix, or operator, defined by the exponential series. The following three identities are easily shown by direct use of the exponential series:

1.
$$e^{\alpha A} e^{\beta A} = e^{(\alpha+\beta)A} \tag{11.698}$$

2.

$$\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A \tag{11.699}$$

3.

$$e^A e^B e^{-A} = e^{e^A B e^{-A}} \tag{11.700}$$

Now we prove some identities that are not directly obvious from the exponential series:

**Definition**: For $A \in M_n(\kappa)$ we denote by $\mathrm{Ad}(A)$ the linear transformation $M_n(\kappa) \to M_n(\kappa)$ defined by

$$\mathrm{Ad}(A) : B \mapsto [A, B] \tag{11.701}$$

We also denote:

$$(\mathrm{Ad}(A))^m B = \overbrace{[A, [A, \cdots [A, B] \cdots ]}^{m \text{ times}} \tag{11.702}$$

where there are $m$ commutators on the RHS.

First we prove

$$e^A B e^{-A} = e^{\mathrm{Ad}(A)} B \tag{11.703}$$

in other words:

$$\begin{aligned} e^A B e^{-A} &= e^{\mathrm{Ad}(A)} B \\ &= B + \mathrm{Ad}(A)B + \frac{1}{2!}(\mathrm{Ad}(A))^2 B + \cdots \\ &= B + [A, B] + \frac{1}{2!}[A, [A, B]] + \cdots \end{aligned} \tag{11.704}$$

To prove this define $B(t) := e^{tA} B e^{-tA}$. So $B(0) = B$ and $B(1) = e^A B e^{-A}$ is the quantity we want. Now it is easy to derive the differential equation:

$$\frac{d}{dt}B(t) = \mathrm{Ad}(A)B(t) \tag{11.705}$$

so

$$B(t) = e^{t\mathrm{Ad}(A)} B(0) \tag{11.706}$$

Now set $t = 1$.

Combining with (11.700) we now have the somewhat less trivial identity:

$$e^A e^B e^{-A} = e^{e^{\mathrm{Ad}(A)} B} \tag{11.707}$$

All these identities follow from a much more nontrivial formula, known as the Baker-Campbell-Hausdorff formula that expresses the operator $C$ defined by

$$e^A e^B = e^C \tag{11.708}$$

as a power series in $A, B$. We have

$$C = A + B + s(A, B) \tag{11.709}$$

where $s(A,B)$ is an infinite series and every term involves nested commutators.

We will give a complete statement and proof of the BCH formula below. In order to do that we first state the extremely useful

**Lemma** : Let

$$f(z) = \frac{e^z - 1}{z} = 1 + \frac{z}{2!} + \frac{z^2}{3!} + \cdots \tag{11.710}$$

Then

$$(\frac{d}{dt} e^{A(t)}) e^{-A(t)} = -e^{A(t)} \frac{d}{dt} e^{-A(t)} = f(\mathrm{Ad}(A(t))) \cdot \dot{A}(t) \tag{11.711}$$

where $A(t)$ is any differentiable matrix function of $t$.

Note that this is nontrivial because $\dot{A}(t)$ does not commute with $A(t)$ in general! Indeed, from the exponential series you can easily show that

$$\begin{aligned} \frac{d}{dt} e^{A(t)} - \dot{A}(t) e^{A(t)} &= \frac{1}{2}[A(t), \dot{A}(t)] + \cdots \\ \frac{d}{dt} e^{A(t)} - e^{A(t)} \dot{A}(t) &= -\frac{1}{2}[A(t), \dot{A}(t)] + \cdots \end{aligned} \tag{11.712}$$

*Proof*: Introduce a matrix function of two variables and take derivatives wrt $s$:

$$\begin{aligned} B(s,t) &:= e^{sA(t)} \frac{d}{dt} e^{-sA(t)} \\ \frac{\partial B}{\partial s} &= A(t) e^{sA(t)} \frac{d}{dt} e^{-sA(t)} - e^{sA(t)} \frac{d}{dt} \left[ e^{-sA(t)} A(t) \right] \\ &= \mathrm{Ad}(A(t)) B(s,t) - \dot{A}(t) \\ \frac{\partial^j B}{\partial s_j} &= (\mathrm{Ad}(A(t)))^j B(s,t) - (\mathrm{Ad} A(t))^{j-1} \dot{A}(t) \end{aligned} \tag{11.713}$$

$B(0,t) = 0$ therefore again by Taylor:

$$\frac{1}{j!} \frac{\partial^j}{\partial s^j} B(s,t) \mid_{s=0} = -\mathrm{Ad}(A(t))^{j-1} \dot{A}(t) \qquad j \geq 1 \tag{11.714}$$

So

$$e^{sA(t)} \frac{d}{dt} (e^{-sA(t)}) = -\sum_{j=1}^{\infty} \frac{s^j (\mathrm{Ad}(A(t)))^{j-1}}{j!} \dot{A}(t) \tag{11.715}$$

Now set s=1. ♠

Note: you can rewrite this lemma as the statement:

$$\frac{d}{dt} e^{A(t)} = \int_0^1 e^{sA(t)} \dot{A}(t) e^{(1-s)A(t)} ds \tag{11.716}$$

because:

$$\begin{aligned}\frac{d}{dt}e^{A(t)} &= \int_0^1 e^{sA(t)}\dot{A}(t)e^{(1-s)A(t)}ds \\ &= \int_0^1 e^{sAd(A(t))}ds\dot{A}(t)e^{A(t)} \\ &= \left[\left(\frac{e^{Ad(A(t))}-1}{Ad(A(t))}\right)\dot{A}(t)\right]e^{A(t)}\end{aligned} \tag{11.717}$$

**Remark**: Equation (11.716) is an intuitively appealing formula. For a finite product we have:

$$\begin{aligned}\frac{d}{dt}\Big[M_1(t)M_2(t)M_3(t)\cdots M_n(t)\Big] &= (\frac{d}{dt}M_1(t))M_2(t)M_3(t)\cdots \\ &+ (M_1(t))(\frac{d}{dt}M_2(t))M_3(t)\cdots \\ &+ (M_1(t))(M_2(t))(\frac{d}{dt}M_3(t))\cdots \\ &+ \cdots + M_1(t)M_2(t)\cdots(\frac{d}{dt}M_n(t))\end{aligned} \tag{11.718}$$

Now write

$$e^{A(t)} = \prod_{i=1}^{N}[e^{A(t)\Delta s}] \tag{11.719}$$

where $\Delta s = 1/N$. By equation (11.712), we can replace $\frac{d}{dt}e^{\Delta sA(t)}$ by $e^{\Delta sA(t)}\Delta s\dot{A}(t)$ up to order $(\Delta s)^2$. Then write the general term in the sum (11.718) as

$$\left(\prod_{i<s}e^{\Delta sA(t)}\right)e^{\Delta sA(t)}\dot{A}(t)\left(\prod_{s<i}e^{\Delta sA(t)}\right)\Delta s \tag{11.720}$$

up to terms of order $(\Delta s)^2$. Next we sum over these terms and take $N \to \infty$ to get (11.716).

Now we are finally ready to state the main theorem:

**Theorem**: (Baker-Campbell-Hausdorff formula)
Let:

$$g(w) = \frac{\log w}{w-1} = \sum_{j=0}^{\infty}\frac{(1-w)^j}{j+1} = 1 + \frac{1-w}{2} + \frac{(1-w)^2}{3} + \cdots \tag{11.721}$$

be a power series in $w$ about 1. Then when $A, B$ are $n \times n$ matrices with $\parallel A \parallel$, $\parallel B \parallel$ sufficiently small, the matrix $C$ given by the expansion:

$$\boxed{C = B + \int_0^1 g(e^{t\mathrm{Ad}A}e^{\mathrm{Ad}B})(A)dt} \tag{11.722}$$

satisfies $C = \log(e^A e^B)$.

*Proof*:

Introduce the matrix-valued function $C(t)$ via:

$$e^{C(t)} = e^{tA} e^B \tag{11.723}$$

and note that $C(0) = B$, and $C(1)$ is the matrix we want. We derive a differential equation for $C(t)$. By our lemma we have:

$$e^{C(t)} \frac{d}{dt} e^{-C(t)} = -f(\mathrm{Ad}C(t))\dot{C}(t) \tag{11.724}$$

with

$$f(z) = \frac{e^z - 1}{z} \tag{11.725}$$

On the other hand, plugging in the definition (11.723) we compute directly the simple result

$$e^{C(t)} \frac{d}{dt} e^{-C(t)} = e^{tA} \frac{d}{dt} e^{-tA} = -A \tag{11.726}$$

Therefore we get a differential equation:

$$f(\mathrm{Ad}C(t))\dot{C}(t) = A \tag{11.727}$$

Now, $f$ is a power series about 1 so it immediately follows that

$$\dot{C}(t) = f(Ad(C(t)))^{-1} A \tag{11.728}$$

Let us make this more explicit: Using the power series $g(w)$ above with $w = e^z$ note that

$$f(z)g(e^z) = \frac{e^z - 1}{z} \cdot \frac{z}{e^z - 1} = 1 \tag{11.729}$$

regarded as an identity of power series in $z$. Now we can substitute for $z$ any operator $\mathcal{O}$, and use

$$g(e^{\mathcal{O}}) = f(\mathcal{O})^{-1}, \tag{11.730}$$

and therefore we can solve for $\dot{C}$:

$$\begin{aligned} \dot{C}(t) &= f(\mathrm{Ad}(C(t)))^{-1} \cdot A \\ &= g(\exp(\mathrm{Ad}(C(t))) \cdot A \end{aligned} \tag{11.731}$$

where we applied (11.730) with $\mathcal{O} = \mathrm{Ad}(C(t)$. This hardly seems useful, since we still don't know $C(t)$, but now since we have power series we can say

$$e^{\mathcal{O}} = e^{Ad(C(t))} = e^{\mathrm{Ad}(tA)} e^{\mathrm{Ad}(B)} \tag{11.732}$$

To prove (11.732) note that for all $H$ we have:

$$\begin{aligned} e^{\mathrm{Ad}C(t)}H &= e^{C(t)}He^{-C(t)} \\ &= e^{tA}e^{B}He^{-B}e^{-tA} \\ &= e^{\mathrm{Ad}(tA)}e^{\mathrm{Ad}(B)}H \\ \Rightarrow e^{\mathrm{Ad}(C(t))} &= e^{\mathrm{Ad}(tA)}e^{\mathrm{Ad}(B)} \end{aligned} \tag{11.733}$$

Therefore:

$$\dot{C}(t) = g(e^{\mathrm{Ad}(tA)}e^{\mathrm{Ad}(B)}) \cdot A \tag{11.734}$$

Now we integrate equation (11.734)

$$C(t) = C(0) + \int_0^t g(e^{\mathrm{Ad}(sA)}e^{\mathrm{Ad}(B)})Ads \tag{11.735}$$

but $C(0) = B$, so

$$C = C(1) = \log(e^A e^B) = B + \int_0^1 g(e^{\mathrm{Ad}(s\ A)}e^{\mathrm{Ad}(B)})A\ ds \tag{11.736}$$

which is what we wanted to show. ♠.

**Remarks**:

1. To evaluate $g(e^{\mathcal{O}})$ for an operator $\mathcal{O}$ we expand $e^{\mathcal{O}}$ around 0 so $e^{\mathcal{O}} = 1 + \mathcal{O} + \cdots$ and then we expand around 1 to get an expansion of $g(e^{\mathcal{O}})$ around $\mathcal{O} = 0$. A similar remark applies to $g(e^{\mathcal{O}_1}e^{\mathcal{O}_2})$.

2. Explicitly the first few terms are: [181]

$$\boxed{C = A + B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] + \frac{1}{12}[B,[B,A]] + \frac{1}{24}[A,[B,[A,B]]] + \cdots} \tag{11.737}$$

where the next terms are order $\epsilon^5$ if we scale $A, B$ by $\epsilon$.[182]

3. For suitable operators $A, B$ on Hilbert space the BCH formula continues to hold. But the series has a finite radius of convergence: See the exercises below.

---

**Exercise**
Work out the BCH series to order 5 in $A, B$.

---

[181] It is useful to note that $[A,[B,[A,B]]] = -[B,[A,[B,A]]] = B^2A^2 - A^2B^2$

[182] One can find an algorithm for generating the higher order terms in Varadarajan's book on group theory.

**Exercise**

Show that we can also write:

$$C = \log(e^A e^B) = A + \int_0^1 g(e^{-\mathrm{Ad}(s\ B)} e^{-\mathrm{Ad}(A)}) B\ ds \tag{11.738}$$

**Exercise** *All Orders In B, First Order in A*

Write $A = \epsilon$ and consider it to be small. Show that the formula for $C$ given by BCH to all orders in $B$ and first order in $\epsilon$ is

$$\begin{aligned} C &= B + \frac{\mathrm{Ad}B}{e^{\mathrm{Ad}B} - 1}(\epsilon) \\ &= B + \epsilon - \frac{1}{2}[B,\epsilon] + \frac{1}{12}[B,[B,\epsilon]] - \frac{1}{720}[B,[B,[B,[B,\epsilon]]]] + \cdots \end{aligned} \tag{11.739}$$

Note:

$$\begin{aligned} \frac{x}{e^x - 1} &= 1 - \frac{1}{2}x + \frac{1}{12}x^2 - \frac{x^4}{720} + \frac{x^6}{30240} - \frac{x^8}{1209600} + \frac{x^{10}}{47900160} \\ &\quad - \frac{691}{1307674368000}x^{12} + \cdots \\ &\equiv \sum_{n=0}^{\infty} \frac{B_n x^n}{n!} \end{aligned} \tag{11.740}$$

is an important expansion in classical function theory -the numbers $B_n$ are known as the Bernoulli numbers

There are many applications of this formula. One in particle physics is to spontaneous symmetry breaking where the formula above gives the chiral transformation law of the pion field. Here $B = \pi(x)$ is the pion field and $\epsilon$ is the chiral transformation parameter.

**Exercise** *Eigenvalues of* Ad($A$) *For Diagonalizable A*

a.) Show that if $A \in M_n(\kappa)$ is diagonalizable $A \sim \mathrm{Diag}\{\lambda_1, \ldots, \lambda_n\}$ then the eigenvalues of Ad($A$) acting on $M_n(\kappa)$ are $\lambda_i - \lambda_j$.

b.) Using (a) and the previous exercise conclude that the BCH formula has finite radius of convergence.

### 11.21.2 Lie Algebras

We recall two basic definitions - the definition of an *algebra* and of a *Lie algebra*. See Chapter 2, Linear Algebra User's Manual for more discussion and examples. The relation between Lie groups and Lie algebras is covered in much more detail in Chapter 8. But here is a preview.

**Definition** An *algebra* over a field $\kappa$ is a vector space $A$ over $\kappa$ with a notion of multiplication of two vectors

$$A \times A \to A \tag{11.741}$$

denoted:

$$a_1, a_2 \in A \to a_1 \odot a_2 \in A \tag{11.742}$$

which has a ring structure compatible with the scalar multiplication by the field. Concretely, this means we have axioms:

i.) $(a_1 + a_2) \odot a_3 = a_1 \odot a_3 + a_2 \odot a_3$
ii.) $a_1 \odot (a_2 + a_3) = a_1 \odot a_2 + a_1 \odot a_3$
iii.) $\alpha(a_1 \odot a_2) = (\alpha a_1) \odot a_2 = a_1 \odot (\alpha a_2), \qquad \forall \alpha \in \kappa.$

The product $\odot$ might, or might not, be associative. When it is associative, it is called an *associative algebra*.

**Example**: A basic example of an algebra is the vector space of $n \times n$ matrices over a field $\kappa$. The vector addition is simply addition of matrices. The algebra product $\odot$ is matrix multiplication.

**Definition** A *Lie algebra* over a field $\kappa$ is an algebra $A$ over $\kappa$ where the multiplication of vectors $a_1, a_2 \in A$, satisfies in addition the two conditions:

1. $\forall a_1, a_2 \in A$:
$$a_2 \odot a_1 = -a_1 \odot a_2 \tag{11.743}$$

2. $\forall a_1, a_2, a_3 \in A$:
$$((a_1 \odot a_2) \odot a_3) + ((a_3 \odot a_1) \odot a_2) + ((a_2 \odot a_3) \odot a_1) = 0 \tag{11.744}$$
This is known as the *Jacobi relation*.

Now, tradition demands that the product on a Lie algebra be denoted not as $a_1 \odot a_2$ but rather as $[a_1, a_2]$ where it is usually referred to as the *bracket*. So then the two defining conditions (11.743) and (11.744) are written as:

1. $\forall a_1, a_2 \in A$:
$$[a_2, a_1] = -[a_1, a_2] \tag{11.745}$$

2. $\forall a_1, a_2, a_3 \in A$:

$$[[a_1, a_2], a_3] + [[a_3, a_1], a_2] + [[a_2, a_3], a_1] = 0 \tag{11.746}$$

**Remark**: Note that if we consider the Lie algebra product as the algebra product on a vector space, then the algebra is <u>non-associative</u>. If we have a Lie algebra product and use it to define an algebra product: $a_1 \odot a_2 := [a_1, a_2]$ then

$$\begin{aligned}(a_1 \odot a_2) \odot a_3 - a_1 \odot (a_2 \odot a_3) &= [[a_1, a_2], a_3] - [a_1, [a_2, a_3]] \\ &= -[[a_3, a_1], a_2]]\end{aligned} \tag{11.747}$$

by the Jacobi identity.

**Example 1**: Whenever we have an <u>associative</u> algebra $A$, we can automatically turn it into a Lie algebra by defining the Lie product as a commutator:

$$[a_1, a_2] := a_1 \odot a_2 - a_2 \odot a_1 \tag{11.748}$$

However, it should be stressed that not all Lie algebras arise in this way. A good example would be first order smooth differential operators on, say, $\mathbb{R}^n$, (or more generally, smooth vector fields on a manifold). If

$$V = v^\mu(x) \frac{\partial}{\partial x^\mu} \tag{11.749}$$

then the product of the differential operators - which makes sense as a second order operator - is not a vector field, but we can define

$$[V_1, V_2] = \left( v_1^\mu \frac{\partial}{\partial x^\mu} v_2^\nu - v_2^\mu \frac{\partial}{\partial x^\mu} v_1^\nu \right) \frac{\partial}{\partial x^\nu} \tag{11.750}$$

**Example 2**: An important example arises by applying the general remark above to the algebra $M_n(\kappa)$. In this case the Lie algebra is often denoted $\mathfrak{gl}(n, \kappa)$. Now $M_n(\kappa)$ and $\mathfrak{gl}(n, \kappa)$ are precisely the same as sets, and are precisely the same as vector spaces. The notation simply puts a different stress on the algebraic structure which the author is considering.

**Example 3**: If $\mathfrak{g} \subset \mathfrak{gl}(n, \kappa)$ is a vector subspace of matrices that is closed under matrix commutation, then $\mathfrak{g}$ is a Lie subalgebra of $\mathfrak{gl}(n, \kappa)$. For example, there is a vector subspace of $\mathfrak{gl}(n, \kappa)$ consisting of <u>traceless</u> matrices. The property of being traceless is preserved by commutator. Therefore this subspace is a sub-Lie algebra. It is denoted as $\mathfrak{sl}(n, \kappa)$. In the classification of semi-simple Lie algebras over $\mathbb{C}$ these Lie algebras have special names: $A_n := \mathfrak{sl}(n+1, \kappa)$.

**Example 4**: Consider $\mathfrak{so}(n, \kappa) \subset M_n(\kappa)$ defined to be the vector subspace of $n \times n$ <u>antisymmetric</u> matrices. The matrix product of anti-symmetric matrices is not anti-symmetric, but the matrix commutator is. In the classification of semi-simple Lie algebras over $\mathbb{C}$ these Lie algebras have special names, and for good reasons the even and odd cases are considered as separate families. They are denoted $B_n = \mathfrak{so}(2n+1)$ and $D_n = \mathfrak{so}(2n)$.

**Example 5**: $\mathfrak{u}(n) \subset M_n(\mathbb{C})$ is defined to be the vector subspace of $n \times n$ antihermitian matrices. The matrix product of anti-hermitian matrices is not anti-hermitian, but the matrix commutator is. Note carefully that $\mathfrak{u}(n)$ is a real vector space and is not a complex vector space! After all, if $A^\dagger = -A$ then $(\mathrm{i}A)^\dagger = \mathrm{i}A$ is not anti-hermitian. We can further define a sub-Lie algebra $\mathfrak{su}(n) \subset \mathfrak{u}(n)$ of anti-hermitian matrices that are, in a addition, traceless.

**Example 6**: Finally $\mathfrak{sp}(2n,\kappa) \subset M_n(\kappa)$ is the Lie subalgebra such that $(Ja)^{tr} = Ja$ where

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{11.751}$$

**Example 7**: The intersection of two Lie algebras is a Lie algebra and a particularly important example is $\mathfrak{usp}(2n) = \mathfrak{su}(2n) \cap \mathfrak{sp}(2n,\mathbb{C}) \subset M_{2n}(\mathbb{C})$. ♣say why ♣

Let $\mathfrak{g} \subset \mathfrak{gl}(n,\kappa)$ be a sub-Lie algebra with $\kappa = \mathbb{Q}, \mathbb{R}, \mathbb{C}$. The BCH formula gives one way to understand a relation between Lie algebras and Lie groups: Provided the BCH series converges it follows that if $A, B \in \mathfrak{g}$ then

$$e^A e^B = e^C \tag{11.752}$$

with $C \in \mathfrak{g}$. Thus, up to convergence issues, the invertible operators $e^A$ with $A \in \mathfrak{g}$ will close to form a group. The series will converge for $A, B$ in an open neighborhood of the origin. So, taking the closure (under group multiplication) of the set of matrices $\exp[A]$ with $A \in \mathfrak{g}$ will give the Lie group.

One can show that for compact connected Lie groups the exponential map

$$\exp : \mathfrak{g} \to G \tag{11.753}$$

is indeed surjective, although it will not be injective. A good example is the case $U(1)$. Here the Lie algebra is $\mathfrak{g} = \mathrm{i}\mathbb{R}$ (and the commutator is zero). Then all of $2\pi\mathrm{i}\mathbb{Z}$ is in the kernel of the exponential map.

Conversely, from a matrix Lie group one can recover the Lie algebra by considering the general one-parameter subgroups $g(t)$ with $g(0) = 1$ and computing $g^{-1}(t)\frac{d}{dt}g(t)$ at $t = 0$. We will elaborate on this idea by first closing a gap in our discussion at the very beginning of the course and prove that the classical Lie groups are manifolds as follows:

**Exercise** *Structure Constants*

If $\mathfrak{g}$ is a Lie algebra over $\kappa$ then we can choose a basis $T^i$ for $\mathfrak{g}$ and necessarily we have an expansion

$$[T^i, T^j] = \sum_k f_k^{ij} T^k \qquad f_k^{ij} \in \kappa \tag{11.754}$$

from which one can construct all commutators. The constants $f_k^{ij}$ are known as *structure constants*.

a.) Show that

$$f_k^{ij} = -f_k^{ji} \tag{11.755}$$

$$f_\ell^{ij} f_m^{\ell k} + f_\ell^{jk} f_m^{\ell i} + f_\ell^{ki} f_m^{\ell j} = 0 \tag{11.756}$$

b.) Conversely, show that given tensors $f_k^{ij} \in \kappa$ satisfying equations (11.755) and (11.756) one can define a Lie algebra over $\kappa$.

**Remark**: Sometimes Lie algebras are presented by giving a list of structure constants. If someone tries to sell you a Lie algebra by giving you a list of structure constants don't buy it until you have checked equations (11.755) and (11.756).

---

### 11.21.3 The Classical Matrix Groups Are Lie Groups

This section assumes some knowledge of differential geometry. Some readers might wish to skip it.

First we recall the pre-image theorem. A map $f : M_1 \to M_2$ between two manifolds $M_1$ and $M_2$ is said to be a *submersion* if at every $p \in M_1$ the map of tangent spaces:

$$df : T_p M_1 \to T_{f(p)} M_2 \tag{11.757}$$

is surjective. In this case $f(p)$ is said to be a *regular value*. Using results from the calculus of many variables one can show that if $f$ is a submersion at $p$ then there are local coordinates so that in a neighborhood of $p$ it has the form:

$$f : (x^1, \dots, x^{n_1}) \mapsto (x^1, \dots, x^{n_2}) \tag{11.758}$$

That is, locally, in suitable coordinate systems $f$ is literally the map $f : \mathbb{R}^{n_1} \to \mathbb{R}^{n_2}$ keeping only the first $n_2$ coordinates. Then, every point in the target $\mathbb{R}^{n_2}$ is a regular value of $f$ and the inverse image of a regular value is

$$f^{-1}(c^1, \dots, c^{n_2}) = \{x \in \mathbb{R}^{n_1} | x^1 = c^1, \dots, x^{n_2} = c^{n_2}\} \cong \mathbb{R}^{n_1 - n_2} \tag{11.759}$$

is a submanifold of dimension $n_1 - n_2$.

Therefore we have

**Theorem**: [Preimage Theorem] If $f : M_1 \to M_2$ and $q \in M_2$ is a regular value in the image of $f$ then the preimage $f^{-1}(q)$ is a submanifold of $M_1$ of dimension $\dim M_1 - \dim M_2$. That is, the preimage $f^{-1}(q)$ is a submanifold of $M_1$ of codimension dim$M_2$. The tangent space to $f^{-1}(q)$ at any point $p$ is $\ker(df_p)$.

Now we can apply this idea to describe subsets defined by equations. If $f : M \to \mathbb{R}^\ell$ and dim$M \geq \ell$ then for $\vec{c} \in \mathbb{R}^\ell$ the sets

$$M_{\vec{c}} := f^{-1}(\vec{c}) = \cap_{i=1}^{\ell}\{p \in M | f^i(p) = c^i\} \tag{11.760}$$

are called level sets. If $\vec{c}$ is a regular value then the level set is a submanifold of $M$ of codimension $\ell$. Note that for each $i$, $f^i : M \to \mathbb{R}$ so $df^i : T_pM \to T_{c^i}\mathbb{R} \cong \mathbb{R}$, and hence $df^i$ is a linear functional on $T_pM$. The regularity condition is the condition that these linear functionals are all linearly independent. We say that the functions $f^i$ are *independent*.

We can apply these ideas to give easy proofs that the classical matrix groups are in fact Lie groups. First we prove they are all manifolds:

**Example 1**: $GL(n,\mathbb{R}) \subset \mathbb{R}^{n^2}$ and $GL(n,\mathbb{C}) \subset \mathbb{C}^{n^2} \cong \mathbb{R}^{2n^2}$ are both manifolds. The coordinates are the matrix elements. The inverse image $\det^{-1}(0)$ is a closed subset of $M_n(\kappa)$, with $\kappa = \mathbb{R}, \mathbb{C}$ and hence any invertible matrix has an open neighborhood of invertible matrices. The matrix elements thus serve as global coordinates. Unless we insist that our coordinate patches are diffeomorphic to $\mathbb{R}^{n^2}$ there is no need to use more than one coordinate patch. The tangent space at any point $A \in GL(n,\kappa)$ is isomorphic to the vector space $M_n(\kappa)$ of $n \times n$ matrices over $\kappa$.

**Example 2**: Now, for $SL(n,\kappa)$ consider $f : M_n(\kappa) \to \kappa$ defined by $f(A) := \det A - 1$. We claim that $0 \in \kappa$ is a regular value of $f$. Indeed, if $A$ is invertible then for any

$$M \in T_A M_n(\kappa) \cong M_n(\kappa) \tag{11.761}$$

we have

$$df_A(M) = \det A \mathrm{Tr}(A^{-1}M) \tag{11.762}$$

This is usually written as the (very useful) identity [183]

$$\delta \log\det A = \mathrm{Tr}\left(A^{-1}\delta A\right) \tag{11.763}$$

for $A$ invertible. When $A$ is invertible the kernel of $df_A$ is the linear subspace of $n \times n$ matrices $M$ such that $A^{-1}M$ is traceless, which is linearly equivalent to the linear subspace of traceless matrices, and therefore has dimension (over $\kappa$) equal to $n - 1$. Therefore the rank of $df_A$ is 1, and $f = \det$ is a submersion. So the inverse image is a manifold.

**Example 3**: For $O(n;\kappa)$ we define $f : M_n(\kappa) \to S_n(\kappa)$ where $S_n(\kappa) \cong \kappa^{\frac{1}{2}n(n+1)}$ is the vector space over $\kappa$ of $n \times n$ symmetric matrices. We take $f$ to be

$$f(A) = AA^{tr} - 1 \tag{11.764}$$

[183] For a proof see the Linear Algebra User's Manual, ch. 3.

Then $O(n) = f^{-1}(0)$. We aim to show it is a manifold. Note that $df_A$ is a linear operator $M_n(\kappa) \to S_n(\kappa)$. It is just

$$df_A(M) = MA^{tr} + AM^{tr} \tag{11.765}$$

Therefore $\ker(df_A)$ is the linear subspace of $M_n(\kappa)$ of matrices such that $MA^{tr}$ is anti-symmetric. When $A$ is invertible this subspace is isomorphic to the linear subspace of anti-symmetric matrices and hence has dimension $\frac{1}{2}n(n-1)$. It follows that 0 is a regular value of $f$ and $O(n,\kappa)$ is a manifold.

**Example 4**: For $Sp(2n;\kappa)$ we define $f : M_{2n}(\kappa) \to A_{2n}(\kappa)$ where $A_{2n}(\kappa)$ is the set of $(2n) \times (2n)$ matrices over $\kappa$ such that $(Jm)$ is antisymmetric. This is isomorphic to the vector space over $\kappa$ of dimension $\frac{1}{2}(2n)(2n-1) = n(2n-1)$. Now we take $f$ to be

$$f(A) = AJA^{tr}J^{tr} - 1 \tag{11.766}$$

so that $Sp(2n;\kappa) = f^{-1}(0)$. Again we claim that 0 is a regular value of $f$. Now $df_A$ is the linear operator $M_n(\kappa) \to A_{2n}(\kappa)$. It is just

$$df_A(M) = MJA^{tr}J^{tr} + AJM^{tr}J^{tr} \tag{11.767}$$

Therefore $\ker(df_A)$ is the linear subspace of $M_n(\kappa)$ of matrices such that $MJA^{tr}$ is symmetric. When $A$ is invertible this subspace is isomorphic to the linear subspace of symmetric matrices and hence has dimension $\frac{1}{2}2n(2n+1)$, which is complementary to the dimension of the image $\frac{1}{2}(2n)(2n-1)$ and hence $df_A$ is surjective. It follows that 0 is a regular value of $f$ and $Sp(2n;\kappa)$ is a manifold.

**Example 5**: Finally, for $U(n)$ consider $f : M_n(\mathbb{C}) \to \mathcal{H}_n$ where $\mathcal{H}_n$ is the *real* vector space of $n \times n$ Hermitian matrices in $M_n(\mathbb{C})$. This has real dimension $n + 2 \times \frac{1}{2}n(n-1) = n^2$. We now take $f(A) = AA^\dagger - 1$. Then

$$df_A(M) = MA^\dagger + AM^\dagger \tag{11.768}$$

When $A$ is invertible the kernel is the subspace of $M_n(\mathbb{C})$ of matrices such that $MA^\dagger$ is anti-hermitian. This is again a real vector space of real dimension $n + 2 \times \frac{1}{2}n(n-1) = n^2$. Since $M_n(\mathbb{C})$ is a real vector space of real dimension $2n^2$ it follows that $df_A$ is surjective and hence 0 is a regular value of $f$. Therefore $U(n)$ is a manifold.

**Example 6**: It is useful to combine the previous two examples and define the Lie group:

$$USp(2n) := U(2n) \cap Sp(2n, \mathbb{C}) \tag{11.769}$$

Now we have a map: $f : M_{2n}(\mathbb{C}) \to A_{2n}(\mathbb{C}) \oplus \mathcal{H}_{2n}$ defined by taking the direct sum. Again, one must check that the *real linear* map $df_A$ at an invertible matrix in the preimage of 0 has a kernel of the correct dimension. [184]

[184] We have not covered quaternions yet, but a superior viewpoint is to view $USp(2n)$ as the group of $n \times n$ unitary matrices over the quaternions. In this viewpoint we should define $f : M_n(\mathbb{H}) \to \mathcal{H}_n(\mathbb{H})$

All the examples above are submanifolds of $GL(n,\kappa)$. The group operation of multiplication is polynomial in the matrix elements and hence certainly a $\mathcal{C}^\infty$ function. The group operation of inversion is a rational function of the matrix elements and is also $\mathcal{C}^\infty$ on $GL(n,\kappa)$. It follows from the exercise of Section §?? that the group operations of multiplication and inversion are $\mathcal{C}^\infty$ maps in all the cases of the matrix subgroups. This concludes the argument that the above examples are all Lie groups.

In general, the *Lie algebra* of a Lie group $G$ is defined, as a vector space, to be the tangent space at the identity:

$$Lie(G) := T_1 G \tag{11.770}$$

As we will demonstrate, it is in fact a Lie algebra.

Our proof above that the classical matrix groups are manifolds also leads nicely to an immediate computation of the Lie algebras of these groups.

♣For consistency, use small gothic font for Lie algebras ♣

1. $GL(n,\kappa)$:

$$T_1 GL(n,\kappa) \cong M_n(\kappa) = \mathfrak{gl}(n,\kappa) \tag{11.771}$$

2. $SL(n,\kappa)$:

$$T_1 SL(n,\kappa) \cong \{M \in M_n(\kappa) | \mathrm{Tr}(M) = 0\} = \mathfrak{sl}(n,\kappa) \tag{11.772}$$

3. $O(n,\kappa)$:

$$\mathrm{so}(n;\kappa) := T_1 O(n,\kappa) \cong \{M \in M_n(\kappa) | M^{tr} = -M\} = \mathfrak{o}(n,\kappa) = \mathfrak{so}(n,\kappa) \tag{11.773}$$

4. $Sp(2n,\kappa)$:

$$T_1 Sp(2n,\kappa) \cong \{M \in M_{2n}(\kappa) | (MJ)^{tr} = +MJ\} = \mathfrak{sp}(2n,\kappa) \tag{11.774}$$

5. $U(n)$:

$$T_1 U(n) \cong \{M \in M_n(\mathbb{C}) | M^\dagger = -M\} = \mathfrak{u}(n) \tag{11.775}$$

6. $SU(n)$:

$$T_1 SU(n) \cong \{M \in M_n(\mathbb{C}) | M^\dagger = -M \quad \& \quad \mathrm{Tr}(M) = 0\} = \mathfrak{su}(n) \tag{11.776}$$

7. $USp(2n)$:

$$\begin{aligned} T_1 USp(2n) &\cong \{M \in M_{2n}(\mathbb{C}) | M^\dagger = -M \quad \& \quad (MJ)^{tr} = MJ\} \\ &\cong \{M \in M_n(\mathbb{H}) | M^\dagger = -M\} := \mathfrak{usp}(2n) \end{aligned} \tag{11.777}$$

---

where $\mathcal{H}_n(\mathbb{H})$ is the space of $n \times n$ quaternionic Hermitian matrices. The above arguments work in the same way: $\dim_{\mathbb{R}} M_n(\mathbb{H}) = 4n^2$, while $\dim_{\mathbb{R}} \mathcal{H}_n(\mathbb{H}) = n + 4\frac{1}{2}n(n-1) = 2n^2 - n$. As before, the kernel of $df_A$, for $A$ invertible is the space of $n \times n$ quaternionic-antihermitian matrices. This has real dimension $3n + 4\frac{1}{2}n(n-1) = 2n^2 + n$. (The $3n$ is there because one can have an arbitrary imaginary quaternion on the diagonal.) Therefore 0 is a regular value, and $USp(2n)$ is a manifold. Many authors denote this group simply as $Sp(n)$.

We now recognize that the above vector spaces of matrices are in fact the Lie algebras we introduced earlier. We can get back the groups by exponentiation. It is a good exercise to check, from the defining relations of $T_1G$ above that the exponentiated matrix indeed satisfies the defining relations of the group. Thus, for example, one should check that if $M$ is anti-Hermitian, i.e. if $M^\dagger = -M$ then $\exp(M)$ is unitary.

We can argue more generally that they must be Lie algebras as follows: Given a matrix $V \in T_1G \subset M_N(\kappa)$ we can form the family of group elements

$$g_V(t) = \exp[tV] := \sum_{n=0}^{\infty} \frac{(tV)^n}{n!} \tag{11.778}$$

Note that for $t \in \mathbb{R}$ these elements form a subgroup of $G$:

$$g_V(t_1)g_V(t_2) = g_V(t_1 + t_2) \tag{11.779}$$

In general, in differential geometry, there is an exponential map from the tangent space $T_pM$ of a manifold $M$ to a neighborhood of $p$ in $M$. It is not as explicit as an exponential series of a matrix, but involves using the vector to define a differential equation. An important property of the tangent spaces $T_1G$ for the various groups above is that:

*If $V_1, V_2 \in T_1G$ then the matrix commutator $[V_1, V_2]$ is also in $T_1G$.*

This can be verified by directly checking each case. For example, in the case of $so(n, \kappa)$, if $V_1, V_2$ are antisymmetric matrices over $\kappa$ then neither $V_1V_2$, nor $V_2V_1$ is antisymmetric, but $[V_1, V_2]$ is. The reader should check the other cases in this way. Nevertheless, this fact also follows from more general principles, and that is important because as we will see not every Lie group is a classical matrix group. In fact, it is not true that every finite-dimensional Lie group is a subgroup of $GL(N, \mathbb{R})$ for some $N$. [185] Given $V_1, V_2$ we can consider the path through $g = 1$ at $t = 0$ given by the group commutator:

$$\lambda(t) = [g_{V_1}(\sqrt{t}), g_{V_2}(\sqrt{t})] \tag{11.780}$$

Now, usin the BCH formula one can show that for $t_1, t_2$ small we have

$$g_{V_1}(t_1)g_{V_2}(t_2) = \exp[t_1V_1 + t_2V_2 + \frac{1}{2}t_1t_2[V_1, V_2] + \mathcal{O}(t_1^a t_2^b)] \tag{11.781}$$

where the higher order terms have $a + b > 2$, and therefore the tangent vector to the path through $\lambda(t)$ is the matrix commutator.

Therefore, just based on group theory and manifold theory, one can deduce that the tangent space at the identity $\mathfrak{g} = T_1G$ is indeed a Lie algebra.

One of the main theorems about the relation of Lie groups and Lie algebras is the following:

**Theorem**:

[185] A counterexample is the metaplectic group, a group which arises as a central extension of the symplectic group when one tries to implement symplectic transformations on a the quantum mechanics of a system of free particles.

a.) Every finite dimensional Lie algebra $\mathfrak{g}$ over $\kappa = \mathbb{R}$ arises from a unique (up to isomorphism) connected and simply connected Lie group $G$.

b.) Under this correspondence, Lie group homomorphisms $f : G_1 \to G_2$ are in $1-1$ correspondence with Lie algebra homomorphisms $\mu : T_1 G_1 \to T_1 G_2$.

For more about this see Chapter 8. The best statement makes use of the language of categories. What is described here is an equivalence of categories. See Section **** below.

**Examples**:

1. $\mathfrak{su}(2)$ has a standard basis
$$T^a := -\frac{\mathrm{i}}{2}\sigma^a \tag{11.782}$$
with structure constants
$$[T^a, T^b] = \epsilon^{abc} T^c \tag{11.783}$$
Note that any traceless anti-Hermitian matrix can be diagonalized: For any $A \in \mathfrak{su}(2)$ there is a $u \in SU(2)$ with
$$u^{-1} A u = \mathrm{i}\lambda\sigma^3 \tag{11.784}$$
for some $\lambda \in \mathbb{R}$. Note that it follows that the exponentiation of any one-parameter subgroup is subgroup of $SU(2)$ isomorphic to $U(1)$. In particular, it is compact. Finally, as we have seen, every $SU(2)$ group element can be written as
$$g = \cos\chi + \mathrm{i}\sin\chi \hat{n}\cdot\vec{\sigma} = \exp[\mathrm{i}\chi\hat{n}\cdot\vec{\sigma}] \tag{11.785}$$
where $\hat{n} \in S^2 \subset \mathbb{R}^3$. So the exponential map is surjective.

2. $\mathfrak{sl}(2, \mathbb{R})$ has a standard basis:
$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad h = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad f = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \tag{11.786}$$
You should check the structure constants, taking careful note of signs:
$$[h, e] = -2e \qquad [e, f] = h \qquad [h, f] = +2f \tag{11.787}$$
Note that $\mathfrak{sl}(2, \mathbb{R})$ is qualitatively different: The elements $e, f \in \mathfrak{sl}(2, \mathbb{R})$ have non-trivial Jordan form and cannot be diagonalized (even within $GL(2, \mathbb{C})$) So this Lie algebra is inequivalent to $\mathfrak{su}(2)$. It is not hard to show that if $A \in \mathfrak{sl}(2, \mathbb{R})$ then it is conjugate (via $SL(2, \mathbb{R})$ conjugation) to one of three distinct forms
$$SAS^{-1} = xe = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \qquad x \in \mathbb{R} \tag{11.788}$$
$$SAS^{-1} = xh = \begin{pmatrix} -x & 0 \\ 0 & x \end{pmatrix} \qquad x \in \mathbb{R} \tag{11.789}$$

$$SAS^{-1} = x(e-f) = \begin{pmatrix} 0 & x \\ -x & 0 \end{pmatrix} \qquad x \in \mathbb{R} \tag{11.790}$$

and hence there are three distinct maximal Abelian subgroups (up to conjugation):

$$\exp[xe] = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \qquad x \in \mathbb{R} \tag{11.791}$$

$$\exp[xh] = \begin{pmatrix} e^{-x} & 0 \\ 0 & e^{x} \end{pmatrix} \qquad x \in \mathbb{R} \tag{11.792}$$

$$\exp[x(e+f)] = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix} \qquad x \in \mathbb{R} \tag{11.793}$$

Note that the last subgroup is compact and is just a copy of $SO(2,\mathbb{R})$ and $x \sim x+2\pi$ parametrize the same group element.

It is easy to show these subgroups are not conjugate by considering the trace - see the exercise below. It is now easy to see that the exponential map cannot be surjective onto $SL(2,\mathbb{R})$. Consider the group elements

$$\begin{pmatrix} -1 & x \\ 0 & -1 \end{pmatrix} \qquad x \neq 0 \tag{11.794}$$

If this were of the form $\exp[A]$ then it would be in the one-parameter group $\exp[tA]$. But the trace is $-2$ and the only way one of the above one-parameter groups can have trace $= -2$ is to take (11.793) with $x = \pi(2n+1)$ with $n \in \mathbb{Z}$. Nevertheless, via the Gram-Schmidt procedure (see chapter 2) one can prove the so-called *KAN decomposition*: Every $g \in SL(2,\mathbb{R})$ can be uniquely written in the form

$$g = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \cdot \begin{pmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{pmatrix} \cdot \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \tag{11.795}$$

with $\lambda > 0$ and $x \in \mathbb{R}$. So the products of the one-parameter groups generate all group elements.

**Remark**: *The group manifold $SL(2,\mathbb{R})$ as anti-deSitter space:* Every $2 \times 2$ real matrix can plainly be written as:

$$g = \begin{pmatrix} T_1 + X_1 & X_2 + T_2 \\ X_2 - T_2 & T_1 - X_1 \end{pmatrix} \tag{11.796}$$

for real variables $T_1, T_2, X_1, X_2$. Restricting to matrices with $\det g = 1$ gives the hyperboloid in $\mathbb{R}^{2,2}$:

$$-T_1^2 - T_2^2 + X_1^2 + X_2^2 = -1 \tag{11.797}$$

Thus one picture of $SL(2,\mathbb{R})$ can be given as a hyperboloid in $\mathbb{R}^{2,2}$. The induced metric has signature $(-^1,+^2)$. Indeed, a global set of coordinates is:

$$\begin{aligned} T_1 &= \cosh\rho\cos t \\ T_2 &= \cosh\rho\sin t \\ X_1 &= \sinh\rho\cos\phi \\ X_2 &= \sinh\rho\sin\phi \end{aligned} \tag{11.798}$$

These coordinates smoothly cover the manifold once for $0 \le \rho < \infty$, $t \sim t+2\pi$, $\phi \sim \phi+2\pi$. Substituting into

$$ds^2 = -(dT_1)^2 - (dT_2)^2 + (dX_1)^2 + (dX_2)^2 \tag{11.799}$$

the induced metric becomes

$$ds^2 = -\cosh^2\rho dt^2 + \sinh^2\rho d\phi^2 + d\rho^2. \tag{11.800}$$

This is one form of the *anti-deSitter metric* of constant curvature $-1$. Note that the time coordinate $t$ is periodic. What is usually meant by anti-deSitter space is the universal cover of $SL(2,\mathbb{R})$.

---

**Exercise** *Due Diligence*
Show that if matrices $a_1, a_2$ satisfy $(Ja)^{tr} = Ja$ then $[a_1, a_2]$ has the same property.

---

---

**Exercise** *Group Commutators And Lie Algebra Commutators*
a.) Use the BCH theorem to show that if

$$g_1 = e^{t_1 A_1}, \quad g_2 = e^{t_2 A_2} \tag{11.801}$$

the <u>*group commutator*</u>, $g_1 g_2 g_1^{-1} g_2^{-1}$ corresponds to the Lie algebra commutator:

$$g_1 \cdot g_2 \cdot g_1^{-1} \cdot g_2^{-1} = 1 + t_1 t_2 [A_1, A_2] + \mathcal{O}(t_1^2, t_2^2) \tag{11.802}$$

b.) We say a Lie algebra is "Abelian" if $[A_1, A_2] = 0$ for all $A_1, A_2 \in \mathfrak{g}$. That that such a Lie algebra exponentiates to form an Abelian group.

c.) If $A_i = \frac{d}{dt}|_0 g_i(t)$ then $[A_1, A_2]$ is the Lie algebra element associated to the curve

$$g_{12}(t) = g_1(\sqrt{t}) \cdot g_2(\sqrt{t}) \cdot g_1^{-1}(\sqrt{t}) \cdot g_2^{-1}(\sqrt{t}) \tag{11.803}$$

---

---

**Exercise** *Inequivalent One-Parameter Subgroups Of* $SL(2,\mathbb{R})$

a.) Show that the three maximal Abelian subgroups (11.791), (11.792), (11.793) are non-conjugate. [186]

b.) Consider the one-parameter subgroup $\exp[xf]$. Is it conjugate to one of the above three?

---

### 11.21.4 Representations Of Lie Algebras

A *representation of a Lie algebra* is a linear map

$$\rho : \mathfrak{g} \to \mathrm{End}(V) \tag{11.804}$$

for some vector space $V$ (again sometimes called the *carrier space*) such that

$$[\rho(x), \rho(y)] = \rho([x, y]) \tag{11.805}$$

for all $x, y \in \mathfrak{g}$.

Note that any Lie algebra $\mathfrak{g}$ has a canonical representation with $V = \mathfrak{g}$ and

$$\dot{\rho}(x) : y \mapsto \mathrm{Ad}(x)(y) = [x, y] \tag{11.806}$$

This follows because, as one easily checks, the equation $[\dot{\rho}(x), \dot{\rho}(y)] = \dot{\rho}([x, y])$ is equivalent to the Jacobi identity. This representation is known as the *adjoint representation.*

Given a representation $\dot{\rho}$ of a Lie algebra $\mathfrak{g}$ we get a representation of a corresponding (connected) Lie group $G$ so that $\mathfrak{g} = T_1 G$. Recall a *representation of a Lie group* is a group homomorphism

$$\rho : G \to GL(V) \tag{11.807}$$

We do this by setting [187]

$$\rho(e^x) := e^{\dot{\rho}(x)} \tag{11.808}$$

For $\mathfrak{g} = \mathfrak{gl}(n, \kappa)$ the group representation associated to the adjoint representation has carrier space $\mathfrak{g}$ and the group action is:

$$\rho(g)(x) = gxg^{-1} \tag{11.809}$$

Conversely, given a representation $\rho$ of a Lie group we can define a representation of the Lie algebra by

$$\dot{\rho}(X) = \frac{d}{dt}|_0 \rho(e^{tX}) \tag{11.810}$$

**Remarks**:

---

[186] *Answer*: Consider the traces of these matrices. They are $= 2$, $\geq 2$, and $\leq 2$, respectively.

[187] We are assuming there is a corresponding Lie group. Some Lie algebras can, in fact, not be exponentiated to form Lie groups.

1. *Adjoint Representation Of* $\mathfrak{su}(2)$ *And Rotations.* We have already seen this in a slightly different form when we defined the natural double-cover homomorphism $\pi : SU(2) \to SO(3)$. We can identify $\mathfrak{su}(2) \cong \mathbb{R}^3$ as a vector space by identifying $\vec{x} \in \mathbb{R}^3$ with
$$M_{\vec{x}} := \mathrm{i}\vec{x} \cdot \vec{\sigma} \tag{11.811}$$
The adjoint representation is $\rho_{adj}(u) \cdot M = uMu^{-1}$, so $\rho_{adj}$ acts linearly on $\vec{x} \in \mathbb{R}^3$. Note that $M_{\vec{x}}^2 = -\vec{x}^2 1_{2\times 2}$, so this linear action preserves the norm and therefore $\rho_{adj}(u) \in O(3)$. Since
$$\mathrm{tr}\,(M_{\vec{x}_1} M_{\vec{x}_2} M_{\vec{x}_3}) = 2\mathrm{i}\vec{x}_1 \cdot (\vec{x}_2 \times \vec{x}_3) \tag{11.812}$$
is also preserved $\rho_{adj}(u) \in SO(3)$.

2. *Adjoint Representation Of* $\mathfrak{sl}(2,\mathbb{R})$ *And Lorentz Transformations* Every element of $\mathfrak{sl}(2,\mathbb{R})$ can be written as
$$M_{x^\mu} = \begin{pmatrix} x & y-t \\ y+t & -x \end{pmatrix} = (y-t)e + xh + (y+t)f \tag{11.813}$$
Note $g \in SL(2,\mathbb{R})$ acts by $\rho_{adj}(g) \cdot M = gMg^{-1}$. But now
$$M_{x^\mu}^2 = (-t^2 + x^2 + y^2) 1_{2\times 2} \tag{11.814}$$
So, in a manner similar to the previous example $SL(2,\mathbb{R})$ double covers the connected component of the $2+1$ dimensional Lorentz group $SO_0(1,2)$.

3. $\mathfrak{sl}(2,\mathbb{C})$ *And Lorentz Transformations* For completeness we note that $2\times 2$ Hermitian matrices can be identified with $3+1$ dimensional Minkowski space $\mathbb{M}^{1,3}$ via
$$M_{x^\mu} = x^0 1_{2\times 2} + \vec{x} \cdot \vec{\sigma} \tag{11.815}$$
Note $g \in SL(2,\mathbb{C})$ acts linearly via $\rho(g) \cdot M = gMg^\dagger$ since this preserves Hermiticity. However
$$\det M_{x^\mu} = (x^0)^2 - \vec{x}^2 \tag{11.816}$$
and the determinant is preserved by this actions so $g \mapsto \rho(g)$ describes $SL(2,\mathbb{C})$ as a double cover of the connected component $SO_0(1,3)$ of the $3+1$ dimensional Lorentz group.

4. $\mathfrak{su}(2) \oplus \mathfrak{su}(2)$ *And Rotations In* $\mathbb{R}^4$ Finally, a very similar construction gives the double cover $\pi : SU(2) \times SU(2) \to SO(4)$. But this is best discussed in the context of the *quaternions*. See the section on quaternions in Chapter 2.

---

**Exercise** *Another Proof That* $T_1 GL(n,\kappa) = \mathfrak{gl}(n,\kappa)$

Show that for $G = GL(n, \kappa)$ the corresponding Lie algebra is $\mathfrak{gl}(n, \kappa)$ by considering the 1-parameter subgroups $\exp[te_{ij}]$ where $e_{ij}$ are the matrix units.

---

---

**Exercise**

Interpret equation (11.703) as a special case of (11.808) for the adjoint representation. Use this to derive (11.732) from the group homomorphism property of $\rho$.

---

---

**Exercise** *Tensor Products Of Representations*

We have noted that if $\rho_1 : G \to \mathrm{Aut}(V_1)$ and $\rho_2 : G \to \mathrm{Aut}(V_2)$ are two representations of any group $G$ then there is a tensor product representation

$$\rho_{12}(g) := \rho_1(g) \otimes \rho_2(g) \tag{11.817}$$

Show that if $G$ is a Lie gorup then the corresponding Lie algebra representation on $V_1 \otimes V_2$ representations $X \in \mathfrak{g} = T_1 G$ by

$$\dot{\rho}_{12}(X) = \dot{\rho}_1(X) \otimes 1_{V_2} + 1_{V_1} \otimes \dot{\rho}_2(X) \tag{11.818}$$

---

### 11.21.5 Finite Dimensional Irreducible Representations Of $\mathfrak{sl}(2,\mathbb{C})$,$\mathfrak{sl}(2,\mathbb{R})$, and $\mathfrak{su}(2)$

*Relation of* $\mathfrak{sl}(2,\mathbb{R})$*,* $\mathfrak{su}(2)$ *and* $\mathfrak{sl}(2,\mathbb{C})$:

If we regard $\mathfrak{sl}(2,\mathbb{C})$ as a Lie algebra over $\kappa = \mathbb{R}$ there are two inequivalent real Lie subalgebras: $\mathfrak{sl}(2,\mathbb{R})$ and $\mathfrak{su}(2)$, as we have described in detail above. However, if we allow ourselves to multiply by complex numbers, that is, if we consider $\mathfrak{sl}(2,\mathbb{R}) \otimes \mathbb{C}$ and $\mathfrak{su}(2) \otimes \mathbb{C}$ then we obtain a single Lie algebra $\mathfrak{sl}(2,\mathbb{C})$. Indeed, in the complexification, we have

$$\begin{aligned} e &= \mathrm{i}T^1 - T^2 \\ h &= -2\mathrm{i}T^3 \\ f &= -\mathrm{i}T^1 - T^2 \end{aligned} \tag{11.819}$$

Therefore, there is no distinction between the finite-dimensional representations of $\mathfrak{su}(2)$, $\mathfrak{sl}(2,\mathbb{R})$, and $\mathfrak{sl}(2,\mathbb{C})$ on complex vector spaces. It is actually easiest to construct the finite dimensional representation of $\mathfrak{sl}(2,\mathbb{R})$ or $\mathfrak{sl}(2,\mathbb{C})$ on a complex vector space. But by the above formulae these will immediately give the finite dimensional representations of $\mathfrak{su}(2)$.

Suppose we have a finite-dimensional complex vector space $V$ and linear operators $\rho(e), \rho(f), \rho(h)$ on $V$ satisfying the commutation relations:

$$\begin{aligned}[\rho(h), \rho(f)] &= 2\rho(f) \\ [\rho(h), \rho(e)] &= -2\rho(e) \\ [\rho(e), \rho(f)] &= \rho(h)\end{aligned} \tag{11.820}$$

As shown in Chapter two, any linear operator on a complex vector space has at least one eigenvector. (It might have only one eigenvector.)

Suppose we choose an eigenvector $v$ of $\rho(h)$ and suppose the eigenvalue is $\lambda$. Then, we claim that, so long as $\rho(e)^n v \neq 0$ the vector $\rho(e)^n v$ has eigenvalue $\lambda - 2n$. To prove this apply the general identity

$$[A, B^n] = \sum_{i=0}^{n-1} B^i [A, B] B^{n-1-i} \tag{11.821}$$

to conclude:

$$[\rho(h), \rho(e)^n] = -2n\rho(e)^n \tag{11.822}$$

and the result follows. Now, it is general fact of linear algebra that if we have nonzero vectors $v_1, \ldots, v_n$ with <u>distinct</u> eigenvalues $\lambda_1, \ldots, \lambda_n$ for some operator then the $v_1, \ldots, v_n$ are linearly independent. (Prove this as an exercise.) Therefore, if $\rho(e)^n v \neq 0$ then the vectors $v, \rho(e)v, \ldots, \rho(e)^n v$ are linearly independent. Therefore, since we have a finite-dimensional representation there must be a nonnegative integer $n$ so that $\rho(e)^n v \neq 0$ but $\rho(e)^{n+1} v = 0$. Let us denote $v_0 := \rho(e)^n v$. So $\rho(e) v_0 = 0$ with $\rho(h) v_0 = \lambda_0 v_0$ and $v_0 \neq 0$. Now, using (11.821) again we get:

$$[\rho(h), \rho(f)^k] = 2k\rho(f)^k \tag{11.823}$$

and therefore

$$\rho(h)(\rho(f)^k v_0) = (\lambda_0 + 2k)\rho(f)^k v_0 \tag{11.824}$$

By a similar argument to that above, we know that if $\rho(f)^n v_0$ is nonzero then the vectors $v_0, \rho(f)v_0, \ldots, \rho(f)^n v_0$ are linearly independent, and therefore there must exist an integer $N$ so that $\rho(f)^N v_0 \neq 0$ but $\rho(f)^{N+1} v_0 = 0$. Therefore

$$[\rho(e), \rho(f)^{N+1}] v_0 = 0 \tag{11.825}$$

Now, using

$$[\rho(e), \rho(f)^{N+1}] = \sum_{i=0}^{N} \rho(f)^i [\rho(e), \rho(f)] \rho(f)^{N-i} = \sum_{i=0}^{N} \rho(f)^i \rho(h) \rho(f)^{N-i} \tag{11.826}$$

and applying this identity to $v_0$ we get:

$$0 = \left( \sum_{i=0}^{N} (\lambda_0 + 2(N-i)) \right) \rho(f)^N v_0 \tag{11.827}$$

But $\rho(f)^N v_0 \neq 0$, by the definition of $N$ so

$$\left(\sum_{i=0}^{N}(\lambda_0 + 2(N-i))\right) = 0 \tag{11.828}$$

which implies $\lambda_0 = -N$. Thus, we have produced a set of vectors

$$v_0, \rho(f)v_0, \cdots, \rho(f)^N v_0 = v_0, v_1, \ldots, v_N \tag{11.829}$$

spanning an $(N+1)$-dimensional representation of $\mathfrak{sl}(2,\mathbb{R})$ inside any finite-dimensional representation. Indeed, if we define:

$$\tilde{v}_s := \rho(f)^s v_0 \tag{11.830}$$

understanding that $v_{N+1} = v_{N+2} = \cdots = 0$ then

$$\begin{aligned} \rho(f)\tilde{v}_s &= \tilde{v}_{s+1} \\ \rho(h)\tilde{v}_s &= (2s-N)\tilde{v}_s \\ \rho(e)\tilde{v}_s &= -s(N+1-s)\tilde{v}_{s-1} \end{aligned} \tag{11.831}$$

So the span, which we will denote $W_N \subset V$ is a subrepresentation. Note that in this ordered basis for $W_N$ the representation matrices are: (See chapter 2 for the proper way to associate a matrix to a linear transformation on a basis with an ordered basis.)

$$\rho(f) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{11.832}$$

$$\rho(h) = \begin{pmatrix} -N & 0 & 0 & \cdots & 0 & 0 \\ 0 & -N+2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -N+4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & N-2 & 0 \\ 0 & 0 & 0 & \cdots & 0 & N \end{pmatrix} \tag{11.833}$$

$$\rho(e) = \begin{pmatrix} 0 & -N & 0 & \cdots & 0 & 0 \\ 0 & 0 & -2(N-1) & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -N \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \tag{11.834}$$

Note that the Jordan form for $\rho(f)$ shows that the representation $W_N$ is irreducible.

As mentioned above, we automatically get a representation of $\mathfrak{su}(2)$, and hence of the group $SU(2)$. Therefore $V$ is completely reducible. Therefore, we have recovered - this time at the Lie algebra level - that there is one irreducible representation of $SU(2)$ of dimension $n$ for each positive integer $n$. In fact, identifying $N = 2j$ the representation $W_N$ gives another model for the spin $j$ representation $V_j$.

**Remarks**

1. In physics the *angular momentum* operators in quantum mechanics generate the Lie algebra $\mathfrak{su}(2)$. See the exercise below. In physics observable quantities (like angular momentum) are represented by Hermitian operators. The observable angular momentum operators are related to the anti-Hermitian generators of $\mathfrak{su}(2)$ by a factor of $\sqrt{-1}$. Physicists usually define

$$J^a := \mathrm{i}T^a = \frac{1}{2}\sigma^a \tag{11.835}$$

and they satisfy the commutation relations

$$[J^a, J^b] = \mathrm{i}\epsilon^{abc}J^c \tag{11.836}$$

which, unfortunately, obscures the fact that we are working with a real Lie algebra. Group elements are of the form $\exp[i\vec{\theta}\cdot\vec{J}]$ where $\vec{\theta}$ is real. When working with representations physicists generally define

$$\begin{aligned} J^+ &= J^1 + \mathrm{i}J^2 = e \\ J^- &= J^1 - \mathrm{i}J^2 = -f \end{aligned} \tag{11.837}$$

so that

$$\begin{aligned} [J^3, J^\pm] &= \pm J^\pm \\ [J^+, J^-] &= 2J^3 \end{aligned} \tag{11.838}$$

More generally, in a physical system with $SU(2)$ symmetry the $J^a$ are the conserved "Noether charges" for that symmetry. Recall that we called the irreducible representations of $SU(2)$ of dimension $n = 2j+1$ the "spin $j$ representations." This is the origin of that terminology.

2. $V_j$ is a unitary representation of $\mathfrak{su}(2)$ and $SU(2)$ but is definitely <u>not</u> unitary as a representation of $\mathfrak{sl}(2,\mathbb{R})$. (There are unitary irreps of $\mathfrak{sl}(2,\mathbb{R})$, but they are infinite-dimensional.) The standard physics notation for an ON basis diagonalizing $\rho(J^3)$ is $|j,m\rangle$ with

$$\begin{aligned} \rho(J^+)|j,m\rangle &= \sqrt{(j-m)(j+m+1)}|j,m+1\rangle \\ \rho(J^3)|j,m\rangle &= m|j,m\rangle \\ \rho(J^-)|j,m\rangle &= \sqrt{(j+m)(j-m+1)}|j,m-1\rangle \end{aligned} \tag{11.839}$$

with $m = -j, -j+1, \dots, j-1, j$. We can relate this to our basis by identifying $N = 2j$ and

$$\rho(f)^s v_0 = C_{j,s} |j, j-s\rangle \tag{11.840}$$

for a suitable normalization factor $C_{j,s}$.

---

**Exercise** *Justifying The Relation To Angular Momentum*

In the classical mechanics of a particle moving in $\mathbb{R}^3$ the angular momentum of a particle around the origin is the function on phase space given by

$$L_a = \epsilon_{abc} x_b p_c \qquad a, b, c \in \{1, 2, 3\} \tag{11.841}$$

$L_a$ is a (pseudo)-vector. In quantum mechanics we adopt the same expression and there is no issue of operator ordering because the epsilon symbol prevents $b = c$:

$$\hat{L}_a = \epsilon_{abc} \hat{x}_b \hat{p}_c \qquad a, b, c \in \{1, 2, 3\} \tag{11.842}$$

Using $[\hat{p}_a, \hat{x}_b] = -\mathrm{i}\hbar\delta_{ab}$ show that

$$[\hat{L}_a, \hat{L}_b] = \mathrm{i}\epsilon_{abc}\hbar\hat{L}_c \tag{11.843}$$

so that $\hat{T}^a = -\frac{\mathrm{i}}{\hbar}\hat{L}_a$ are anti-Hermitian operators generating a copy of $\mathfrak{su}(2)$.

**Remark:** In spin $j$ representations if $j$ is order 1 the physical angular momentum is of order $\hbar$ and hence intrinsically quantum mechanical. On the other hand, for fixed $\hat{L}$, the semiclassical $\hbar \to 0$ limit is the large spin limit.

---

---

**Exercise** *Another basis*

Show that if we use the basis

$$v_s = \frac{1}{s!}\rho(f)^s v_0 \tag{11.844}$$

then

$$\begin{aligned} \rho(f) v_s &= (s+1) v_{s+1} \\ \rho(h) v_s &= (2s - N) v_s \\ \rho(e) v_s &= (s - N - 1) v_{s-1} \end{aligned} \tag{11.845}$$

---

---

**Exercise** *Unitarizing $V_j$*

Unitarizing $V_j$ means equipping it with an inner product so that the $\rho(J^a)$ are Hermitian operators. Equivalently, $\rho(h)$ is Hermitian and

$$\rho(J^+)^\dagger = \rho(J^-) \qquad \leftrightarrow \qquad \rho(e)^\dagger = -\rho(f) \tag{11.846}$$

Show that by rescaling $u_s = c_s \tilde{v}_s$ and declaring $u_s$ to be ON we get a unitary structure provided that

$$\left(\frac{c_{s-1}}{c_s}\right)^2 = s(2j+1-s) \tag{11.847}$$

so that

♣Fix this equation: sign issue ♣

$$\begin{aligned} \rho(e)u_s &= -\sqrt{s(2j+1-s)}u_{s-1} \\ \rho(f)u_s &= \sqrt{s(2j+1-s)}u_{s+1} \end{aligned} \tag{11.848}$$

---

### 11.21.6 Casimirs

We have stressed that if two elements $a, b$ in a Lie subalgebra of a matrix Lie algebra are multiplied as matrices $ab$ then in general the result is not in the Lie algebra. Nevertheless, if we have a representation $\rho(a), \rho(b) \in \mathrm{End}(V)$ nothing stops us form multiplying the operators $\rho(a)\rho(b)$. Certain relations among the algebra of the operators $\rho(a)$ for $a \in \mathfrak{g}$ are universal and independent of the representation. They can be expressed in terms of tensor algebra using the *universal enveloping algebra*. See Chapter two for a description. Here we just look at one important aspect of such universal relations.

In any representation of $\mathfrak{su}(2)$ the operator

$$C_2(\rho) := \sum_{a=1}^{3} \rho(T^a)^2 \tag{11.849}$$

commutes with all the operators $\rho(T^a)$. This operator is known as a quadratic Casimir. It is a theorem in the theory of universal enveloping algebras that any representation, any operator that commutes with all operators made by multiplying and adding the $\rho(T^a)$ is a polynomial in the operator $C_2(\rho)$. This fact generalizes to all simple Lie algebras: The center of the universal enveloping algebra is a polynomial in the Casimirs and there are $r$ independent Casimirs where $r$ is the rank. For $SU(N)$ there are $N-1$ independent Casimirs.

Returning to $SU(2)$, we can express $C_2(\rho)$ in terms of the representations of the basis for $\mathfrak{sl}(2,\mathbb{R})$ using

$$C_2(\rho) = \frac{1}{4}\left(2(\rho(e)\rho(f) + \rho(f)\rho(e)) - \rho(h)^2\right) \tag{11.850}$$

In any irreducible representation the Casimir operator must be a multiple of the identity operator. We can easily compute the value by acting on any convenient vector. For example,

$$\begin{aligned} C_2(\rho)v_0 &= \frac{1}{4}\left(2(\rho(e)\rho(f) + \rho(f)\rho(e)) - \rho(h)^2\right)v_0 \\ &= -\frac{N(N+2)}{4}v_0 \end{aligned} \tag{11.851}$$

which the physicists will prefer to write as

$$\rho_j(\vec{J})^2 = j(j+1)\mathbf{1}_{(2j+1)\times(2j+1)} \tag{11.852}$$

where $\rho_j$ is the representation on $V_j$ and we have merely substituted $N = 2j$ above.

**Examples**:

1. Consider a quantum system of two spin 1/2 particles with Hamiltonian:

$$H = J_1^a \otimes J_2^a \tag{11.853}$$

We can easily use representation theory to find the spectrum of this Hamiltonian. We note that

$$H = \frac{1}{2}\Big((J_1^a \otimes 1 + 1 \otimes J_2^a)^2 - \vec{J}_1^2 \otimes 1 - 1 \otimes \vec{J}_2^2\Big) \tag{11.854}$$

But the Hilbert space is $V_{1/2} \otimes V_{1/2} \cong V_0 \oplus V_1$. On the one-dimensional subspace $\cong V_0$ we use (11.852) to compute

$$H|_{V_0} = \frac{1}{2}(0 - \frac{3}{4} - \frac{3}{4}) = -\frac{3}{4} \tag{11.855}$$

But on the three-dimensional subspace $\cong V_1$ we have

$$H|_{V_1} = \frac{1}{2}(2 - \frac{3}{4} - \frac{3}{4}) = +\frac{1}{4} \tag{11.856}$$

---

**Exercise** *Three Qbits On A Ring*

Consider a quantum system of three spin $1/2$ particles on a ring with $\vec{J}_i \cdot J_{i+1}$ interaction between neighboring spins. Compute the spectrum of the Hamiltonian.

$$H = \sum_i J_i^a \otimes J_{i+1}^a \tag{11.857}$$

where $i$ is understood as an index modulo 3. [188]

---

[188] *Answer*: We can write:

$$H = \frac{1}{2}\Big((J_1^a \otimes 1 \otimes 1 + 1 \otimes J_2^a \otimes 1 + 1 \otimes 1 \otimes J_3^a)^2 - \vec{J}_1^2 \otimes 1 \otimes 1 - 1 \otimes \vec{J}_2^2 \otimes 1 - 1 \otimes 1 \otimes \vec{J}_3^2\Big) \tag{11.858}$$

But we have the isotypical decomposition:

$$V_{1/2} \otimes V_{1/2} \otimes V_{1/2} \cong 2V_{1/2} \oplus V_{3/2} \tag{11.859}$$

On the 4 dimensional space $2V_{1/2}$ we have

$$H|_{2V_{1/2}} = -\frac{3}{4} \tag{11.860}$$

and on the 4-dimensional space $V_{3/2}$ we have

$$H|_{V_{3/2}} = +\frac{3}{4} \tag{11.861}$$

### 11.21.7 Lie Algebra Operators In The Induced Representations Of $SU(2)$

Using the above models for the irreducible representations of $SU(2)$ and the relation between elements of the Lie algebra and infinitesimal group elements we get representations of the Lie algebra in terms of differential operators.

In terms of homogeneous polynomials $\psi(u,v)$ of two variables in the space $\mathcal{H}_{2j}$ recall that

$$(\rho(g)\cdot\psi)(\begin{pmatrix} u \\ v \end{pmatrix}) := \psi\left(g^{-1}\begin{pmatrix} u \\ v \end{pmatrix}\right) \tag{11.862}$$

Now suppose $g$ is infinitesimally close to the identity so that

$$(\rho(1+\epsilon X)\cdot\psi)(\begin{pmatrix} u \\ v \end{pmatrix}) := \psi\left((1-\epsilon X)\begin{pmatrix} u \\ v \end{pmatrix}\right) \tag{11.863}$$

up to order $\epsilon^2$. Since $\rho(X)$ is linear in $X$ we can form complex linear combinations, and therefore the same formula will apply to any $X \in \mathfrak{sl}(2,\mathbb{C})$. In particular, we deduce:

$$\begin{aligned}
\rho(e)\cdot\psi &= -v\frac{\partial}{\partial u}\psi \\
\rho(h)\cdot\psi &= \left(u\frac{\partial}{\partial u} - v\frac{\partial}{\partial v}\right)\psi \\
\rho(f)\cdot\psi &= u\frac{\partial}{\partial v}\psi
\end{aligned} \tag{11.864}$$

The reader should check the operators really satisfy $\rho([X,Y]) = [\rho(X),\rho(Y)]$. Note that the Casimir is:

$$C_2(\rho) = -\frac{1}{4}\left(u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v}\right)^2 - \frac{1}{2}\left(u\frac{\partial}{\partial u} + v\frac{\partial}{\partial v}\right) \tag{11.865}$$

so acting on $\mathcal{H}_{2j}$ we immediately get that $C_2(\rho_j)$ acts as the scalar operator of multiplication by $j(j+1)$.

It is also interesting to consider the action in the inhomogeneous representation. As above, we extend the action to all of $SL(2,\mathbb{C})$ so that if

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{C}) \tag{11.866}$$

then

$$\rho(g)\cdot p(z) = (-cz+a)^{2j} p\left(\frac{dz-b}{-cz+a}\right) \tag{11.867}$$

and if $g = 1+\epsilon X + \mathcal{O}(\epsilon^2)$ then $\rho(g) = 1+\epsilon\rho(X)+\mathcal{O}(\epsilon^2)$, and in this way we derive:

$$\begin{aligned}
\rho(e)\cdot p &= -\frac{\partial}{\partial z}p \\
\rho(h)\cdot p &= -(2z\frac{\partial}{\partial z} + 2j)p \\
\rho(f)\cdot p &= -(z^2\frac{\partial}{\partial z} - 2jz)p
\end{aligned} \tag{11.868}$$

For example:

$$
\begin{aligned}
(\rho(1+\epsilon f)\cdot p)(z) &:= (\epsilon z+1)^{2j} p\left(\frac{z}{\epsilon z+1}\right) \\
&= (1+2j\epsilon z)p(z-\epsilon z^2)+\mathcal{O}(\epsilon^2) \\
&= p(z)-\epsilon z^2\frac{\partial}{\partial z}p(z)+\epsilon 2jzp(z)+\mathcal{O}(\epsilon^2)
\end{aligned}
\tag{11.869}
$$

Similarly, acting on the Wigner functions themselves we learn that, after complexification $\epsilon\rho(e)$ translates the group element in the argument of $D^j_{m_L,m_R}$ by

$$
g=\begin{pmatrix}\alpha & -\bar\beta \\ \beta & \bar\alpha\end{pmatrix}\to g-\begin{pmatrix}\epsilon\beta & \epsilon\bar\alpha \\ 0 & 0\end{pmatrix}
\tag{11.870}
$$

and similarly $\epsilon\rho(f)$ translates it by

$$
g=\begin{pmatrix}\alpha & -\bar\beta \\ \beta & \bar\alpha\end{pmatrix}\to g+\begin{pmatrix}0 & 0 \\ \epsilon\alpha & -\epsilon\bar\beta\end{pmatrix}
\tag{11.871}
$$

So that, acting on the Wigner functions considered as polynomials in $\alpha,\bar\alpha,\beta,\bar\beta$ if we treat $\alpha,\bar\alpha,\beta,\bar\beta$ as independent then

$$
\begin{aligned}
\rho(e) &= \bar\alpha\frac{\partial}{\partial\bar\beta}-\beta\frac{\partial}{\partial\alpha} \\
\rho(h) &= \bar\alpha\frac{\partial}{\partial\bar\alpha}-\alpha\frac{\partial}{\partial\alpha}-\bar\beta\frac{\partial}{\partial\bar\beta}+\beta\frac{\partial}{\partial\beta} \\
\rho(f) &= \alpha\frac{\partial}{\partial\beta}-\bar\beta\frac{\partial}{\partial\bar\alpha}
\end{aligned}
\tag{11.872}
$$

One can check explicitly that

$$
\begin{aligned}
\rho(e)\tilde D^j_{m_L,m_R} &= const.\tilde D^j_{m_L+1,m_R} \\
\rho(h)\tilde D^j_{m_L,m_R} &= \tilde D^j_{m_L,m_R} \\
\rho(f)\tilde D^j_{m_L,m_R} &= const.\tilde D^j_{m_L-1,m_R}
\end{aligned}
\tag{11.873}
$$

We can specialize this to get the standard identities on spherical harmonics which are frequently used in mathematical physics: [189] ♣NEED TO GIVE MORE DETAILS ON THIS LAST REMARK ♣

$$
\begin{aligned}
L_+ &= e^{\mathrm{i}\phi}\left(\frac{\partial}{\partial\theta}+\mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right) \\
L_- &= e^{-\mathrm{i}\phi}\left(-\frac{\partial}{\partial\theta}+\mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right) \\
L_3 &= -\mathrm{i}\frac{\partial}{\partial\phi}
\end{aligned}
\tag{11.874}
$$

[189] See, for example, J.D. Jackson, *Classical Electrodynamics*, page 743.

$$
\begin{aligned}
L_+ Y_{\ell,m} &= \sqrt{(\ell-m)(\ell+m+1)}Y_{\ell,m+1} \\
L_- Y_{\ell,m} &= \sqrt{(\ell+m)(\ell-m+1)}Y_{\ell,m-1} \\
L_3 Y_{\ell,m} &= mY_{\ell,m} \\
\vec{L}^2 Y_{\ell,m} &= \ell(\ell+1)Y_{\ell,m}
\end{aligned}
\tag{11.875}
$$

In particular $\vec{L}^2$, the Casimir, as a differential shows up in the expression for the Laplacian expressed in terms of spherical coordinates. The equation $\vec{L}^2 Y_{\ell,m} = \ell(\ell+1)Y_{\ell,m}$ becomes the standard differential equation satisfied by associated Legendre functions and Legendre polynomials, whereas the first two lines of (11.875) become standard identities relating associated Legendre functions.

**Remark**: We have seen above that group theory gives a nice perspective on many identities satisfied by the family of special functions associated with Legendre polynomials and spherical harmonics. This viewpoint extends very nicely to many other special functions in mathematical physics. Two references that explain this in some detail are:

1. J.D. Talman, *Special Functions: A Group Theoretic Approach Based on Lectures by Eugene P. Wigner*

2. N. Ja. Vilenkin, *Special Functions and the Theory of Group Representations*

♣The following section on Kernel, Image, Exact Sequence should be moved to be just before the representation theory section, but the quantum mechanics on the circle should be moved either to the Heisenberg section or to the Pontryagin duality section. ♣

## 12. Group Theory And Elementary Number Theory

In this chapter we review some very elementary number theory that has a strong connection to group theory. The facts here can be very useful in thinking about many physics problems.

Two general references are

Hardy and Wright, *An Introduction To The Theory Of Numbers*

Ireland and Rosen, *A Classical Introduction to Modern Number Theory*

### 12.1 Reminder On gcd And The Euclidean Algorithm

Let us recall some basic facts from grade school arithmetic:

First, if $A > B$ are two positive integers then we can write

$$
A = qB + r \qquad 0 \leq r < B \tag{12.1}
$$

for unique nonnegative integers $q$ and $r$ known as the *quotient* and the *residue*, respectively.

Next, let $(A, B) = (\pm A, \pm B) = (\pm B, \pm A)$ denote the greatest common divisor of $A, B$. Then we can find it using the *Euclidean algorithm* by looking at successive quotients. If

$A = qB$ with $r = 0$ we are done! Then $(A, B) = B$. If $r > 0$ then we proceed as follows:

$$\begin{aligned} A &= q_1 B + r_1 \qquad 0 < r_1 < B \\ B &= q_2 r_1 + r_2 \qquad 0 < r_2 < r_1 \\ r_1 &= q_3 r_2 + r_3 \qquad 0 < r_3 < r_2 \\ r_2 &= q_4 r_3 + r_4 \qquad 0 < r_4 < r_3 \\ &\vdots \quad \vdots \\ r_{j-2} &= q_j r_{j-1} + r_j \qquad 0 < r_j < r_{j-1} \\ r_{j-1} &= q_{j+1} r_j \end{aligned} \tag{12.2}$$

Note that $B > r_1 > r_2 > \cdots \geq 0$ is a strictly decreasing sequence of nonnegative integers and hence must terminate at $r_* = 0$ after a finite number of steps.

**Examples**

$A = 96$ and $B = 17$:

$$\begin{aligned} 96 &= 5 \cdot 17 + 11 \\ 17 &= 1 \cdot 11 + 6 \\ 11 &= 1 \cdot 6 + 5 \\ 6 &= 1 \cdot 5 + 1 \\ 5 &= 5 \cdot 1 \end{aligned} \tag{12.3}$$

$A = 96$ and $B = 27$:

$$\begin{aligned} 96 &= 3 \cdot 27 + 15 \\ 27 &= 1 \cdot 15 + 12 \\ 15 &= 1 \cdot 12 + 3 \\ 12 &= 4 \cdot 3 \end{aligned} \tag{12.4}$$

Note well: In (12.1) the remainder might be zero but in the first $j$ lines of the Euclidean algorithm the remainder is positive, unless $B$ divides $A$, in which case rather trivially $(A, B) = B$. The last positive remainder $r_j$ is the gcd $(A, B)$. Indeed if $m_1, m_2$ are integers then the gcd satisfies:

$$(m_1, m_2) = (m_2, m_1) = (m_2, m_1 - x m_2) \tag{12.5}$$

for any integer $x$. Applying this to the Euclidean algorithm above we get:

$$(A, B) = (B, r_1) = (r_1, r_2) = \cdots = (r_{j-1}, r_j) = (r_j, 0) = r_j. \tag{12.6}$$

A corollary of this algorithm is that if $g = (A, B)$ is the greatest common divisor then there exist integers $(x, y)$ so that

$$Ax + By = g \tag{12.7}$$

In particular, two integers $m_1, m_2$ are *relatively prime*, that is, have no common integral divisors other than $\pm 1$, if and only if there exist integers $x, y$ such that

$$m_1 x + m_2 y = 1. \tag{12.8}$$

Of course $x, y$ are not unique. Equation (12.8) is sometimes known as "Bezout's theorem." We can prove these statements from the Euclidean algorithm as follows.

♣Putting this discussion here makes part of the section on $SL(2,\mathbb{Z})$ and continued fractions a little redundant. ♣

For an integer $n$ define

$$T(n) := \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} = T^n \tag{12.9}$$

where

$$T := T(1) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \ . \tag{12.10}$$

Now let us write the first line of the Euclidean algorithm as a matrix identity as

$$T(-q_1) \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} r_1 \\ B \end{pmatrix} \tag{12.11}$$

and better, we write this as:

$$\sigma^1 T(-q_1) \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} B \\ r_1 \end{pmatrix} \tag{12.12}$$

Then the second line of the Euclidean algorithm becomes:

$$\sigma^1 T(-q_2) \sigma^1 T(-q_1) \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \tag{12.13}$$

Thus we have

$$\sigma^1 T(-q_j) \cdots \sigma^1 T(-q_2) \sigma^1 T(-q_1) \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} r_{j-1} \\ r_j \end{pmatrix} \tag{12.14}$$

and in the final step:

$$\sigma^1 T(-q_{j+1}) \sigma^1 T(-q_j) \cdots \sigma^1 T(-q_2) \sigma^1 T(-q_1) \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} r_j \\ 0 \end{pmatrix} \tag{12.15}$$

Multiplying out the matrices on the LHS gives an expression:

$$\begin{pmatrix} x & y \\ u & v \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} r_j \\ 0 \end{pmatrix} \tag{12.16}$$

(Note that $x, y, u, v$ are polynomials in the $q_i$. See comments on continued fractions below.)

**Remarks**:

1. *The Euclidean algorithm is fast*: A theorem of Lamé asserts that the Euclidean algorithm is very efficient. It should be completely obvious to you that the number of steps cannot exceed $B$. (Recall that $A > B$.) However, Lamé asserts that in fact the number of steps never exceeds $5\log_{10} B$. This is important for RSA (see below).

2. *Relation to continued fractions*: Note that from equation (12.15) we can also write

$$\begin{pmatrix} A \\ B \end{pmatrix} = T(q_1)\sigma^1 T(q_2)\sigma^1 \cdots T(q_j)\sigma^1 T(q_{j+1})\sigma^1 \begin{pmatrix} r_j \\ 0 \end{pmatrix} \tag{12.17}$$

Let us write:

$$M(q) := T(q)\sigma^1 = \begin{pmatrix} q & 1 \\ 1 & 0 \end{pmatrix} \tag{12.18}$$

We now define two sequences of polynomials in $n$ variables that we call $N_n(q_1, \ldots, q_n)$ and $D_n(q_1, \ldots, q_n)$ for all $n \geq 1$. It is convenient to define $N_0 = 1$ and $D_0 = 1$ and then we can write:

$$M(q_1) \cdots M(q_n) := \begin{pmatrix} N_n(q_1, \ldots, q_n) & N_{n-1}(q_1, \ldots, q_{n-1}) \\ D_n(q_1, \ldots, q_n) & D_{n-1}(q_1, \ldots, q_{n-1}) \end{pmatrix} \tag{12.19}$$

(The reader should check that this is a consistent definition for all $n$.) One easily generates:

$$\begin{aligned} N_1(q_1) &= q_1 \\ N_2(q_1, q_2) &= 1 + q_1 q_2 \\ N_3(q_1, q_2, q_3) &= q_1 + q_3 + q_1 q_2 q_3 \end{aligned} \tag{12.20}$$

$$\begin{aligned} D_1(q_1) &= 1 \\ D_2(q_1, q_2) &= q_2 \\ D_3(q_1, q_2, q_3) &= 1 + q_2 q_3 \end{aligned} \tag{12.21}$$

These polynomials are closely related to continued fractions, defined as:

$$[q_1, q_2, q_3, \cdots, q_j] := q_1 + \cfrac{1}{q_2 + \frac{1}{q_3 + \cdots + \frac{1}{q_j}}} \tag{12.22}$$

Indeed, now that

$$M(q_1) \cdot (M(q_2) \cdots M(q_{n+1})) = M(q_1) \begin{pmatrix} N_n(q_2, \ldots, q_{n+1}) & N_{n-1}(q_2, \ldots, q_n) \\ D_n(q_2, \ldots, q_{n+1}) & D_{n-1}(q_2, \ldots, q_n) \end{pmatrix} \tag{12.23}$$

from which one deduces the recursion relations:

$$\begin{aligned} N_{n+1}(q_1, \ldots, q_{n+1}) &= q_1 N_n(q_2, \ldots, q_{n+1}) + D_n(q_2, \ldots, q_{n+1}) \\ D_{n+1}(q_1, \ldots, q_{n+1}) &= N_n(q_2, \ldots, q_{n+1}) \end{aligned} \tag{12.24}$$

On the other hand, writing

$$[q_1, q_2, q_3, \cdots, q_n] := \frac{P_n(q_1, \ldots, q_n)}{Q_n(q_1, \ldots, q_n)} \tag{12.25}$$

we see that

$$\begin{aligned}[q_1, q_2, q_3, \cdots, q_{n+1}] &= q_1 + \frac{1}{[q_2, \dots, q_{n+1}]} \\ &= \frac{q_1 P_n(q_2, \dots, q_{n+1}) + Q_n(q_2, \dots, q_{n+1})}{P_n(q_2, \dots, q_{n+1})}\end{aligned} \tag{12.26}$$

So, $P_n, Q_n$ satisfy the same recursion relations as $N_n, D_n$, respectively, and since the initial values are also the same we conclude that $P_n = N_n$ and $Q_n = D_n$.

---

**Exercise**
Check the Lamé bound for the two examples above.

---

---

**Exercise**
Given one solution for (12.7), find all the others.

---

---

**Exercise** *Continued fractions and the Euclidean algorithm*
a.) Show that the quotients $q_i$ in the Euclidean algorithm define a continued fraction expansion for $A/B$:

$$\frac{A}{B} = q_1 + \frac{1}{q_2 + \frac{1}{q_3 + \cdots + \frac{1}{q_j}}} := [q_1, q_2, q_3, \cdots, q_j] \tag{12.27}$$

The fractions $[q_1], [q_1, q_2], [q_1, q_2, q_3], \dots$ are known as the *convergents* of the continued fraction.
b.) Show that [190]

$$\begin{aligned}N_{n+1}(q_1, \dots, q_{n+1}) &= q_{n+1} N_n(q_1, \dots, q_n) + N_{n-1}(q_1, \dots, q_{n-1}) \\ D_{n+1}(q_1, \dots, q_{n+1}) &= q_{n+1} D_n(q_1, \dots, q_n) + D_{n-1}(q_1, \dots, q_{n-1})\end{aligned} \tag{12.28}$$

c.) Show that

$$N_n D_{n-1} - D_n N_{n-1} = (-1)^n \tag{12.29}$$

---

[190] *Answer*: Write $M(q_1) \cdots M(q_{n+1}) = (M(q_1) \cdots M(q_n)) \cdot M(q_{n+1})$.

## 12.2 Application: Expressing elements of $SL(2,\mathbb{Z})$ as words in $S$ and $T$

The group $SL(2,\mathbb{Z})$ is generated by

$$S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \qquad \& \qquad T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{12.30}$$

Here is an algorithm for decomposing an arbitrary element

$$h = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in SL(2,\mathbb{Z}) \tag{12.31}$$

as a word in $S$ and $T$.

First, note the following simple

**Lemma** Suppose $h \in SL(2,\mathbb{Z})$ as in (12.31). Suppose moreover that $g \in SL(2,\mathbb{Z})$ satisfies:

$$g \cdot \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{12.32}$$

Then

$$gh = T^n \tag{12.33}$$

for some integer $n \in \mathbb{Z}$.

The proof is almost immediate by combining the criterion that $gh \in SL(2,\mathbb{Z})$ has determinant one and yet must have the first column $(1,0)$.

Now, suppose $h$ is such that $A > C > 0$. Then $(A,C) = 1$ and hence we have the Euclidean algorithm to define integers $q_\ell$, $\ell = 1, \dots N+1$, where $N \geq 1$, such that

$$\begin{aligned} A &= q_1 C + r_1 & 0 < r_1 < C \\ C &= q_2 r_1 + r_2 & 0 < r_2 < r_1 \\ r_1 &= q_3 r_2 + r_3 & 0 < r_3 < r_2 \\ &\vdots \quad \vdots & \\ r_{N-2} &= q_N r_{N-1} + r_N & 0 < r_N < r_{N-1} \\ r_{N-1} &= q_{N+1} r_N & \end{aligned} \tag{12.34}$$

with $r_N = (A,C) = 1$. (Note you can interpret $r_0 = C$, as is necessary if $N = 1$.) Now, write the first line in the Euclidean algorithm in matrix form as: ♣$N = 0$ here? ♣

$$\begin{pmatrix} 1 & -q_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} r_1 \\ C \end{pmatrix} \tag{12.35}$$

We would like to have the equation in a form that we can iterate the algorithm, so we need the larger integer on top. Therefore, rewrite the identity as:

$$\sigma^1 \begin{pmatrix} 1 & -q_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} C \\ r_1 \end{pmatrix} \tag{12.36}$$

We can now iterate the procedure. So the Euclidean algorithm implies the matrix identity:

$$\tilde{g}\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{12.37}$$

$$\tilde{g} = (\sigma^1 T^{-q_{N+1}}) \cdots (\sigma^1 T^{-q_1}) \tag{12.38}$$

Now, to apply the Lemma we need $g$ to be in $SL(2,\mathbb{Z})$, but

$$\det\tilde{g} = (-1)^{N+1} \tag{12.39}$$

We can easily modify the equation to obtained a desired element $g$. We divide the argument into two cases:

1. Suppose first that $N+1 = 2s$ is even. Then we group the factors of $\tilde{g}$ in pairs and write

$$\begin{aligned}(\sigma^1 T^{-q_{2\ell}})(\sigma^1 T^{-q_{2\ell-1}}) &= (\sigma^1\sigma^3)(\sigma^3 T^{-q_{2\ell}}\sigma^3)(\sigma^3\sigma^1)T^{-q_{2\ell-1}} \\ &= -ST^{q_{2\ell}}ST^{-q_{2\ell-1}}\end{aligned} \tag{12.40}$$

   where we used that $\sigma^1\sigma^3 = -\mathrm{i}\sigma^2 = S$. Therefore, we can write

$$\tilde{g} = g = (-1)^s \prod_{\ell=1}^{s} (ST^{q_{2\ell}}ST^{-q_{2\ell-1}}) \tag{12.41}$$

2. Now suppose that $N+1 = 2s+1$ is odd. Then we rewrite the identity (12.37) as:

$$\sigma^1\tilde{g}\begin{pmatrix} A \\ C \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{12.42}$$

   so now we simply take

$$g = \sigma^1\tilde{g} = (-1)^{s+1}(ST^{-q_{2s+1}}) \prod_{\ell=1}^{s} (ST^{q_{2\ell}}ST^{-q_{2\ell-1}}) \tag{12.43}$$

Thus we can summarize both cases by saying that

$$g = (-1)^{\lfloor \frac{N+1}{2} \rfloor} \prod_{\ell=1}^{N+1} (ST^{(-1)^\ell q_\ell}) \tag{12.44}$$

Then we can finally write

$$h = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = g^{-1}T^n \tag{12.45}$$

as a word in $S$ and $T$ for a suitable integer $n$. (Note that $S^2 = -1$.)

♣It would be good to give an algorithm for determining $n$. ♣

Now we need to show how to bring the general element $h \in SL(2,\mathbb{Z})$ to the form with $A > C > 0$ so we can apply the above formula. Note that

$$\begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C+mA & D+mB \end{pmatrix} \tag{12.46}$$

while

$$\begin{pmatrix} 1 & 0 \\ -m & 1 \end{pmatrix} = ST^m S^{-1} \tag{12.47}$$

Thus, if $A > 0$ we can use this operation to shift $C$ so that $0 \le C < A$. In case $A < 0$ we can multiply by $S^2 = -1$ to reduce to the case $A > 0$. Finally, if $A = 0$ then

$$h = \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & n \end{pmatrix} \tag{12.48}$$

and we write

$$ST^n = \begin{pmatrix} 0 & -1 \\ 1 & n \end{pmatrix} \tag{12.49}$$

♣Need to summarize the result in a useful way ♣

## 12.3 Products Of Cyclic Groups And The Chinese Remainder Theorem

Recall the elementary definition we met in the last exercise of section 3.

**Definition** Let $H, G$ be two groups. The *direct product* of $H$ and $G$, denoted $H \times G$, is the set $H \times G$ with product:

$$(h_1, g_1) \cdot (h_2, h_2) = (h_1 \cdot h_2, g_1 \cdot g_2) \tag{12.50}$$

We will consider the direct product of cyclic groups. According to our general notation we would write this as $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$. However, since $\mathbb{Z}_m$ is also a ring the notation $\mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2}$ also often used, and we will use it below, especially when we write our Abelian groups additively.

Let us begin with the question: Is it true that

$$\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \overset{?}{\cong} \mathbb{Z}_{m_1 m_2}. \tag{12.51}$$

In general (12.51) is *false*!

---

**Exercise**

a.) Show that $\mathbb{Z}_4$ is not isomorphic to $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. (There is a one-line proof.) [191]

b.) Is $p$ is prime is $\mathbb{Z}_p \oplus \mathbb{Z}_p$ isomorphic to $\mathbb{Z}_{p^2}$ ?

c.) Is $\mathbb{Z}_3 \oplus \mathbb{Z}_5$ isomorphic to $\mathbb{Z}_{15}$ ?

---

Write $g = \gcd(m_1, m_2)$ and $\ell = \mathrm{lcm}(m_1, m_2)$. Then there are two natural exact sequences:

$$1 \to \mathbb{Z}_g \to \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \to \mathbb{Z}_\ell \to 1 \tag{12.52}$$

$$0 \to \mathbb{Z}/\ell\mathbb{Z} \to \mathbb{Z}/m_1\mathbb{Z} \oplus \mathbb{Z}/m_2\mathbb{Z} \to \mathbb{Z}/g\mathbb{Z} \to 0 \tag{12.53}$$

[191] *Answer*: Every element in $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ is of order two. But some elements of $\mathbb{Z}_4$ have order four. The but the order of a group element is preserved under isomorphism.

In fact, we will show below that

$$\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \cong \mathbb{Z}_g \times \mathbb{Z}_\ell \tag{12.54}$$

**Remarks**:

1. The sequence (12.52) is easier to write down multiplicatively, while (12.53) is easier to write down additively. See the discussion below. (Of course, both are true in either formulation!)
2. If $g = 1$ since $\mathbb{Z}_1 = \mathbb{Z}/\mathbb{Z}$ is the trivial group we can indeed conclude that $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \cong \mathbb{Z}_{m_1 m_2}$ but otherwise this is false. We will return to this point.

Now, let us prove (12.52) and (12.53).

Recall that

$$m_1 m_2 = g\ell \tag{12.55}$$

a fact that will be useful momentarily. (If you do not know this we will prove it below.) It will also be useful to write $m_1 = \mu_1 g$ and $m_2 = \mu_2 g$ where $\mu_1, \mu_2$ are relatively prime. Thus there are integers $\nu_1, \nu_2$ with

$$\mu_1 \nu_1 + \mu_2 \nu_2 = 1 \tag{12.56}$$

and hence

$$m_1 \nu_1 + m_2 \nu_2 = g. \tag{12.57}$$

To prove (12.52) think of $\mathbb{Z}_m$ as the multiplicative group of $m^{th}$ roots of 1, so they are all subgroups of $U(1)$. Now define a group homomorphism:

$$\pi : \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \to \mathbb{Z}_\ell \tag{12.58}$$

by:

$$\pi : (\xi_1, \xi_2) \to \xi_1 \xi_2 \tag{12.59}$$

That is, we merely multiply the two entries. (This makes it clear that it is a group homomorphism since the group law is multiplication of complex numbers and that multiplication is commutative.) Here $\xi_1$ is an $m_1^{th}$ root of unity and $\xi_2$ is an $m_2^{th}$ root of unity. The only thing you need to check is that indeed then $\xi_1 \xi_2$ is an $\ell^{th}$ root of unity, so $\pi$ indeed maps into $\mathbb{Z}_\ell$.

Now we prove that $\pi$ is surjective: Let $\omega_1 = e^{\frac{2\pi i}{m_1}}$ and $\omega_2 = e^{\frac{2\pi i}{m_2}}$. These are generators of $\mathbb{Z}_{m_1}$ and $\mathbb{Z}_{m_2}$. Choose integers $\nu_1, \nu_2$ so that $\nu_1 m_1 + \nu_2 m_2 = g$ then $\pi$ maps

$$\begin{aligned}
\pi : (\omega_1^{\nu_2}, \omega_2^{\nu_1}) \mapsto\ & \omega_1^{\nu_2} \omega_2^{\nu_1} \\
&= \exp\left[2\pi \mathrm{i} \left(\frac{\nu_2}{m_1} + \frac{\nu_1}{m_2}\right)\right] \\
&= \exp\left[2\pi \mathrm{i} \left(\frac{m_2 \nu_2 + m_1 \nu_1}{m_1 m_2}\right)\right] \\
&= \exp\left[2\pi \mathrm{i} \frac{g}{m_1 m_2}\right] \\
&= \exp\left[2\pi \mathrm{i} \frac{1}{\ell}\right]
\end{aligned} \tag{12.60}$$

But $\exp\left[2\pi i \frac{1}{\ell}\right]$ is a generator of the multiplicative group of $\ell^{th}$ roots of unity, isomorphic to $\mathbb{Z}_\ell$, and hence the homomorphism $\pi$ is onto. Thus, we have checked exactness of the sequence at $\mathbb{Z}_\ell$.

On the other hand the injection map

$$\iota : \mathbb{Z}_g \to \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \tag{12.61}$$

is defined by identifying $\mathbb{Z}_g$ with the multiplicative group of $g^{th}$ roots of unity and just sending:

$$\iota(\xi) = (\xi, \xi^{-1}) \tag{12.62}$$

Note that a $g^{th}$ root of unity $\xi$ has the property that $\xi^{\pm 1}$ is also both an $m_1^{th}$ and an $m_2^{th}$ root of unity. So this makes sense. It is now easy to check that indeed the kernel of $\pi$ is the image of $\iota$. Since $\pi$ takes the product of the two entries it is immediate from the definition (12.62) that $im(\iota) \subset \ker(\pi)$. On the other hand, if $\pi(\xi_1, \xi_2) = \xi_1\xi_2 = 1$ then clearly $\xi_2 = \xi_1^{-1}$, so this must be in the image of $\iota$. Now we have checked exactness at the middle of the sequence. Exactness at $\mathbb{Z}_g$ is trivial. This concludes the proof of (12.52) ♠

It is worth noting that we can write "additive" version of the maps $\iota$ and $\pi$ as:

$$\begin{aligned} \iota(x) &= \mu_1 x \oplus (-\mu_2 x) \\ \pi(x_1 \oplus x_2) &= \mu_2 x_1 + \mu_1 x_2 \end{aligned} \tag{12.63}$$

You should check that written this way it is well defined, and the sequence is exact.

---

**Exercise**

a.) Show that there is an exact sequence

$$0 \to \mathbb{Z}/\ell\mathbb{Z} \xrightarrow{\iota} \mathbb{Z}/m_1\mathbb{Z} \oplus \mathbb{Z}/m_2\mathbb{Z} \xrightarrow{\pi} \mathbb{Z}/g\mathbb{Z} \to 0 \tag{12.64}$$

where

$$\pi : x_1 \oplus x_2 \mapsto (x_1 - x_2) \mathrm{mod} g \ . \tag{12.65}$$

$$\iota : x \mapsto (x \mathrm{mod} m_1 \oplus x \mathrm{mod} m_2) \tag{12.66}$$

b.) Show that if we think of these groups as groups of roots of unity then we have $\pi(\xi_1, \xi_2) = \xi_1^{\mu_1}\xi_2^{-\mu_2}$ and $\iota(\omega) = (\omega^{\mu_2}, \omega^{\mu_1})$ with $m_1 = \mu_1 g$ and $m_2 = \mu_2 g$.

---

Now we prove that in general

$$\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \cong \mathbb{Z}_g \times \mathbb{Z}_\ell \tag{12.67}$$

First, it follows from either of the two exact sequences we proved above that if $(m_1, m_2) = 1$ then indeed

$$\mathbb{Z}_{m_1 m_2} \cong \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \tag{12.68}$$

Next, recall that any integer can be decomposed into its prime factors:

$$m = \prod_p p^{v_p(m)} \tag{12.69}$$

where $v_p(m) \in \mathbb{Z}_+$, known as the *valuation of $m$ at $p$* is zero for all but finitely many primes. (So we have an infinite product of 1's on the RHS of the above equation.)

Now in terms of the prime factorizations of $m_1, m_2$ we can write:

$$\begin{aligned} g = \gcd(m_1, m_2) &= \prod_p p^{\min[v_p(m_1), v_p(m_2)]} \\ \ell = \mathrm{lcm}(m_1, m_2) &= \prod_p p^{\max[v_p(m_1), v_p(m_2)]} \end{aligned} \tag{12.70}$$

Now, from the above we know that $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} \cong \mathbb{Z}_{m_1 m_2}$ if $m_1$ and $m_2$ are relatively prime. Therefore we can write

$$\mathbb{Z}/m\mathbb{Z} \cong \prod_p (\mathbb{Z}/p^{v_p(m)}\mathbb{Z}) \tag{12.71}$$

Applying this to each of the two factors in $\mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2}$ and using $G_1 \times G_2 \cong G_2 \times G_1$ to arrange the factors so the minimum power is on the left and maximum on the right and regrouping gives (12.67). In equations:

$$\begin{aligned} \mathbb{Z}_{m_1} \times \mathbb{Z}_{m_2} &\cong \prod_p \mathbb{Z}_{p^{\nu_p(m_1)}} \times \prod_p \mathbb{Z}_{p^{\nu_p(m_2)}} \\ &\cong \prod_p \mathbb{Z}_{p^{\min[\nu_p(m_1), \nu_p(m_2)]}} \times \prod_p \mathbb{Z}_{p^{\max[\nu_p(m_1), \nu_p(m_2)]}} \\ &\cong \mathbb{Z}_g \times \mathbb{Z}_\ell \end{aligned} \tag{12.72}$$

A second proof gives some additional insight by providing an interesting visual picture of what is going on, as well as relating this fact to lattices. It is related to the first by "taking a logarithm" and involves exact sequences of infinite groups which induce sequences on finite quotients.

Consider the sublattice of $\mathbb{Z} \oplus \mathbb{Z}$ given by

$$\Lambda = m_1\mathbb{Z} \oplus m_2\mathbb{Z} = \{ \begin{pmatrix} m_1\alpha \\ m_2\beta \end{pmatrix} | \alpha, \beta \in \mathbb{Z}\} \tag{12.73}$$

Then $\Lambda \subset \mathbb{Z} \oplus \mathbb{Z}$ is a sublattice and it should be pretty clear that

$$\mathbb{Z}^2/\Lambda = \mathbb{Z}_{m_1} \oplus \mathbb{Z}_{m_2} \tag{12.74}$$

Now, write $m_1 = \mu_1 g, m_2 = \mu_2 g$ as above. Choose integers $\nu_1, \nu_2$ so that $\mu_1\nu_1 + \mu_2\nu_2 = 1$ and consider the matrix

$$\begin{pmatrix} \mu_2 & \mu_1 \\ -\nu_1 & \nu_2 \end{pmatrix} \in SL(2, \mathbb{Z}) \tag{12.75}$$

This is an invertible matrix over the integers, so we can change coordinates on the lattice from $x = m_1\alpha, y = m_2\beta$ to

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \mu_2 & \mu_1 \\ -\nu_1 & \nu_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{12.76}$$

that is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \nu_2 & -\mu_1 \\ \nu_1 & \mu_2 \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} \tag{12.77}$$

which we prefer to write as:

$$\begin{pmatrix} x \\ y \end{pmatrix} = x' \begin{pmatrix} \nu_2 \\ \nu_1 \end{pmatrix} + y' \begin{pmatrix} -\mu_1 \\ \mu_2 \end{pmatrix} \tag{12.78}$$

We interpret this as saying that $x', y'$ are the coordinates of the vector $(x, y) \in \mathbb{Z}^2$ relative to the new basis vectors for $\mathbb{Z}^2$.

$$v_1 = \begin{pmatrix} \nu_2 \\ \nu_1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} -\mu_1 \\ \mu_2 \end{pmatrix} \tag{12.79}$$

The good property of this basis is that the smallest multiple of $v_1$ that sits in $\Lambda$ is $\ell v_1$ (prove this) [192] Similarly, the smallest multiple of $v_2$ in $\Lambda$ is $gv_2$. Thus, we have a way of writing $\mathbb{Z}^2$ as $\mathbb{Z}v_1 \oplus \mathbb{Z}v_2$ such that the projection of $\Lambda$ to the $v_1$ axis is the group $\ell\mathbb{Z}$ while the kernel is the subgroup of $\mathbb{Z}v_2$ that maps into $\Lambda$, and that is just $\cong g\mathbb{Z}$.

Put differently, there is a homomorphism $\psi : \mathbb{Z}^2 \to \mathbb{Z}$ that takes

$$\psi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x' \ . \tag{12.80}$$

This is the projection on the $v_1$ axis. This defines a surjective homomorphism onto $\mathbb{Z}$. (Explain why.) On the other hand, using (12.76) and $\mu_1\mu_2 g = \ell$ we see that the image of $\Lambda$ under $\psi$ is $\ell\mathbb{Z}$. Therefore, using the exercise result (7.144) $\psi$ descends to a map

$$\bar{\psi} : \mathbb{Z}^2/\Lambda \to \mathbb{Z}/\ell\mathbb{Z} \tag{12.81}$$

Now note from (12.78) that

$$\begin{pmatrix} -\mu_1 \\ \mu_2 \end{pmatrix} \mathrm{mod}\Lambda \tag{12.82}$$

is in the kernel of $\bar{\psi}$, and moreover it generates a cyclic subgroup of order $g$ in $\mathbb{Z}^2/\Lambda$. By counting, this cyclic subgroup must be the entire kernel of $\bar{\psi}$. Therefore we have an exact sequence

$$0 \to \mathbb{Z}_g \to \mathbb{Z}^2/\Lambda \to \mathbb{Z}_\ell \to 0 \tag{12.83}$$

**************************

AND THERE IS A MAP TO $y'$ AND TOGETHER THESE GIVE ISOMORPHISM TO $\mathbb{Z}/\ell\mathbb{Z} \oplus \mathbb{Z}/g\mathbb{Z}$.

**************************

This conclues our second proof. ♠

♣Should really add a figure to illustrate this. Unfortunately the first really nontrivial case is $m_1 = 2 \cdot 3$ and $m_2 = 2 \cdot 5$ so $\ell = 30$. ♣

---

**Exercise**

Using the Kronecker theorem show that if a finite Abelian group $G$ is <u>not</u> a cyclic group then there is a nontrivial divisor $n$ of $|G|$ so that $g^n = 1$ for all $g \in G$.

---

### 12.3.1 The Chinese Remainder Theorem

In fact, there is an important generalization of this statement known as the *Chinese remainder theorem*:

**Theorem** Suppose $m_1, \dots, m_r$ are pairwise relatively prime positive integers, (i.e. $(m_i, m_j) = 1$ for all $i \neq j$) then

$$(\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z}) \cdots \oplus (\mathbb{Z}/m_r\mathbb{Z}) \cong \mathbb{Z}/M\mathbb{Z} \tag{12.84}$$

where $M = m_1 m_2 \cdots m_r$.

*Proof*:

The fastest proof makes use of the previous result and induction on $r$.

A second proof offers some additional insight into solving simultaneous congruences: We construct a homomorphism

$$\psi : \mathbb{Z} \to (\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z}) \cdots \oplus (\mathbb{Z}/m_r\mathbb{Z}) \tag{12.85}$$

by

$$\psi(x) = (x \text{mod} m_1, x \text{mod} m_2, \dots, x \text{mod} m_r) \tag{12.86}$$

We first claim that $\psi(x)$ is *onto*. That is, for any values $a_1, \dots, a_r$ we can solve the simultaneous congruences:

$$\begin{aligned} x &= a_1 \text{mod} m_1 \\ x &= a_2 \text{mod} m_2 \\ \vdots & \quad \vdots \\ x &= a_r \text{mod} m_r \end{aligned} \tag{12.87}$$

for some common value $x \in \mathbb{Z}$.

[192] *Answer*: We have $xv_1 \in \Lambda$ iff $x\nu_2 = 0\text{mod}(g\mu_1)$ and $x\nu_1 = 0\text{mod}(g\mu_2)$. Multiply these equations by $\mu_2$ and $\mu_1$, respectively, and add them. Find that $x = 0\text{mod}\ell$.

To prove this note that $\hat{m}_i := M/m_i = \prod_{j\neq i} m_j$ is relatively prime to $m_i$ (by the hypothesis of the theorem). Therefore there are integers $x_i, y_i$ such that

$$x_i m_i + y_i \hat{m}_i = 1 \tag{12.88}$$

Let $g_i = y_i \hat{m}_i$. Note that

$$g_i = \delta_{i,j} \text{mod} m_j \qquad \forall 1 \leq i, j \leq r \tag{12.89}$$

Therefore if we set

$$x = \sum_{i=1}^{r} a_i g_i \tag{12.90}$$

then $x$ is a desired solution to (12.87) and hence is a preimage under $\psi$.

On the other hand, the kernel of $\psi$ is clearly $M\mathbb{Z}$. Therefore:

$$0 \to M\mathbb{Z} \to \mathbb{Z} \to (\mathbb{Z}/m_1\mathbb{Z}) \oplus (\mathbb{Z}/m_2\mathbb{Z}) \cdots \oplus (\mathbb{Z}/m_r\mathbb{Z}) \to 0 \tag{12.91}$$

and hence the desired isomorphism follows. ♠

**Remarks**

1. Equation (12.67) is used implicitly all the time in physics, whenever we have two degrees of freedom with different but commensurable frequencies. Indeed, it is used all the time in everyday life. As a simple example, suppose you do X every other day. You will then do X on Mondays every other week, i.e., every 14 days, because 2 and 7 are relatively prime. More generally, consider a system with a discrete configuration space $\mathbb{Z}/p\mathbb{Z}$ thought of as the multiplicative group of $p^{th}$ roots of 1. Suppose the time evolution for $\Delta t = 1$ is $\omega_p^r \to \omega_p^{r+1}$ where $\omega_p$ is a primitive $p^{th}$ root of 1. The basic period is $T = p$. Now, if we have *two* oscillators of periods $p, q$, the configuration space is $\mathbb{Z}_p \times \mathbb{Z}_q$. The basic period of this system is - obviously - the least common multiple of $p$ and $q$. That is the essential content of (**??**).

2. Our second proof shows that in fact equation (12.84) is a statement of an isomorphism of rings.

3. One might wonder how the theorem got this strange name. (Why don't we refer to the "Swiss-German theory of relativity?") The theorem is attributed (see, e.g. Wikipedia) to Sun-tzu Suan-ching in the 3rd century A.D. (He should not be confused with Sun Tzu who lived in the earlier Spring and Autumn period and wrote *The Art of War*.) For an interesting historical commentary see [193] which documents the historical development in India and China up to the definitive treatments by Euler, Lagrange, and Gauss who were probably unaware of previous developments hundreds of years earlier. The original motivation was apparently related to construction of calendars, and this is certainly mentioned by Gauss in his renowned book *Disquisitiones Arithmeticae*. The Chinese calendar is based on *both* the lunar and solar cycles. Roughly speaking, one starts the new year based on both the winter solstice

*and* the new moon. Thus, to find periods of time in this calendar one needs to solve simultaneous congruences. I suspect the name "Chinese Remainder Theorem" is an invention of 19th century mathematicians. Hardy & Wright (1938) do not call it that, but do recognize Sun Tzu.

♣Some students with Chinese background say this is wrong. Check it out. ♣

---

**Exercise** *Counting your troops*

Suppose that you are a general and you need to know how many troops you have from a cohort of several hundred. Time is too short to take attendance.

So, you have your troops line up in rows of 5. You observe that there are 3 left over. Then you have your troops line up in rows of 11. Now there are 2 left over. Finally, you have your troops line up in rows of 13, and there is only one left over.

How many troops are there? [194]

---

---

**Exercise**

a.) Show that the Chinese Remainder theorem is false if the $m_i$ are not pairwise relatively prime.

b.) Show that the obstruction to finding a solution $x$ to $x = a_i \text{mod} m_i$ is given by the reductions $(a_i - a_j)\text{mod}(m_i, m_j)$ over all pairs $i \neq j$. That is, a solution exists iff all of these vanish.

---

## 13. The Group Of Automorphisms

Recall that an *automorphism* of a group $G$ is an isomorphism $\mu : G \to G$, i.e. an isomorphism of $G$ onto itself.

One easily checks that the composition of two automorphisms $\mu_1, \mu_2$ is an automorphism. The identity map is an automorphism, and every automorphism is invertible. In this way, the set of automorphisms, $\text{Aut}(G)$, is *itself a group* with group law given by composition.

Given a group $G$ there are God-given automorphisms given by conjugation. That is, if $a \in G$ then

$$I(a) : g \to aga^{-1} \tag{13.1}$$

[193] Kang Sheng Shen, "Historical development of the Chinese remainder theorem," Arch. Hist. Exact Sci. 38 (1988), no. 4, 285-305.

[194] Apply the Chinese remainder theorem with $m_1 = 5, m_2 = 11, m_3 = 13$. Then $M = 715$, $\hat{m}_1 = 143$, $\hat{m}_2 = 65$ and $\hat{m}_3 = 55$. Using the Euclidean algorithm you find convenient lifts to the integers $g_1 = 286$, $g_2 = -65$ and $g_3 = -220$. Then the number of troops is $3 \times 286 - 2 \times 65 - 1 \times 220 = 508 \text{mod} 715$. Therefore there are 508 soldiers.

defines an automorphism of $G$. Indeed $I(a) \circ I(b) = I(ab)$ and hence $I : G \to \mathrm{Aut}(G)$ is a homomorphism. The subgroup $\mathrm{Inn}(G)$ of such automorphisms is called the group of *inner automorphisms*. Note that if $a \in Z(G)$ then $I(a)$ is trivial, and conversely. Thus we have:

$$\mathrm{Inn}(G) \cong G/Z(G). \tag{13.2}$$

Moreover, $\mathrm{Inn}(G)$ is a normal subgroup of $\mathrm{Aut}(G)$, since for any automorphism $\phi \in \mathrm{Aut}(G)$:

$$\phi \circ I(a) \circ \phi^{-1} = I(\phi(a)). \tag{13.3}$$

Therefore we have another group

$$\mathrm{Out}(G) := \mathrm{Aut}(G)/\mathrm{Inn}(G) \tag{13.4}$$

known as the group of "outer automorphisms." Thus

$$1 \to \mathrm{Inn}(G) \to \mathrm{Aut}(G) \to \mathrm{Out}(G) \to 1 \tag{13.5}$$

Note we can also write and exact sequence of length four:

$$1 \to Z(G) \to G \to \mathrm{Aut}(G) \to \mathrm{Out}(G) \to 1 \tag{13.6}$$

**Remarks**

1. In practice one often reads or hears the statement that an element $\varphi \in \mathrm{Aut}(G)$ is an "outer automorphism." What this means is that it projects to a nontrivial element of $\mathrm{Out}(G)$. However, strictly speaking this is an abuse of terminology and an outer automorphism is in the quotient group (13.4). These notes might sometimes perpetrate this abuse of terminology.

2. Note that for any abelian group $G$ all nontrivial automorphisms are outer automorphisms.

**Example 13.1**: Consider $\mathrm{Aut}(\mathbb{Z}_3)$. This group is Abelian so all automorphisms are outer. Thinking of it multiplicatively, the only nontrivial choice is $\omega \to \omega^{-1}$. If we think of $A_3 \cong \mathrm{Aut}(\mathbb{Z}_3)$ then we are taking

$$(123) \to (132) \tag{13.7}$$

So: $Aut(\mathbb{Z}_3) \cong \mathbb{Z}_2$.

**Example 13.2**: Consider $Aut(\mathbb{Z}_4)$. Think of $\mathbb{Z}_4$ as the group of fourth roots of unity, generated by $\omega = \exp[i\pi/2] = i$. A generator must go to a generator, so there is only one possible nontrivial automorphism: $\phi : \omega \to \omega^3$. Note that $\omega \to \omega^2$ is a nontrivial homomorphism of $\mathbb{Z}_4 \to \mathbb{Z}_4$, but it is not an automorphism. Thus $Aut(\mathbb{Z}_4) \cong \mathbb{Z}_2$.

**Example 13.3**: Consider $Aut(\mathbb{Z}_5)$. Think of $\mathbb{Z}_5$ as the group of fifth roots of unity, generated by $\omega = \exp[2\pi i/5]$. Now there are several automorphisms: $\phi_2$ defined by its action on the generator $\omega \to \omega^2$. Similarly, we can define $\phi_3$, by $\omega \to \omega^3$ and $\phi_4$, by $\omega \to \omega^4$. Letting $\phi_1$ denote the identity we have

$$\phi_2^2 = \phi_4 \qquad \phi_2^3 = \phi_3 \qquad \phi_2^4 = \phi_4^2 = \phi_1 = 1 \tag{13.8}$$

So $Aut(\mathbb{Z}_5) \cong \mathbb{Z}_4$. The explicit isomorphism is

$$\begin{aligned} \phi_2 &\to \bar{1} \\ \phi_4 &\to \bar{2} \\ \phi_3 &\to \bar{3} \end{aligned} \tag{13.9}$$

**Example 13.4**: Consider Aut($\mathbb{Z}_N$), and let us think of $\mathbb{Z}_N$ multiplicatively as the group of $N^{th}$ roots of 1. An automorphism $\phi$ of $\mathbb{Z}_N$ must send $\omega \mapsto \omega^r$ for some $r$. On the other hand, $\omega^r$ must also be a generator of $\mathbb{Z}_N$. Automorphisms must take generators to generators. Hence $r$ is relatively prime to $N$. This is true iff there is an $s$ with

$$rs = 1 \mathrm{mod} N \tag{13.10}$$

*Thus, $Aut(\mathbb{Z}_N)$ is the group of transformations $\omega \to \omega^r$ where $r$ admits a solution to $rs = 1\mathrm{mod}N$.* We will examine this interesting group in a little more detail in §13.1 below.

**Example 13.5**: *Automorphisms Of The Symmetric Group $S_n$*: There are no outer automorphisms of $S_n$ so

$$Aut(S_n) \cong Inn(S_n) \cong S_n, \qquad n \neq 2,6 \tag{13.11}$$

Note the exception: $n = 2, 6$. Note the striking contrast from an abelian group, all of whose automorphisms are outer.

This is not difficult to prove: Note that an automorphism $\phi$ of $S_n$ must take conjugacy classes to conjugacy classes. Therefore we focus on how it acts on transpositions. These are involutions, and involutions must map to involutions so the conjugacy class of transpositions must map to a conjugacy class of the form $(1)^k(2)^\ell$ with $k + 2\ell = n$. We will show below that, just based on the order of the conjugacy class, $\phi$ must map transpositions to transpositions. We claim that any automorphism that maps transpositions to transpositions must be inner. Let us say that

$$\phi((ab)) = (xy) \qquad \phi((ac)) = (zw) \tag{13.12}$$

where $a, b, c$ are all distinct. We claim that $x, y, z, w$ must comprise precisely three distinct letters. We surely can't have $(xy) = (zw)$ because $\phi$ is 1-1, and we also can't have $(xy)$ and $(zw)$ commuting because the group commutator of $(ab)$ and $(ac)$ is $(abc)$. Therefore we can write

$$\phi((ab)) = (xy) \qquad \phi((ac)) = (xz) \tag{13.13}$$

Therefore, we have defined a permutation $a \to x$ and $\phi$ is the inner automorphism associated with this permutation.

Now let us consider the size of the conjugacy classes. This was computed in exercise *** above. The size of the conjugacy class of transpositions is of course

$$\binom{n}{2} = \frac{n!}{(n-2)!2!} \tag{13.14}$$

The size of a conjugacy class of the form $(1)^k(2)^\ell$ with $k + 2\ell = n$ is

$$\frac{n!}{(n-2\ell)!\ell!2^\ell} \tag{13.15}$$

Setting these equal results in the identity

$$\frac{(n-2)!}{(n-2\ell)!} = \ell!2^{\ell-1} \qquad n \geq 2\ell \tag{13.16}$$

For a fixed $\ell$ the LHS is a polynomial in $n$ which is growing for $n \geq 2\ell$ and therefore bounded below by $(2\ell - 2)!$. Therefore we consider whether there can be a solution with $n = 2\ell$:

$$(2\ell - 2)! = \ell!2^{\ell-1} \tag{13.17}$$

For $\ell = 3$, corresponding to $n = 6$, there is a solution, but for $\ell > 3$ we have $(2\ell - 2)! > \ell!2^{\ell-1}$. The peculiar exception $n = 6$ is related to the symmetries of the icosahedron. For more information see

1.http://en.wikipedia.org/wiki/Automorphisms of the symmetric and alternating groups

2. http://www.jstor.org/pss/2321657

3. I.E. Segal, "The automorphisms of the symmetric group," *Bulletin of the American Mathematical Society* **46**(1940) 565.

**Example 13.6**: *Automorphisms Of Alternating Groups.* For the group $A_n \subset S_n$ there is an automorphism which is not obviously inner: Conjugation by any odd permutation. Recall that $Out(G) = Aut(G)/Inn(G)$ is a quotient group so conjugation by any odd permutation represents the same element in $Out(G)$. If we consider $A_3 \subset S_3$ then

$$(12)(123)(12)^{-1} = (132) \tag{13.18}$$

is indeed a nontrivial automorphism of $A_3$ and since $A_3$ is abelian this automorphism must be an outer automorphism. In general conjugation by an odd permutation defines an outer automorphism of $A_n$. For example suppose conjugation by (12) were inner. Then there would be an even permutation $a$ so that conjugation by $a \cdot (12)$ centralizes every $h \in A_n$. But $a \cdot (12)$ together with $A_n$ generates all of $S_n$ and then $a \cdot (12)$ would have to be in the center of $S_n$, a contradiction. Thus, the outer automorphism group of $A_n$ contains a nontrivial involution. Again for $n = 6$ there is an exceptional outer automorphism.

The above example nicely illustrates a general idea: If $N \triangleleft G$ is a normal subgroup of $G$ and $g \notin H$ then conjugation by $g$ defines an automorphism $H \to H$ which is, in general, not an inner automorphism.

**Example 13.7**: Consider $G = GL(n, \mathbb{C})$. Then $A \to A^*$ is an outer automorphism: That is, there is no invertible complex matrix $S \in GL(n, \mathbb{C})$ such that, for every invertible matrix $A \in GL(n, \mathbb{C})$ we have

$$A^* = SAS^{-1} \tag{13.19}$$

---

**Exercise** *Outer Automorphisms Of Some Matrix Groups*

a.) Prove (13.19). [195]

b.) Consider maps of $GL(n,\mathbb{C})$ given by $A \to A^{tr}$, $A \to A^{-1}$ and $A \to A^{tr,-1}$. Which of these are automorphisms? Which of these are outer automorphisms?

c.) Consider $G = SU(2)$. Is $A \to A^*$ an outer automorphism? [196]

d.) Consider the automorphism of $G = SO(2)$

$$R(\phi) \to R(-\phi) \tag{13.20}$$

Is this inner or outer?

---

---

**Exercise** *Automorphisms of* $\mathbb{Z}$

Show that $\mathrm{Aut}(\mathbb{Z}) \cong \mathbb{Z}_2$. [197]

---

---

**Exercise**

Although $\mathbb{Z}_2$ does not have any automorphisms the product group $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ certainly does.

a.) Show that an automorphism of $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ must be of the form

$$\phi(x_1, x_2) = (a_1 x_1 + a_2 x_2, a_3 x_3 + a_4 x_4) \tag{13.21}$$

where we are writing the group additively, and

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \in GL(2, \mathbb{Z}_2) \tag{13.22}$$

b.) Show that $GL(2,\mathbb{Z}_2) \cong S_3$. [198]

---

[195] *Hint*: Consider the invertible matrix $A = \mathrm{i}1_{n\times n}$.

[196] *Answer*: No! Note that $(\mathrm{i}\sigma^k)^* = -\mathrm{i}(\sigma^k)^* = (\mathrm{i}\sigma^2)(\mathrm{i}\sigma^k)(\mathrm{i}\sigma^2)^{-1}$. But $\mathrm{i}\sigma^2 \in SU(2)$ and every $SU(2)$ matrix is a real linear combination of 1 and $\mathrm{i}\sigma^k$. This has an important implication for the representation theory of $SU(2)$: Every irreducible representation is either real or "pseudoreal" (quaternionic).

[197] *Answer*: The most general homomorphism $\mathbb{Z} \to \mathbb{Z}$ is the map $n \mapsto an$ for some integer $a$. But for an automorphism $a$ must be mutliplicatively invertible in the integers. Therefore $a$ is $+1$ or $-1$.

[198] *Hint*: Consider what the group does to the three nontrivial elements $(1,0)$, $(0,1)$, and $(1,1)$. The three transpositions correspond to $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and the two elements of order 3 are $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$.

c.) Now describe $\mathrm{Aut}(\mathbb{Z}_4 \times \mathbb{Z}_4)$. [199]

---

---

**Exercise** *Automorphisms of $\mathbb{Z}_p^N$*

(Warning: This is hard and uses some other ideas from algebra.)

Let $p$ be prime. Describe the automorphisms of $\mathbb{Z}_p^N$, and show that the group has order [200]

$$|\mathrm{Aut}(\mathbb{Z}_p^N)| = (p^N - 1)(p^N - p)(p^N - p^2) \cdots (p^N - p^{N-1}) \tag{13.23}$$

---

---

**Exercise** *Automorphisms Of The Quaternion Group*

Show that the group of automorphisms of the quaternion group $Q = \{\pm 1, \pm \mathfrak{i}, \pm \mathfrak{j}, \pm \mathfrak{k}\}$ is isomorphic to $S_4$. [201]

♣This exercise should go later, perhaps in the section on extensions. Perhaps in the chapter on symmetries of regular objects. ♣

(This assumes you know what the quaternions are. See below for various descriptions of the quaternion group $Q$.)

---

---

**Exercise** *Isomorphisms between two different groups*

Let $G_1, G_2$ be two groups which are isomorphic, but not presented as the same set with the same multiplication table. Let $\mathrm{Isom}(G_1, G_2)$ be the set of all isomorphisms from $G_1 \to G_2$.

Show that

a.) Any two isomorphisms $\Psi, \Psi' \in \mathrm{Isom}(G_1, G_2)$ are related by $\Psi' = \Psi \circ \phi$ where $\phi \in \mathrm{Aut}(G_1)$.

---

[199] *Answer*: This group has a homomorphism onto $\mathrm{Aut}(\mathbb{Z}_2 \times \mathbb{Z}_2)$ with a kernel isomorphic to $\mathbb{Z}_2^4$.

[200] *Answer*: The group is the group of $N \times N$ invertible matrices over the ring $\mathbb{Z}_p$. These are in one-one correspondence with the possible bases for the vector space $\mathbb{Z}_p^N$. How many ordered bases are there? Note that any nonzero vector can serve as the first basis vector, and there are $p^N - 1$ nonzero vectors. Choose one and call it $e_1$. Now, $e_2$ can be any vector not in the linear span of $e_1$. But the linear span of $e_1$ is a one-dimensional subspace of $p$ elements. These are all excluded so $e_2$ must be chosen from a set of $p^N - p$ vectors. Make a choice of $e_2$. Then $e_3$ must be chosen from a vector not in the span of $e_1, e_2$. The span of $e_1, e_2$ consists of $p^2$ vectors so there are $p^N - p^2$ choices for $e_3$, and so on.

[201] *Answer*: First, the group of inner automorphisms is $Q/Z(Q) \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. The three nontrivial elements are given by conjugation by $\mathfrak{i}, \mathfrak{j}, \mathfrak{k}$. Now, any automorphism must permute the three normal subgroups generated by $\mathfrak{i}, \mathfrak{j}, \mathfrak{k}$, and automorphisms leading to nontrivial permutations of normal subgroups must be outer. So the outer automorphism group must be a subgroup of $S_3$. Now, in fact, one can construct such outer automorphisms. In fact, it suffices to say what the image of $\mathfrak{i}$ and $\mathfrak{j}$ are since these generate the whole group. Thus, the automorphism group is an extension of $S_3$ by $\mathbb{Z}_2 \times \mathbb{Z}_2$ and one can then map this isomorphically to $S_4$.

b.) Any two isomorphisms $\Psi, \Psi' \in \mathrm{Isom}(G_1, G_2)$ are related by $\Psi' = \phi \circ \Psi$ where $\phi \in \mathrm{Aut}(G_2)$.

The set $Isom(G_1, G_2)$ with $G_1, G_2$ not equal but isomorphic is a good example of what is called a *torsor*. A *torsor* for a group $G$ is a set $X$ with a free transitive action.

---

### 13.1 The group of units in $\mathbb{Z}_N$

We have seen that $\mathbb{Z}/N\mathbb{Z}$ is a group inherited from the *additive* law on $\mathbb{Z}$. For an integer $n \in \mathbb{Z}$ denote its image in $\mathbb{Z}/N\mathbb{Z}$ by $\bar{n}$. With this notation the group law on $\mathbb{Z}/N\mathbb{Z}$ is

$$\bar{n}_1 + \bar{n}_2 = \overline{n_1 + n_2}, \tag{13.24}$$

and $\bar{0}$ is the unit element.

However, note that since

$$(n_1 + N\ell_1)(n_2 + N\ell_2) = n_1 n_2 + N\ell'' \tag{13.25}$$

we do have a well-defined operation on $\mathbb{Z}/N\mathbb{Z}$ inherited from *multiplcation* in $\mathbb{Z}$:

$$\bar{n}_1 \cdot \bar{n}_2 := \overline{n_1 \cdot n_2}. \tag{13.26}$$

In general, even if we omit $\bar{0}$, $\mathbb{Z}/N\mathbb{Z}$ is *not* a group with respect to the multiplication law (find a counterexample). Nevertheless, $\mathbb{Z}/N\mathbb{Z}$ with $+, \times$ is an interesting object which is an example of something called a *ring*. See the next chapter for a general definition of a ring.

Let us define *the group of units in the ring* $\mathbb{Z}/N\mathbb{Z}$:

$$(\mathbb{Z}/N\mathbb{Z})^* := \{\bar{m} : 1 \le m \le N-1, \gcd(m, N) = 1\} \tag{13.27}$$

where $(m, N)$ is the *greatest common divisor* of $m$ and $N$. We will also denote this group as $\mathbb{Z}_N^*$.

Then, $(\mathbb{Z}/N\mathbb{Z})^*$ *is* a group with the law (13.26) ! Clearly the multiplication is closed and $\bar{1}$ is the unit. The existence of multiplicative inverses follows from (12.8).

Moreover, as we have seen above, we can identify

$$\mathrm{Aut}(\mathbb{Z}/N\mathbb{Z}) \cong (\mathbb{Z}/N\mathbb{Z})^* \tag{13.28}$$

The isomorphism is that $a \in (\mathbb{Z}/N\mathbb{Z})^*$ is mapped to the transformation

$$\psi_a : n \ \mathrm{mod} N \to an \ \mathrm{mod} N \tag{13.29}$$

if we think of $\mathbb{Z}/N\mathbb{Z}$ additively or

$$\psi_a : \omega \to \omega^a \tag{13.30}$$

if we think of it multiplicatively. Note that $\psi_{a_1} \circ \psi_{a_2} = \psi_{a_1 a_2}$.

The order of the group $(\mathbb{Z}/N\mathbb{Z})^*$ is denoted $\phi(N)$ and is called the Euler $\phi$-function or *Euler's totient function.* [202] One can check that

$$\begin{aligned} \phi(2) &= 1 \\ \phi(3) &= 2 \\ \phi(4) &= 2 \end{aligned} \tag{13.31}$$

What can we say about the structure of $\mathbb{Z}_N^*$? Now, in general it is not true that $\mathrm{Aut}(G_1 \times G_2)$ and $\mathrm{Aut}(G_1) \times \mathrm{Aut}(G_2)$ are isomorphic. Counterexamples abound. For example $\mathrm{Aut}(\mathbb{Z}) \cong \mathbb{Z}_2$ but $\mathrm{Aut}(\mathbb{Z} \oplus \mathbb{Z}) \cong GL(2, \mathbb{Z})$. Nevertheless, it actually is true that $\mathrm{Aut}(\mathbb{Z}_n \times \mathbb{Z}_m) \cong \mathrm{Aut}(\mathbb{Z}_n) \times \mathrm{Aut}(\mathbb{Z}_m)$ when $n$ and $m$ are relatively prime. To prove this, let $v_1$ be a generator of $\mathbb{Z}_n$ and $v_2$ a generator of $\mathbb{Z}_m$ and let us write our Abelian group additively. The general endomorphism of $\mathbb{Z}_n \oplus \mathbb{Z}_m$ is of the form

$$\begin{aligned} v_1 &\to \alpha v_1 + \beta v_2 \\ v_2 &\to \gamma v_1 + \delta v_2 \end{aligned} \tag{13.32}$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$. Now impose the conditions $nv_1 = 0$ and $mv_2 = 0$ and the fact that $\bar{n}$ is multiplicatively invertible in $\mathbb{Z}_m$ and $\bar{m}$ is multiplicatively invertible in $\mathbb{Z}_n$ to learn that in fact an endomorphism must have $\beta = 0 \mathrm{mod}\ m$ and $\gamma = 0 \mathrm{mod}\ n$. Therefore $\beta v_2 = 0$ and $\gamma v_1 = 0$. Therefore, an automorphism of $\mathbb{Z}_n \oplus \mathbb{Z}_m$ is determined by $v_1 \to \alpha v_1$ with $\bar{\alpha} \in \mathbb{Z}_n^*$ and $v_2 \to \delta v_2$ with $\bar{\delta} \in \mathbb{Z}_m^*$ and hence $\mathrm{Aut}(\mathbb{Z}_n \oplus \mathbb{Z}_m) \cong \mathrm{Aut}(\mathbb{Z}_n) \times \mathrm{Aut}(\mathbb{Z}_m)$ when $n$ and $m$ are relatively prime. (The corresponding statement is absolutely false when they are not relatively prime.) So we have:

$$\mathbb{Z}_{nm}^* \cong \mathbb{Z}_n^* \times \mathbb{Z}_m^* \tag{13.33}$$

In particular, $\phi$ is a multiplicative function: $\phi(nm) = \phi(n)\phi(m)$ if $(n, m) = 1$. Therefore, if $N = p_1^{e_1} \cdots p_r^{e_r}$ is the decomposition of $N$ into distinct prime powers then

$$(\mathbb{Z}/N\mathbb{Z})^* \cong (\mathbb{Z}/p_1^{e_1}\mathbb{Z})^* \times \cdots (\mathbb{Z}/p_r^{e_r}\mathbb{Z})^* \tag{13.34}$$

Moreover, $(\mathbb{Z}/p^e\mathbb{Z})^*$ is of order $\phi(p^e) = p^e - p^{e-1}$, as is easily shown [203] and hence

$$\phi(N) = \prod_i (p_i^{e_i} - p_i^{e_i - 1}) = N \prod_{p|N} (1 - \frac{1}{p}) \tag{13.35}$$

**Remark**: For later reference in our discussion of cryptography note one consequence of this: If we choose, randomly - i.e. with uniform probability density - a number between 1 and $N$ the probability that it will be relatively prime to $N$ is

$$\frac{\phi(N)}{N} = \prod_{p|N} (1 - \frac{1}{p}) \tag{13.36}$$

[202] Do not confuse $\phi(N)$ with the $\phi_a$ above!

[203] Proof: The numbers between 1 and $p^e$ which have gcd larger than one must be of the form $px$ where $1 \le x \le p^{e-1}$. So the rest are relatively prime.

This means that, if $N$ is huge and a product of just a few primes, then a randomly chosen number will almost certainly be relatively prime to $N$.

In elementary number theory textbooks it is shown that if $p$ is an odd prime then $(\mathbb{Z}/p^e\mathbb{Z})^*$ is a cyclic group.

♣Finish proof for $e > 1$. ♣

To prove this let us begin with $(\mathbb{Z}/p\mathbb{Z})^*$. (This proof uses some ideas from the algebra of fields.) Suppose this group were not cyclic. Then there would be some $n$ which is a nontrivial divisor of the order, $\phi(p) = p - 1$ such that $x^n = 1$ for all $x \in (\mathbb{Z}/p\mathbb{Z})^*$. That would imply that in the field $\mathbb{F}_p$ the equation $x^n - 1$ would have $p - 1$ distinct roots. On the other hand, the equation $x^n - 1$ can have at most $n$ roots, and that is a contradiction. We conclude that, in fact,$(\mathbb{Z}/p\mathbb{Z})^*$ must be cyclic.

Of course, all primes are odd, and two is the oddest prime of all. If $p = 2$ the result is a little different and we have:

$$(\mathbb{Z}/4\mathbb{Z})^* \cong \{\pm 1\} \tag{13.37}$$

is cyclic but

$$(\mathbb{Z}/2^e\mathbb{Z})^* = \{(-1)^a 5^b | a = 0, 1, 0 \le b < 2^{e-2}\} \cong (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2^{e-2}\mathbb{Z}) \tag{13.38}$$

when $e \ge 3$.

In fact, more generally, it turns out that $(\mathbb{Z}/n\mathbb{Z})^*$ is cyclic iff $n \in \{1, 2, 4, p^k, 2p^k\}$ where $p$ runs over odd primes and $k > 0$.

Note that if we take a product of two distinct odd prime powers then

$$(\mathbb{Z}/(p_1^{k_1} p_2^{k_2}\mathbb{Z})^* \cong (\mathbb{Z}/p_1^{k_1}\mathbb{Z})^* \times (\mathbb{Z}/p_2^{k_2}\mathbb{Z})^* \tag{13.39}$$

But $\phi(p_1^{k_1})$ and $\phi(p_2^{k_2})$ are both even, being divisible by $p_1 - 1$ and $p_2 - 1$, respectively, and hence are not relatively prime, and hence $(\mathbb{Z}/(p_1^{k_1} p_2^{k_2}\mathbb{Z})^*$ is not cyclic.

**Examples**

1. $(\mathbb{Z}/7\mathbb{Z})^* = \{1, 2, 3, 4, 5, 6\}\text{mod}7 \cong \mathbb{Z}_6$. Note that 3 and 5 are generators:

$$3^1 = 3, \quad 3^2 = 2, \quad 3^3 = 6, \quad 3^4 = 4, \quad 3^5 = 5, \quad 3^6 = 1 \qquad \text{mod}7 \tag{13.40}$$

$$5^1 = 5, \quad 5^2 = 4, \quad 5^3 = 6, \quad 5^4 = 2, \quad 5^5 = 3, \quad 5^6 = 1 \qquad \text{mod}7 \tag{13.41}$$

   However, $2 = 3^2\text{mod}7$ is *not* a generator, even though it is prime. Rather, it generates an index 2 subgroup $\cong \mathbb{Z}_3$, as does 4, while 6 generates an index 3 subgroup $\cong \mathbb{Z}_2$. Do not confuse this isomorphic copy of $\mathbb{Z}_6$ with the additive presentation $\mathbb{Z}_6 \cong \mathbb{Z}/6\mathbb{Z} = \{0, 1, 2, 3, 4, 5\}$ with the *additive* law. Then 1 and 5 are generators, but not $2, 3, 4$.

2. $(\mathbb{Z}/9\mathbb{Z})^* = \{1, 2, 4, 5, 7, 8\}\text{mod}9 \cong \mathbb{Z}_6$. It is a cyclic group generated by 2 and $2^5 = 5\text{mod}9$, but it is not generated by $2^2 = 4$, $2^3 = 8$ or $2^4 = 7\text{mod}9$, because $2, 3, 4$ are not relatively prime to 6.

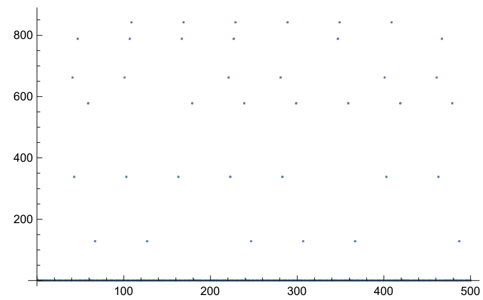



**Figure 37:** A plot of the residues of $2^x\delta_x$ modulo $N = 2\cdot3\cdot5\cdot7\cdot11 = 2310$, for $1 \le x \le 500$. Here $\delta_x = 0$ if $gcd(x, N) > 1$ so that we only see the values in $(\mathbb{Z}/N\mathbb{Z})^*$. Notice the apparently random way in which the value jumps as we increase $x$.

3. $(\mathbb{Z}/8\mathbb{Z})^* = \{1, 3, 5, 7\} \cong \mathbb{Z}_2 \times \mathbb{Z}_2$. Note that $3^2 = 5^2 = 7^2 = 1\text{mod}8$ and $3\cdot5 = 7\text{mod}8$, so we can take 3 and 5 to be the generators of the two $\mathbb{Z}_2$ subgroups.

4. $(\mathbb{Z}/15\mathbb{Z})^* \cong \mathbb{Z}_2 \times \mathbb{Z}_4$ is not cyclic.

**Remarks**

1. When $(\mathbb{Z}/n\mathbb{Z})^*$ is cyclic a generator is called a *primitive root modulo* $n$, and is not to be confused with a primitive $n^{th}$ root of one. It is trivial to find examples of the latter and highly nontrivial to find examples of the former.

2. The values of $f(x) = a^x \text{mod} N$ for $(a, N) = 1$ appear to jump about randomly as a function of $x$, as shown in Figure 37. Therefore, finding the period of this function, that is, the smallest positive integer $r$ so that $f(x + r) = f(x)$ is not easy. This is significant because of the next remark.

3. *Factoring Integers.* Suppose $N$ is a positive integer and $a$ is a positive integer so that $(a, N) = 1$, and the order, denoted $r$, of $\bar{a} \in (\mathbb{Z}/N\mathbb{Z})^*$ is <u>even</u> and finally suppose that $b := a^{r/2} \neq \pm 1 \text{mod} N$. Note that $b^2 = 1 \text{mod} N$ so $\bar{b}$ is a nontrivial squareroot of $\bar{1} \in (\mathbb{Z}/N\mathbb{Z})^*$. Then we claim that $d_\pm := gcd(b \pm 1, N)$ are in fact <u>nontrivial</u> factors of $N$. To see this we need to rule out $d_\pm = 1$ and $d_\pm = N$, the trivial factors of $N$. If we had $d_\pm = N$ then $N$ would divide $b \pm 1$ but that would imply $b = \mp 1 \text{mod} N$, contrary to assumption. Now, suppose $d_\pm = 1$, then by Bezout's theorem there would be integers $\alpha_\pm, \beta_\pm$ so that
$$(b \pm 1)\alpha_\pm + N\beta_\pm = 1 \tag{13.42}$$
But then multiply the equation by $b \mp 1$ to get
$$(b^2 - 1)\alpha_\pm + N\beta_\pm(b \mp 1) = b \mp 1 \tag{13.43}$$
But now, $N$ divides the LHS so $b \mp 1 = 0 \text{mod} N$ which implies $b = \pm 1 \text{mod} N$, again contrary to assumption. Thus, $d_\pm$ are <u>nontrivial</u> divisors of $N$.

   To give a concrete example, take $N = 3 \cdot 5 \cdot 7 = 105$, so $\phi(N) = 48$. Then the period of $f(x) = 2^x$ is $r = 12$, and $b = 2^{12/2} = 64$. Well $gcd(64 + 1, 105) = 5$ and $gcd(64 - 1, 105) = 21$ are both divisors of 105. In fact $105 = 5 \cdot 21$.

4. *Artin's Conjecture*: Finding a generator is not always easy, and it is related to some deep conjectures in number theory. For example, the Artin conjecture on primitive roots states that for any positive integer $a$ which is not a perfect square there are an infinite number of primes so that $\bar{a}$ is a generator of the cyclic group $(\mathbb{Z}/p\mathbb{Z})^*$. In fact, if $a$ is not a power of another integer, and the square-free part of $a$ is not 1mod4 then Artin predicts the density of primes for which $a$ is a generator to be
$$\prod_{Artin\ primes} \left(1 - \frac{1}{p(p-1)}\right) = 0.37.... \tag{13.44}$$
According to the Wikipedia page, there is not a single number $a$ for which the conjecture is known to be true. For example, the primes $p < 500$ for which $a = 2$ is a

generator of $(\mathbb{Z}/p\mathbb{Z})^*$ is

$$\{3,5,11,13,19,29,37,53,59,61,67,83,101,107,131,139,149,163,173,179,181,$$
$$197,211,227,269,293,317,347,349,373,379,389,419,421,443,461,467,491\} \tag{13.45}$$

5. A good reference for this material is Ireland and Rosen, *A Classical Introduction to Modern Number Theory* Springer GTM

---

**Exercise** *Euler's theorem and Fermat's little theorem*

a.) Let $G$ be a finite group of order $n$. Show that if $g \in G$ then $g^n = e$ where $e$ is the identity element.

b.) Prove *Euler's theorem*: For all integers $a$ relatively prime to $N$, $g.c.d(a,N) = 1$,

$$a^{\phi(N)} = 1 \mathrm{mod} N \tag{13.46}$$

Note that a special case of this is Fermat's little theorem: If $a$ is an integer and $p$ is prime then

$$a^p = a \mathrm{mod} p \tag{13.47}$$

**Remark**: This theorem has important practical applications in *prime testing*. If we want to test whether an odd integer $n$ is prime we can compute $2^n \mathrm{mod} n$. If the result is $\neq 2 \mathrm{mod} n$ then we can be sure that $n$ is not prime. Now $2^n \mathrm{mod} n$ can be computed *much* more quickly with a computer than the traditional test of seeing whether the primes up to $\sqrt{n}$ divide $n$. If $2^n \mathrm{mod} n$ is indeed $= 2 \mathrm{mod} n$ then we can suspect that $n$ is prime. Unfortunately, there are composite numbers which will masquerade as primes in this test. They are called "base 2 pseudoprimes." In fact, there are numbers $n$, known as *Carmichael numbers* which satisfy $a^n = a \mathrm{mod} n$ for all integers $a$. The good news is that they are rare. The bad news is that there are infinitely many of them. According to Wikipedia the first few Carmichael numbers are

$$561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, \dots, \tag{13.48}$$

The first Carmichael number is $561 = 3 \cdot 11 \cdot 17$ and Erdös proved that the number $C(X)$ of Carmichael numbers smaller than $X$ is bounded by

$$C(X) < X \exp\left(-\frac{\kappa \log X \log\log\log X}{\log\log X}\right) \tag{13.49}$$

where $\kappa$ is a positive real number.

---

---

**Exercise** *Periodic Functions*

a.) Consider the function

$$f(x) = 2^x \text{mod } N \tag{13.50}$$

for an odd integer $N$. Show that this function is periodic $f(x + r) = f(x)$ for a minimal period $r$ which divides $\phi(N)$.

b.) Compute the period for $N = 15, 21, 105$. [204]

c.) More generally, if $(a, N) = 1$ show that $f(x) = a^x$ mod $N$ is a periodic function.

---

---

**Exercise** *How Many Primitive Roots Of n Are There?*

Show that $n$ has either zero or $\phi(\phi(n))$ different primitive roots.

---

---

## 13.2 Group theory and cryptography

Any invertible map $f : \mathbb{Z}/N\mathbb{Z} \to \mathbb{Z}/N\mathbb{Z}$ can be used to define a code. For example, if $N = 26$ we may identify the elements in $\mathbb{Z}/26\mathbb{Z}$ with the letters in the Latin alphabet:

$$a \leftrightarrow \bar{0}, b \leftrightarrow \bar{1}, c \leftrightarrow \bar{2}, \ldots \tag{13.51}$$

---

**Exercise** *Caesar Shift*

a.) Show that $f(m) = (m - 3)$mod26 defines a code. In fact, the above remark, and this example in particular, is attributed to Julius Caesar. Using this decode the message:

$$ZOLPPQEBORYFZLK! \tag{13.52}$$

b.) Is $f(m) = (3m)$mod26 a valid code? By adding symbols or changing the alphabet we can change the value of $N$ above. Is $f(m) = (3m)$mod27 a valid code?

---

The RSA public key encryption system is a beautiful application of Euler's theorem and works as follows. The basic idea is that with numbers with thousands of digits it is relatively easy to compute powers $a^n$mod$m$ and greatest common divisors, but it is very difficult to factorize such numbers into their prime parts. For example, for a 1000 digit number the brute force method of factorization requires that we sample up to

$$\sqrt{10^{1000}} = 10^{500} \tag{13.53}$$

[204] *Answer*: $r = 4, 6, 12$ divides $\phi(N) = 8, 12, 48$.

divisors. Bear in mind that our universe is about $\pi \times 10^7 \times 13.79 \times 10^9 \cong 4 \times 10^{17}$ seconds old. [205] There are of course more efficient algorithms, but all the publicly known ones are still far too slow.

Now, Alice wishes to receive and decode secret messages sent by any member of the public. She chooses two large primes (thousands of digits long) $p_A, q_A$ and computes $n_A := p_A q_A$. These primes are to be kept secret. How does she find her secret thousand-digit primes? She chooses a random thousand digit number and applies the Fermat primality test. By the prime number theorem she need only make a few thousand attempts, and she will find a prime. [206]

Next, Alice computes $\phi(n_A) = (p_A - 1)(q_A - 1)$, and then she chooses a random thousand-digit number $d_A$ such that $gcd(d_A, \phi(n_A)) = 1$ and computes an inverse $d_A e_A = 1 \text{mod} \phi(n_A)$. All these steps are relatively fast and easy, because Euclid's algorithm is very fast. Thus there is some integer $f$ so that

$$d_A e_A - f\phi(n_A) = 1 \tag{13.54}$$

That is, she solves the congruence $x = 1\text{mod}\phi(n_A)$ and $x = 0\text{mod}d_A$, for the smallest positive $x$ and then computes $e_A = x/d_A$.

Finally, she publishes for the world to see the encoding key: $\{n_A, e_A\}$, but she keeps the numbers $p_A, q_A, \phi(n_A), d_A$ secret. This means that if anybody, say Bob, wants to send Alice a secret message then he can do the following:

Bob converts his plaintext message into a number less than $n_A$ by writing $a \leftrightarrow 01$, $b \leftrightarrow 02$, $\dots$, $z \leftrightarrow 26$. (Thus, when reading a message with an odd number of digits we should add a 0 in front. If the message is long then it should be broken into pieces of length smaller than $n_A$.) Let Bob's plaintext message thus converted be denoted $m$. It is a positive integer smaller than $n_A$.

Now to compute the ciphertext Bob looks up Alice's numbers $\{n_A, e_A\}$ on the public site and uses these to compute the ciphertext:

$$c := m^{e_A} \text{mod} n_A \tag{13.55}$$

Bob sends the ciphertext $c$ to Alice over the internet. Anyone can read it.

Then Alice can decode the message by computing

$$\begin{aligned} c^{d_A} \text{mod} n_A &= m^{e_A d_A} \text{mod} n_A \\ &= m^{1+f\phi(n_A)} \text{mod} n_A \\ &= m \text{mod} n_A \end{aligned} \tag{13.56}$$

Thus, to decode the message Alice just needs one piece of private information, namely the integer $d_A$.

---

[205]There are $\pi \times 10^7$ seconds in a year, to 0.3% accuracy.

[206]The prime number theorem says that if $\pi(x)$ is the number of primes between 1 and $x$ then as $x \to \infty$ we have $\pi(x) \sim \frac{x}{\log x}$. Equivalently, the $n^{th}$ prime is asymptotically like $p_n \sim n \log n$. This means that the density of primes for large $x$ is $\sim 1/\log x$, so if $x \sim 10^n$ then the density is $1/n$ so if we work with thousand-digit primes then after about one thousand random choices we will find a prime.

Now Eve, who has a reputation for making trouble, cannot decode the message without knowing $d_A$. Just knowing $n_A$ and $e_A$ but not the prime factorization $n_A = p_A q_A$ there is no obvious way to find $d_A$. The reason is that even though the number $n_A$ is public it is hard to compute $\phi(n_A)$ without knowing the prime factorization of $n_A$. Of course, if Eve finds out about the prime factorization of $n_A$ then she can compute $\phi(n_A)$ immediately and then quickly (using the Euclidean algorithm) invert $e_A$ to get $d_A$. Thus, the security of the method hinges on the inability of Eve to factor $n_A$ into primes.

In summary,

1. The <u>intended receiver</u> of the message, namely Alice in our discussion, knows

$$(p_A, q_A, n_A = p_A q_A, \phi(n_A) = (p_A - 1)(q_A - 1), e_A, d_A). \tag{13.57}$$

2. Alice publishes $(n_A, e_A)$. Anybody can look these up.
3. The <u>sender</u> of the message, namely Bob in our discussion, takes a secret message $m_B$ and computes the ciphertext $c = m_B^{e_A} \mathrm{mod} n_A$.
4. Alice can decode Bob's message by computing $m_B = c^{d_A} \mathrm{mod} n_A$ using her secret knowledge of $d_A$.
5. The <u>attacker</u>, namely Eve in our discussion, knows $(n_A, e_A, c)$ but will have to work to find $d_A$ or some other way of decoding the ciphertext.

**Remarks**

1. Note that the decoding will *fail* if $m$ and $n_A$ have a common factor. However, $n_A = p_A q_A$ and $p_A, q_A$ are primes with thousands of digits. The probability that Bob's message is one of these is around 1 in $10^{1000}$.

---

**Exercise** *Your turn to play Eve*
Alice has published the key

$$(n = 661643, e = 325993) \tag{13.58}$$

Bob sends her the ciphertext in four batches:

$$c_1 = 541907 \quad c_2 = 153890 \quad c_3 = 59747 \quad c_4 = 640956 \tag{13.59}$$

What is Bob's message? [207]

---

[207] Factor the integer $n = 541 * 1223$. Then you know $p, q$ and hence $\phi(n) = 659880$. Now take $e$ and compute $d$ by using the Chinese Remainder theorem to compute $x = 1 \mathrm{mod} \phi$ and $x = 0 \mathrm{mod} e$. This gives $x = 735766201 = de$ and hence $d = 2257$. Now you can compute the message from the ciphertext $m = c^d \mathrm{mod} n$.

### 13.2.1 How To Break RSA: Period Finding

The attacker, Eve, can read the ciphertext $c\text{mod}n_A$. That means the attacker can try to compute the period of the function

$$f(x) := c^x \text{mod} n_A \tag{13.60}$$

Suppose (as is extremely likely when $n_A$ is a product of two large primes) that $c$ is relatively prime to $n_A$. Then the cyclic group $\langle c \rangle \in (\mathbb{Z}/n_A\mathbb{Z})^*$ generated by $c$ must coincide with the cyclic group generated by the message $m_B$ and in particular they both have the same period $r$, which divides $\phi(n_A)$. Suppose Eve figures out the period $r$. Since the published value $e_A$ is relatively prime to $\phi(n_A)$ it will be relatively prime to $r$ and therefore there exists a new decoding method: Compute $d_E$ such that

$$e_A d_E = 1 \bmod r \tag{13.61}$$

Then

$$c^{d_E} = m^{e_A d_E} \text{mod} n_A = m_B^{1+\ell r} \text{mod} n_A = m_B \text{mod} n_A \tag{13.62}$$

decodes the message.

Thus, if the attacker can find the period of $f(x)$ the message can be decoded.

Another way in which finding the period leads to rapid decoding is through explicit factoring:

We saw in our discussion of $\mathbb{Z}_N^*$ that, if one has an element $\bar{a} \in \mathbb{Z}_N^*$ with even period $r$ and $\bar{b} = \bar{a}^{r/2} \neq \overline{\pm 1}$ then $d_\pm = gcd(b \pm 1, N)$ are nontrivial factors of $N$. Suppose there were a quick method to find the period $r$. Then we could quickly factor $N$ as follows:

1. Choose a random integer $a$ and using Euclid check that $(a, N) = 1$. If $N$ is a product of two large primes you will only need to make a few choices of $a$ before succeeding.

2. Compute the period $r$ of the function $f(x) = a^x \text{mod} N$.

3. If $r$ is odd go back and choose another $a$ until you get one with $r$ even.

4. Then check that $b = a^{r/2} \neq -1 \text{mod} N$. Again this can be done quickly, thanks to Euclid. If you get $b = -1 \text{mod} N$ go back and choose another $a$, until you find one that works. The point is that, with high probability, if you pick $a$ at random you will succeed. So you might have a try a few times, but not many.

So, the only real bottleneck in factoring $N$ is computing the order $r$ of $\bar{a}$ in $\mathbb{Z}_N^*$. Equivalently, this is computing the period of the function $f(x) = a^x \bmod N$ where $(a, N) = 1$. This is where the "quantum Fourier transform" and "phase estimation" come in. Quantum computers give a way to compute this period in polynomial time in $N$, as opposed to classical computers which take exponential time in $N$. We will come back to this.

### 13.2.2 Period Finding With Quantum Mechanics

♣This section is out of place. Goes later in the course ♣

Here we sketch how quantum computation allows one to find the period of the function $f(x) = a^x \text{mod} N$ where $(a, N) = 1$. This is just a sketch. A nice and clear and elementary account (which we used heavily) can be found in D. Mermin's book *Quantum Computer Science* and more details and a more leisurely discussion can be found there.

Quantum computation is based on the action of certain unitary operators on a system of $n$ Qbits, that is, on a Hilbert space

$$\mathcal{H}_n = (\mathbb{C}^2)^{\otimes n} \tag{13.63}$$

equipped with the standard inner product. For each factor $\mathbb{C}^2$ one chooses a basis $\{|0\rangle, |1\rangle\}$, which one should think of as, for example spin up/down eigenstates of an electron or photon helicity polarization states. Then there is a natural basis for $\mathcal{H}_n$:

$$|\vec{x}\rangle := |x_{n-1}\rangle \otimes \cdots \otimes |x_1\rangle \otimes |x_0\rangle \tag{13.64}$$

Here, for each $i$, $x_i \in \{0,1\}$. One can identify the vector $\vec{x} \in \mathbb{F}_2^n$, the $n$-dimensional vector space over the field $\mathbb{F}_2$. In our discussion we will only use its Abelian group structure, so one can also think of it as $\mathbb{Z}_2 \oplus \cdots \oplus \mathbb{Z}_2$ with $n$ summands. The basis of states (13.64) is known as the *computational basis* or, the *Classical basis*. Now to each computational basis vector we can assign an integer by its binary expansion:

$$N(\vec{x}) := 2^{n-1}x_{n-1} + \cdots + 2^2 x_2 + 2^1 x_1 + x_0 \tag{13.65}$$

Now, let $N := 2^n$. We can also define the Hilbert space $L^2(\mathbb{Z}_N)$ of functions on the group $\mathbb{Z}_N$ with the natural Haar measure. Of course $\mathcal{H}_n$ is isomorphic to $L^2(\mathbb{Z}_N)$ and the isomorphism we choose to use is the one which identifies the computational basis vector $|\vec{x}\rangle$ with the delta function supported at $N(\vec{x})\text{mod } N$. We will denote the latter states as $|N(\vec{x})\rangle\rangle_N$, where the subscript indicates which Abelian group $\mathbb{Z}_N$ we are working with.

In quantum computation one works with a Hilbert space decomposed as

$$\mathcal{H} = \mathcal{H}_{input} \otimes \mathcal{H}_{output} \tag{13.66}$$

The two factors have dimension $N_{in} = 2^{n_{in}}$ and $N_{out} = 2^{n_{out}}$, respectively. The quantum gates are unitary operators and, moreover, under identification of $\mathcal{H}$ as a tensor product of Qbits there should be a notion of "locality" in the sense that they only act nontrivially on "a few" adjacent factors. The locality reflects the spatial locality in some realization in the lab in terms of, say, spin systems. Moreover, we should only have to apply " a few" quantum gates in a useful circuit. With an arbitrary number of gates we can construct any unitary out of products of local ones to arbitrary accuracy. The above notions can be made precise, but that is beyond the scope of this section.

♣Still, it is essential to explain more about the notion of "local quantum gate" and quantum circuit and illustrate a few examples of simple gates. ♣

Now suppose we have a function

$$f : (\mathbb{Z}_N)^* \to (\mathbb{Z}_N)^* \tag{13.67}$$

(such as $f(x) = a^x \text{mod} N$ for $(a, N) = 1$). We would like to convert this to a map

$$\check{f} : \mathbb{F}_2^{n_{in}} \to \mathbb{F}_2^{n_{out}} \tag{13.68}$$

We now choose a fundamental domain which is a subset of $\{1, 2, \ldots, N-1\}$ for $(\mathbb{Z}_N)^*$ with $N < N_{out}$ and $N < N_{in}$ (in fact we will eventually assume $N \ll N_{out}$ and $N \ll N_{in}$) so that we can view elements of $(\mathbb{Z}/N\mathbb{Z})^*$ as elements of the set $\{1, 2, \ldots, N-1\}$ which is, in turn, a subset of $\mathbb{Z}/N_{in}\mathbb{Z}$ and $\mathbb{Z}/N_{out}\mathbb{Z}$. We use the function $N(\vec{x})$ above to define $\check{f}$ such that

$$f(N(\vec{x})) = N(\check{f}(\vec{x})) \text{mod} 2^{n_{out}} \tag{13.69}$$

This does not uniquely specify $\check{f}$ but the ambiguity will not affect the discussion. To read this equation, suppose you want to compute $\check{f}(\vec{x})$ for some $\vec{x} \in \mathbb{F}_2^{n_{in}}$. Then you compute $N(\vec{x})$ which is a nonnegative integer between 0 and $N_{in}$. Then you reduce it modulo $N$. If it is relatively prime to $N$ you can compute $f(N(\vec{x}))$ and considerate the result as a number between 1 and $2^{n_{out}} - 1$. The above equation then pins down $\check{f}(\vec{x})$. Using $\check{f}$ we can define a unitary operator $U_f$ by its action on the computational basis:

♣And what if $N(\vec{x})$ is not relatively prime to $N$? Of course, we are thinking this is rare, but it can happen. What is the best way to extend the function? ♣

$$U_f : |\vec{x}\rangle \otimes |\vec{y}\rangle \to |\vec{x}\rangle \otimes |\vec{y} + \check{f}(\vec{x})\rangle \tag{13.70}$$

where on the right-hand side addition is in the Abelian group $(\mathbb{Z}_2)^{n_{out}}$. We will say that the function $f$ *is nice* if $U_f$ can be implemented with a "reasonable" number of local unitary gates. (Of course, one could make this notion much more precise.)

A good example of a local unitary operator on a Qbit is the *Hadamard gate* that acts by

$$\begin{aligned} H : |0\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ H : |1\rangle &\to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \end{aligned} \tag{13.71}$$

This can be summarized by the formula

$$H|y\rangle = \frac{1}{\sqrt{2}} \sum_x (-1)^{xy} |x\rangle \tag{13.72}$$

So,

$$H^{\otimes n}|\vec{y}\rangle = \left(\frac{1}{\sqrt{2}}\right)^n \sum_{\vec{x} \in \mathbb{F}_2^n} (-1)^{\vec{x}\cdot\vec{y}} |\vec{x}\rangle \tag{13.73}$$

and in particular

$$H^{\otimes n}|\vec{0}\rangle = \left(\frac{1}{\sqrt{2}}\right)^n \sum_{\vec{x} \in \mathbb{F}_2^n} |\vec{x}\rangle \tag{13.74}$$

Therefore, (recall that double brackets $|j\rangle\rangle$ refer to the position basis in $L^2(\mathbb{Z}_N)$) we have

$$U_f \left(H^{\otimes n_{in}} \otimes 1\right) : |0\rangle_{in} \times |0\rangle_{out} \to \left(\frac{1}{\sqrt{2}}\right)^{n_{in}} \sum_{j \in \mathbb{Z}_{N_{in}}} |j\rangle\rangle_{N_{in}} \otimes |f(j)\rangle\rangle_{N_{out}} \tag{13.75}$$

Now we wish to apply this to the function $f(x) = a^x \bmod N$ with $(a, N) = 1$ and $N$ is the number we would like to factorize. So, in particular we don't want $N = 2^n$ for some $n$. Nobody will be impressed if you can factor a power of two! Rather, we identify the group $\mathbb{Z}_N^*$, as a set, with the integers $\{1, \dots, N-1\}$ so that it can be considered to be a subset of the natural fundamental domain $\{0, \dots, N_{in} - 1\}$ for $\mathbb{Z}_{N_{in}}$, and similarly for

$\mathbb{Z}_{N_{out}}$. Then to compute $f$ we compute $a^x$, take the residue modulo $N$ to get an integer in the fundamental domain, and then consider that number modulo $N_{out}$. Hence, we should choose $N_{out} = 2^{n_{out}}$ to be some integer larger than $N$. A key claim, explained in textbooks on quantum information theory is that such a function $f$ is nice. That is, it makes sense to compute it with a quantum circuit.

So, we conclude that a suitable quantum circuit can implement:

$$\begin{aligned} U_f\left(H^{\otimes n} \otimes 1\right) : |0\rangle_{N_{in}} \times |0\rangle_{N_{out}} \to & \left(\frac{1}{\sqrt{2}}\right)^{n_{in}} \sum_{k \in \mathbb{Z}_{N_{in}}} |k\rangle\rangle_{N_{in}} \otimes |a^k\rangle\rangle_{N_{out}} \\ & = \left(\frac{1}{\sqrt{2}}\right)^{n_{in}} \sum_{f_0 \in \mathbb{Z}_N^*}{}' \left(|j_0\rangle\rangle_{N_{in}} + |j_0 + r\rangle\rangle_{N_{in}} + |j_0 + 2r\rangle\rangle_{N_{in}} + \cdots\right) \otimes |f_0\rangle\rangle_{N_{out}} \end{aligned} \tag{13.76}$$

In the second line we are considering $\mathbb{Z}_N^*$ as a subset of $\mathbb{Z}_{N_{out}}$ as explained above and the prime on the sum means that we are just summing over the values that are in the image of $f(x) = a^x \mathrm{mod}\, N$. This will be all the values in $\mathbb{Z}_N^*$ if $a$ is a generator of $\mathbb{Z}_N^*$ but in general might be smaller. Also, $j_0$ is some solution of $f_0 = a^{j_0} \bmod N$, and $r$ is the period of $f(x)$, that is, the smallest positive integer so that $f(x+r) = f(x)$ for all $x$. We can choose $j_0$ so that $0 \le j_0 < r$ and write the RHS of (13.76) as

$$\Psi = \left(\frac{1}{\sqrt{2}}\right)^{n_{in}} \sum_{0 \le j_0 < r} \left(\sum_{s=0}^{\mathcal{O}-1} |j_0 + sr\rangle\rangle_{N_{in}}\right) \otimes |a^{j_0}\rangle\rangle_{N_{out}} \tag{13.77}$$

Here $\mathcal{O}$ is the smallest integer such that $j_0 + \mathcal{O}r \ge N_{in}$. So

$$\mathcal{O} = \lfloor \frac{N_{in} - j_0}{r} \rfloor \tag{13.78}$$

In the applications we have in mind $N_{in}$ and $r$ are typically very large numbers so that this is a (very weak) function of $j_0$. At the end of this section we will use the observation that in this case, $r\mathcal{O}/N_{in}$ is, to very good accuracy, just equal to 1.

Now we measure the output system and get some result, say, $f_0 = a^{k_0} \mathrm{mod} N$. Applying the usual Born rule we get the state for the input system

$$P_{f_0}(\Psi) = \frac{1}{\sqrt{\mathcal{O}}} \sum_{s=0}^{\mathcal{O}-1} |k_0 + sr\rangle\rangle_{N_{in}} \tag{13.79}$$

It is some kind of plane wave state in $L^2(\mathbb{Z}_N)$, so measuring position will give no useful information on $r$. Of course, we should therefore go to the Fourier dual basis to learn about the period. In terms of the position basis of $L^2(\mathbb{Z}_N)$ we can apply $V_{FT}$ to get:

$$V_{FT} P_{f_0} \Psi = \frac{1}{\sqrt{N_{in}\mathcal{O}}} \sum_{p \in \mathbb{Z}_{N_{in}}} \sum_{s=0}^{\mathcal{O}-1} e^{2\pi \mathrm{i} \frac{(k_0 + sr)p}{N_{in}}} |p\rangle\rangle_{N_{in}} \tag{13.80}$$

Note that this Fourier transform is quite nontrivial and nontransparent in the computational basis because of the nontrivial isomorphism between $\mathcal{H}_n$ and $L^2(\mathbb{Z}_N)$ with $N = 2^n$.

Nevertheless, and this is nontrivial and part of the magic of Shor's algorithm, the Fourier transform operator $V_{FT}$ can be implemented nicely with quantum gates in the computational basis. Again, the textbooks on quantum information theory give explicit construction of $V_{FT}$ as a quantum circuit in the computational basis. It is exactly at this point that the exponential speed-up of the period finding takes place:

1. <u>Classical Fourier Transform</u>: $N^2 = 2^{2n}$ operations. We learn every Fourier coefficient.

2. <u>Fast Fourier Transform</u>: $nN = n2^n$ operations. We learn every Fourier coefficient.

3. <u>Quantum Fourier Transform</u>: $n^2$ quantum gates. We only learn about correlations of the output state.

Now to find the period we make a measurement of the amplitudes for the various Fourier components $|p\rangle\rangle_{N_{in}}$. (Here we are using the isomorphism between $\mathbb{Z}_{N_{in}}$ and its unitary dual.) The probability to measure $p$ is

$$\begin{aligned} Prob(p) &= \frac{1}{N_{in}\mathcal{O}} |\sum_{s=0}^{\mathcal{O}-1} (e^{2\pi i \frac{p}{N_{in}/r}})^s|^2 \\ &= \frac{1}{N_{in}\mathcal{O}} \frac{\sin^2\left(\pi \cdot p \cdot \frac{\mathcal{O}r}{N_{in}}\right)}{\sin^2\left(\pi \cdot p \cdot \frac{r}{N_{in}}\right)} \end{aligned} \tag{13.81}$$

The basic idea is that the probability, as a function of $p$, will be peaked near values of $p$ from which we can deduce the crucial number $r$, but we need to be a bit careful at this point.

Let us ask what is the probability that we will measure a value $p$ of the form

$$p_j = j\frac{N_{in}}{r} + \delta_j \qquad\qquad |\delta_j| \leq 1/2 \tag{13.82}$$

If $p_j$ is of this form with any value $j = 1, \dots, r-1$ then we can extract $r$. Thus, substituting such a value for $p_j$ into the formula for the probability we have

$$\frac{1}{N_{in}\mathcal{O}} \frac{\sin^2\left(\pi \cdot \delta_j \cdot \frac{\mathcal{O}r}{N_{in}}\right)}{\sin^2\left(\pi \cdot \delta_j \cdot \frac{r}{N_{in}}\right)} \tag{13.83}$$

Now recall that $\mathcal{O}r/N_{in}$ is equal to 1 to excellent accuracy. Suppose we also choose a number of Qbits so that

$$N_{in} \gg N > r \tag{13.84}$$

Then the argument of the sign in the denominator is extremely small and we can replace $\sin(x)$ by $x$. So we get:

$$\begin{aligned} Prob(p_j) &\cong \frac{1}{N_{in}\mathcal{O}} \frac{\sin^2(\pi\delta_j)}{\left(\pi \cdot \delta_j \cdot \frac{r}{N_{in}}\right)^2} \\ &= \frac{N_{in}}{r\mathcal{O}} \cdot \frac{1}{r} \left(\frac{\sin \pi\delta_j}{\pi\delta_j}\right)^2 \\ &\cong \frac{1}{r}\left(\frac{\sin \pi\delta_j}{\pi\delta_j}\right)^2 \end{aligned} \tag{13.85}$$

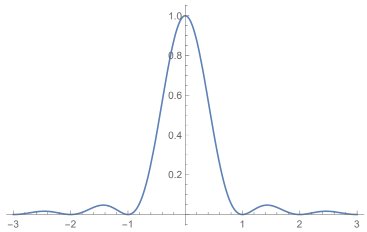



**Figure 38:** A plot of the function $\sin^2(\pi x)/(\pi x)^2$ as a function of $x$. This function, very familiar from the theory of diffraction, is symmetric in $x \to -x$ and monotonically decreasing in the interval $0 \le x \le 1/2$. It therefore takes its minimal value in this interval at $x = 1/2$ where it is about $\cong 0.405$.

Now for $0 \le \delta \le \pi/2$ we have $\frac{\sin x}{x} \ge \frac{2}{\pi}$ so

$$\sum_j Prob(p_j) \ge \frac{r-1}{r}\frac{4}{\pi^2} \cong 0.4 \tag{13.86}$$

Now, using various tricks one can raise this 40% value to near 100%. For these tricks consult Mermin's book. Two other useful textbooks on quantum information theory and quantum computing where one can look these things up include:

1. Nielsen and Chuang, Quantum Information Theory
2. A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, Grad Studies in Math 47, AMS
3. J. Preskill, lecture notes at http://www.theory.caltech.edu/∼preskill/ph219/ph219-2019-20

## 14. Semidirect Products

We have seen a few examples of direct products of groups above. We now study a more subtle notion, the semidirect product. The semidirect product is a twisted version of the direct product of groups $H$ and $G$ which can be defined once we are given one new piece of extra data. The new piece of data we need is a homomorphism

$$\alpha : G \to \mathrm{Aut}(H). \tag{14.1}$$

For an element $g \in G$ we will denote the corresponding automorphism by $\alpha_g$. The value of $\alpha_g$ on an element $h \in H$ is denoted $\alpha_g(h)$. Thus $\alpha_g(h_1 h_2) = \alpha_g(h_1)\alpha_g(h_2)$ because $\alpha_g$ is a homomorphism of $H$ to itself while we also have $\alpha_{g_1 g_2}(h) = \alpha_{g_1}(\alpha_{g_2}(h))$ because $\alpha$ is a homomorphism of $G$ into the group of automorphisms $\mathrm{Aut}(H)$. We also have that $\alpha_1$ is the identity automorphism. (Prove this!)

Using the extra data given by $\alpha$ we can form a more subtle kind of product called the **semidirect product** $H \rtimes G$, or $H \rtimes_\alpha G$ when we wish to stress the role of $\alpha$. In the math literature on group theory the notation $H : G$ is also used. This group is the Cartesian product $H \times G$ as a *set* but has the "twisted" multiplication law:

$$(h_1, g_1) \cdot (h_2, g_2) := (h_1 \alpha_{g_1}(h_2), g_1 g_2) \tag{14.2}$$

A good intuition to have is that "as $g_1$ moves from left to right across the $h_2$ they interact via the action of $g_1$ on $h_2$."

---

**Exercise** *Due diligence*

a.) Show that (14.2) defines an associative group law.

b.) Show that $(1_H, 1_G)$ defines the unit and

$$(h, g)^{-1} = \left(\alpha_{g^{-1}}(h^{-1}), g^{-1}\right) \tag{14.3}$$

c.) Compute the group commutator $[(h_1, g_1), (h_2, g_2)]$ for a semidirect product. [208]

d.) Let $\mathrm{End}(H)$ be the set of all homomorphisms $H \to H$. Note that this set is closed under the operation of composition, and this operation is associative, but $\mathrm{End}(H)$ is not a group because some homomorphisms will not be invertible. Nevertheless, it is a monoid. Show that if $\alpha_g : G \to \mathrm{End}(H)$ is a homomorphism of monoids then (14.2) still defines a monoid. When is it a group?

---

**Example 14.1**: *Infinite dihedral group.* Let $G = \{e, \sigma\} \cong \mathbb{Z}_2$ with generator $\sigma$, and let $H = \mathbb{Z}$, written additively. Then define a nontrivial $\alpha : G \to Aut(H)$ by letting $\alpha_\sigma$ act on $x \in H$ as $\alpha_\sigma(x) = -x$. Then $\mathbb{Z} \rtimes \mathbb{Z}_2$ is a group with elements $(x, e)$ and $(x, \sigma)$, for $x \in \mathbb{Z}$. Note the multiplication laws:

$$\begin{aligned} (x_1, e)(x_2, e) &= (x_1 + x_2, e) \\ (x_1, e)(x_2, \sigma) &= (x_1 + x_2, \sigma) \\ (x_2, \sigma)(x_1, e) &= (x_2 - x_1, \sigma) \\ (x_1, \sigma)(x_2, \sigma) &= (x_1 - x_2, e) \end{aligned} \tag{14.4}$$

and hence the resulting group is nonabelian with this twisted multiplication law. Since $Aut(\mathbb{Z}) \cong \mathbb{Z}_2$ this is the only nontrivial semidirect product we can form. This group is known as the *infinite dihedral group* sometimes denoted $D_\infty$. It has a presentation:

$$\mathbb{Z} \rtimes \mathbb{Z}_2 \cong \langle r, s | s^2 = 1 \qquad srs = r^{-1} \rangle \tag{14.5}$$

[208] *Answer*: $[(h_1, g_1), (h_2, g_2)] = \left(h_1 \alpha_{g_1}(h_2) \alpha_{g_1 g_2 g_1^{-1}}(h_1^{-1}) \alpha_{g_1 g_2 g_1^{-1} g_2^{-1}}(h_2^{-1}), g_1 g_2 g_1^{-1} g_2^{-1}\right)$.

(e.g. take $s = (0, \sigma)$ and $r = (1, e)$)

**Remark**: Taking $x = s$ and $y = rs$ we see that $D_\infty$ also has a presentation as a Coxeter group:

$$\mathbb{Z} \rtimes \mathbb{Z}_2 \cong \langle x, y | x^2 = 1 \qquad y^2 = 1 \rangle \tag{14.6}$$

Indeed it is the Weyl group for the affine Lie group $LSU(2)$.

**Example 14.2**: *Finite dihedral group.* We can use the same formulae as in Example 1, retaining $G = \{e, \sigma\} \cong \mathbb{Z}_2$ but now we take $H = \mathbb{Z}/N\mathbb{Z}$. We still have

$$\alpha_\sigma : \bar{n} \to -\bar{n} \tag{14.7}$$

where we are writing $\mathbb{Z}/N\mathbb{Z}$ additively. The semi-direct product of $\mathbb{Z}/N\mathbb{Z}$ with $\mathbb{Z}_2$ using this automorphism gives one definition of an important group, the *finite dihedral group* $D_N$. Observe that, when we write $\mathbb{Z}_N = \mu_N$ multiplicatively, the automorphism is $\alpha_\sigma(\omega^j) = \omega^{-j}$. In this way we can obtain a presentation of $D_N$ of the form:

$$\mathbb{Z}_N \rtimes_\alpha \mathbb{Z}_2 \cong \langle r, s | s^2 = 1, \qquad srs = r^{-1}, \qquad r^N = 1 \rangle \tag{14.8}$$

Note that here we have switched to a multiplicative model for the group $\mathbb{Z}_N$. The group has the order given by the cardinality of the Cartesian product $\mathbb{Z}_N \times \mathbb{Z}_2$ so it has order $2N$. Note that using the relations, every word in the $r, s$ can be converted to the form $r^x$ or $r^x s$ with $0 \le x \le N-1$, thus accounting for all $2N$ elements.

$$|D_N| = 2N. \tag{14.9}$$

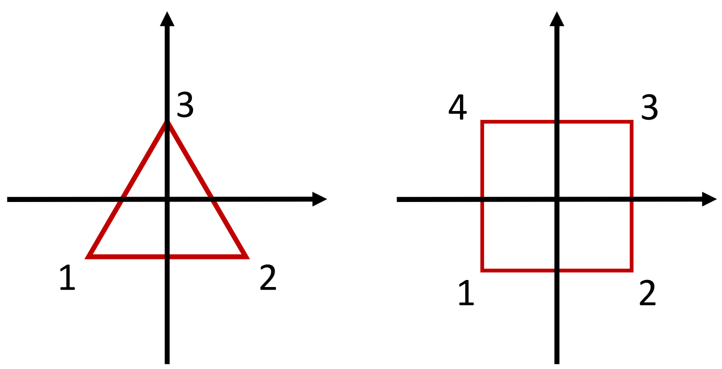



**Figure 39:** A regular 3-gon and 4-gon in the plane, centered at the origin. The subgroup of $O(2)$ that preserves these is $D_3$ and $D_4$, respectively.

**Important Remark**: *The Dihedral Groups And Symmetries Of Regular Polygons.* The group $D_N$ has a natural action on the vector space $\mathbb{R}^2$. The generator $r$ acts by a rotation around the origin: $\phi_r = R(2\pi/N)$. This generates a group isomorphic to $\mathbb{Z}_N$ and in this context it is usually denoted $C_N$. If $P$ is any reflection through a line through the origin then $\phi_s = P$ will satisfy all the relations. The resulting group of transformations of the plane generated by $\phi_r$ and $\phi_s$ is isomorphic to $D_N$. Moreover, if we consider the regular $N$-gon centered at the origin of the plane $\mathbb{R}^2$ then the subgroup of $O(2)$ that maps it to itself is isomorphic to $D_N$, although to preserve the polygon we must choose $P$ carefully so that it is a reflection through an axis of symmetry. For example, if we consider the regular triangle illustrated in 39 then reflection in the $y$-axis is a symmetry, as is rotation by integral amounts of $2\pi/3$. So we have a two-dimensional matrix representation of $D_3$:

$$\begin{aligned} s \to P &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\ r \to R(2\pi/3) &= \begin{pmatrix} \cos(2\pi/3) & \sin(2\pi/3) \\ -\sin(2\pi/3) & \cos(2\pi/3) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix} \end{aligned} \tag{14.10}$$

Note that, if we label the vertices of the triangle $1, 2, 3$ as shown in the figure then the various symmetries are in 1-1 correspondence with permutations of $\{1, 2, 3\}$. So in fact, we have an isomorphism

$$D_3 \cong S_3 \tag{14.11}$$

with $P \to (12)$ and $R(2\pi/3) \to (123)$. Similarly, $D_4$ is the group of symmetries of the square. We can take $s \to P$ as before and now $r \to R(2\pi/4)$. Again we can label the vertices of the square $1, 2, 3, 4$. Again transformations are uniquely determined by permutations of $\{1, 2, 3, 4\}$. However, the group of permutations we get this way is only a subgroup of the permutation group $S_4$.

Clearly there is something quite different about the groups $D_N$ when $N$ is even and odd. This is nicely seen in the set of conjugacy classes. As you show in the exercise below the conjugacy classes in $D_N$ are of the form

$$C(r^x) = \{r^x, r^{-x}\} \tag{14.12}$$

and

$$C(r^x s) = \{r^{x+2y} s | y \in \mathbb{Z}\} \tag{14.13}$$

Now, thinking in terms of symmetry actions on the plane, $r^x$ correspond to rotations by $2\pi x/N$, whereas $r^x s$ correspond to reflections. Now note that for $N$ odd, since $(2, N) = 1$ the conjugacy class $C(r^x s)$ will contain all the elements of the form $r^z s$, in other words all the reflections. However, if $N$ is even then there are two distinct conjugacy classes: $C(r^x s)$ for $x$ even and odd are distinct. This is nicely in accord with symmetries of the $N$-gon: For $N$ odd it is clear that all the symmetry axes can be mapped to each other by using the symmetries of the $N$-gon. Whereas for $N$ even there are two distinct kinds of reflection axes: Those that go through vertices and those that go through edges.

---

**Exercise** $D_4$ *and permutations*

Show that under the map to $S_4$ described above the dihedral group $D_4$ maps to the subgroup containing just the identity, the cyclic permtuations (the rotations)

$$(1234), (1234)^2 = (13)(24), (1234)^3 = (1432) \tag{14.14}$$

and four reflections:

$$(12)(34), \quad (23)(14), \quad (13), \quad , (24) \tag{14.15}$$

---

---

**Exercise**
Show that $D_N$ is a quotient of the infinite dihedral group. [209]

---

**Exercise** *Conjugacy Classes In $D_N$*
a.) Using the presentation of $D_N$ in terms of generators $r, s$ and relations $s^2 = 1$ and $srs = r^{-1}$, and $r^N = 1$ show that we have conjugacy classes given by (14.12) and (14.13).
b.) List the distinct conjugacy classes in $D_N$ for $N$ even and $N$ odd.

♣Should provide answer in a footnote here. ♣

---

**Example 14.3**: In equations (3.20) and (3.49) we found the most general form of a matrix in $O(2)$. It is a disjoint union of two circles, each circle being the elements of determinant $\det(A) = \pm 1$. As a group we have an isomorphism

$$O(2) \cong SO(2) \rtimes \mathbb{Z}_2 \tag{14.17}$$

In fact, there are many such isomorphisms.
We can write an explicit isomorphism as follows: Let $\mathbb{Z}_2 = \{1, \sigma\}$. Then

$$\alpha : \mathbb{Z}_2 \to \mathrm{Aut}(SO(2)) \tag{14.18}$$

is given by

$$\alpha(\sigma) : R(\phi) \to R(-\phi) \tag{14.19}$$

Now choose any nonzero vector $v \in \mathbb{R}^2$ and define an isomorphism

$$\Psi_v : SO(2) \rtimes \mathbb{Z}_2 \to O(2) \tag{14.20}$$

by

$$\begin{aligned} \Psi_v &: (1, \sigma) \to P_v \\ \Psi_v &: (R(\phi), 1) \to R(\phi) \end{aligned} \tag{14.21}$$

---

[209] *Answer*: Note that $\mathcal{N} = \{(x, e) | x = 0 \mathrm{mod} N\} \subset \mathbb{Z} \rtimes \mathbb{Z}_2$ is a normal subgroup and

$$(\mathbb{Z}/N\mathbb{Z}) \rtimes \mathbb{Z}_2 \cong (\mathbb{Z} \rtimes \mathbb{Z}_2)/\mathcal{N}. \tag{14.16}$$

Here $P_v$ is the reflection in the line orthogonal to $v$. We need to check that this is a well-defined homomorphism by checking that the images we have specified are indeed compatible with the relations in the semidirect product. This amounts to checking that $P_v^2 = 1$, which is obvious, and

$$P_v R(\phi) P_v^{-1} = R(-\phi) \tag{14.22}$$

Thanks to (14.22) the relations are indeed satisfied and now it is an easy matter to check that $\Psi_v$ is injective and surjective, so it is an isomorphism.

Note that there are many different such isomorphisms, $\Psi_v$, depending on the choice of $v$. If $v'$ is another nonzero vector in the plane then recall from (10.11) that $P_v P_{v'} = R(2\theta)$ where $\theta$ is the angle between $v$ and $v'$. Then

$$\begin{aligned} \Psi_{v'} : (1, \sigma) \mapsto P_{v'} & \\ &= (P'_v P_v) P_v \\ &= R(2\theta) P_v \\ &= R(\theta) P_v R(-\theta) \end{aligned} \tag{14.23}$$

So

$$\Psi_{v'} = I(R(\theta)) \circ \Psi_{v'} \tag{14.24}$$

and changing $v \to v'$ changes $\Psi_v$ by composition with an inner automorphism of $O(2)$.

More generally, it is true that:

1. When $d$ is odd
$$O(d) \cong SO(d) \times \mathbb{Z}_2 \tag{14.25}$$
2. When $d$ is even
$$O(d) \cong SO(d) \rtimes \mathbb{Z}_2 \tag{14.26}$$

In the case when $d$ is odd the element $-\mathbf{1}_{d\times d}$ has determinant $-1$, so it is in the nontrivial component of $O(d)$, and yet it is also central: so conjugating elements $R \in SO(d)$ acts trivially. The semi-direct product is isomorphic to a direct product in this case. On the other hand, when $d$ is even $-\mathbf{1}_{d\times d}$ has determinant $+1$ and is an element of $SO(d)$. However, it is still true that if $\Psi_v$ is reflection in the hyperplane orthogonal to a nonzero vector $v$ then

$$\alpha_v(R) := P_v R P_v \tag{14.27}$$

is a nontrivial automorphism of $SO(d)$ and we have a family of isomorphisms:

$$\Psi_v : SO(d) \rtimes_{\alpha_v} \mathbb{Z}_2 \to O(d) \tag{14.28}$$

The same discussion as above shows that the dependence on $v$ is through composition with an inner automorphism of $SO(d)$.

**Example 14.4**: *Affine Euclidean Space.* Recall the discussion of Affine Euclidean space in section ******** above. We defined there the Euclidean group.

Some natural examples of isometries are the following: Given any vector $v \in \mathbb{R}^d$ we have the translation operator

$$T_v : p \to p' \tag{14.29}$$

Moreover, if $R \in O(d)$ then, if we choose a point $p$, we can define an operation:

$$R_p : p + v \to p + Rv \tag{14.30}$$

It turns out (this is a nontrivial theorem) that all isometries are obtained by composing such transforamtions. A simple way to express the general transformation, then, is to choose a point $p$ as the "origin" of the affine space thus giving an identification $\mathbb{E}^d \cong \mathbb{R}^d$. Then, to a pair $R \in O(d)$ and $v \in \mathbb{R}^d$ we can associate the isometry: [210]

$$\{v|R\} : x \mapsto v + Rx \qquad\qquad \forall x \in \mathbb{R}^d \tag{14.31}$$

In this notation the group multiplication law is

$$\{v_1|R_1\}\{v_2|R_2\} = \{v_1 + R_1 v_2|R_1 R_2\} \tag{14.32}$$

which makes clear that there is a nontrivial automorphism used to construct the semidirect product of the group of translations, isomorphic to $\mathbb{R}^d$ with the rotation-inversion group $O(d)$.

Put differently: $O(d)$ acts as an automorphism group of $\mathbb{R}^d$:

$$\alpha_R : v \mapsto Rv \tag{14.33}$$

so we can form the abstract group $\mathbb{R}^d \rtimes_\alpha O(d)$. Then, once we choose an origin $p \in \mathbb{E}^d$ we can write an isomorphism

$$\Psi_p : \mathbb{R}^d \rtimes_\alpha O(d) \to \mathrm{Euc}(d) \tag{14.34}$$

To be concrete:

$$\Psi_p(v, R) : p + x \mapsto p + (v + Rx) \tag{14.35}$$

As in our example of $O(d)$ we now have a family of isomorphisms. If $\Psi_{p'}$ is another such based on a different origin $p' = p + v_0$ then the two isomorphisms are related by a translation. See the exercise below.

---

**Exercise** *Internal Definition Of Semidirect Products*

Suppose there is a homomorphism $\alpha : G \to \mathrm{Aut}(H)$ so that we can form the semidirect product $H \rtimes_\alpha G$.

a.) Show that elements of the form $(1, g)$, $g \in G$ form a subgroup $Q \subset H \rtimes G$ isomorphic to $G$, while elements of the form $(h, 1)$, $h \in H$ constitute another subgroup, call it $N$, which is isomorphic to $H$.

[210] Our notation is logically superior to the standard notation in the condensed matter physics literature where it is known as the Seitz notation. In the cond-matt literature we have $\{R|v\} : x \mapsto Rx + v$.

b.) Show that $N = \{(h,1)|h \in H\}$ is a *normal* subgroup of $H \rtimes G$, while $Q = \{(1,g)|g \in G\}$ in general is not a normal subgroup. [211] This explains the funny product symbol $\rtimes$ that looks like a fish: it is a combination of $\times$ with the normal subgroup symbol $\triangleleft$.

c.) Show that we have a short exact sequence:

$$1 \to N \to H \rtimes_\alpha G \to Q \to 1 \tag{14.37}$$

d.) Show that $H \rtimes G = NQ = QN$ and show that $Q \cap N = \{1\}$. Here $NQ$ means the sent of elements $nq$ where $n \in N$ and $q \in Q$. [212]

e.) Conversely, show that if $\tilde{G} = NQ$ where $N$ is a normal subgroup of $\tilde{G}$ and $Q$ is a subgroup of $\tilde{G}$, (that is, every element of $\tilde{G}$ can be written in the form $g = nq$ with $n \in N$ and $q \in Q$ and $N \cap Q = \{1\}$ ) then $\tilde{G}$ is a semidirect product of $N$ and $Q$. Show how to recover the action of $Q$ as a group of automorphisms of $N$ by defining $\alpha_q(n) := qnq^{-1}$. Note that $\alpha_q$ in general is *NOT* an inner automorphism of $N$.

---

---

**Exercise** *When is a semidirect product actually a direct product?*

Show that if $G = NQ$ is a semidirect product and $Q$ is *also* a normal subgroup of $G$, then $G$ is the direct product of $N$ and $Q$. [213]

---

It is useful to think about the Euclidean group in terms of the "internal" characterization of semi-direct products explained in the exercise above. Here we have a normal subgroup $N := \{\{v|1\}|v \in \mathbb{R}^d\}$ and a subgroup $Q$ given by the set of elements of the form $\{0|R\}$. To check that $N$ is normal a short computation using the group law reveals

$$\{v|R\}\{w|1\} = \{Rw|1\}\{v|R\} \tag{14.38}$$

and hence:

$$\{v|R\}\{w|1\}\{v|R\}^{-1} = \{Rw|1\} \tag{14.39}$$

Note that, again, thanks to the group law, $\pi : \{v|R\} \to R$ is a surjective homomorphism $\mathrm{Euc}(d) \to O(d)$. Thus there is an exact sequence:

$$0 \to \mathbb{R}^d \to Euc(d) \to O(d) \to 1 \tag{14.40}$$

[211] Answer to (b): Compute $(h_1,g_1)(h,1)(h_1,g_1)^{-1} = (h_1\alpha_{g_1}(h)h_1^{-1},1)$ and

$$(h_1,g_1)(1,g)(h_1,g_1)^{-1} = (h_1\alpha_{g_1gg_1^{-1}}(h_1^{-1}),g_1gg_1^{-1}). \tag{14.36}$$

[212] The notation is slightly dangerous here: We are considering the group $Q$ both as a subgroup of $G$ and, in equation (14.37), as a quotient. In general, as we will see below in the chapter on exact sequences, there is no way to view a quotient of a group $G$ as a subgroup of $G$. Failure to appreciate this point has led to many, many, many errors in the physics literature.

[213] *Answer:* Note that $n_1q_1n_2q_2 = n_1n_2(n_2^{-1}q_1n_2q_1^{-1})q_1q_2$. However, if both $N$ and $Q$ are normal subgroups then $(n_2^{-1}q_1n_2q_1^{-1}) \in N \cap Q = \{1\}$. Therefore $n_1q_1n_2q_2 = n_1n_2q_1q_2$ is the direct product structure.

Almost identical considerations show that the Poincaré group is isomorphic to the semidirect product of the translation and Lorentz groups:

$$\text{Poincare}(\mathbb{M}^{1,d-1})) \cong \mathbb{M}^{1,d-1} \rtimes O(1,d-1) \tag{14.41}$$

where, once again, the choice of isomorphism depends on a choice of origin.

**Example 14.5**: *Wreath Products.* If $X$ and $Y$ are sets then let $\mathcal{F}[X \to Y]$ be the set of functions from $X$ to $Y$. Recall that

1. If $Y = G_1$ is a group then $\mathcal{F}[X \to G_1]$ is itself a group.
2. If a group $G_2$ acts on $X$ and $Y$ is any set then $G_2$ actions on the function space $\mathcal{F}[X \to Y]$ in a natural way.

We can combine these two ideas as follows: Suppose that $G_2$ acts on a set $X$ and $Y = G_1$ is itself a group. Then let

$$\alpha : G_2 \to \text{Aut}(\mathcal{F}[X \to G_1]) \tag{14.42}$$

be the canonical $G_2$ action on the function space: so if $\phi : G_2 \times X \to X$ is the action on $X$ (part of our given data) then the induced action on the function space is

$$\alpha_g : F \mapsto \alpha_g(F) \in \mathcal{F}[X \to G_1] \tag{14.43}$$

where we define

$$\alpha_g(F)(x) = F(\phi(g^{-1}, x)) \qquad \forall g \in G_2, x \in X \tag{14.44}$$

Then we can form the semidirect product

$$\mathcal{F}[X \to G_1] \rtimes G_2 \tag{14.45}$$

This is a *generalized wreath product.* The traditional wreath product is a special case where $G_2 = S_n$ for some $n$ and $S_n$ acts on $X = \{1, \dots, n\}$ by permutations in the standard way. Note that the group $\mathcal{F}[X \to G_1] \cong G_1^n$. The traditional wreath product $G_1 \text{wr} S_n$, also denoted $G_1 \wr S_n$, is then $\mathcal{F}[X \to G_1] \rtimes S_n$. To be quite explicit, the group elements in $G_1 \wr S_n$ are

$$(h_1, \dots, h_n; \phi) \tag{14.46}$$

with $h_i \in G_1$ and $\phi \in S_n$ and the product is

$$(h_1, \dots, h_n; \phi)(h'_1, \dots, h'_n; \phi') = (h_1 h'_{\phi^{-1}(1)}, h_2 h'_{\phi^{-1}(2)}, \dots, h_n h'_{\phi^{-1}(n)}, \phi \circ \phi') \tag{14.47}$$

**Example 14.6**: *Kaluza-Klein theory.* The basic idea of Kaluza-Klein theory is that we study physics on a product manifold $\mathcal{X} \times \mathcal{Y}$ and partially rigidify the situation by putting some structure on $\mathcal{Y}$. We then regard $\mathcal{Y}$ as "small" and study the physics as "effectively" taking place on $\mathcal{X}$.

The idea is intuitively understood by imagining a $2+1$ dimensional world where space is a cylinder of radius $R$. If we imagine beings in this flatland of a fixed lengthscale, and shrink $R \to 0$ then the beings will end up perceiving themselves as living in a $1+1$ dimensional world.

♣SUITABLE FIGURE NEEDED HERE ♣

Historically, Kaluza-Klein theory arose from attempts to unify the field theories of general relativity with Maxwell's theory of electromagnetism. The basic idea is that pure general relativity on $\mathcal{X} \times \mathcal{Y}$ appears, when $\mathcal{Y}$ is "small" to be a theory of several fields, including electro-magnetism, in $\mathcal{X}$. As originally conceived the idea is very beautiful, but now regarded as too naive and simplistic. Nevertheless, the idea that there might be extra dimensions of spacetime in a compact manifold survives to this day and models that make use of it come astonishingly close to describing the standard model of particle physics and gravity, in the context of "string compactification."

The canonical example of Kaluza-Klein theory is the case where $\mathcal{X} = \mathbb{M}^{1,d-1}$ is $d$-dimensional Minkowski space and $\mathcal{Y} = S^1$ is the circle. We rigidify the situation by putting a metric on the circle $S^1$ so that the metric on space-time is

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu + R^2 (d\theta)^2 \tag{14.48}$$

where $R$ is the radius of the circle, $\theta \sim \theta + 2\pi$ and $0 \le \mu \le d-1$. Our signature is mostly plus. Now we consider a <u>massless</u> scalar field in $(d+1)$ dimensions on this spacetime. A massless scalar field would satisfy the wave equation:

$$\left[ \eta^{\mu\nu} \frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x^\nu} + \frac{1}{R^2} \left( \frac{\partial}{\partial \theta} \right)^2 \right] \phi = 0 \tag{14.49}$$

In Quantum Field Theory one makes a huge leap: The quantization of the field leads to quantum states which are interpreted as the states of a system of <u>particles</u>. An essential step in this feat of magic is that one makes a Fourier-decomposition of the field. The Fourier modes of the field are interpreted as creation/annihilation operators of particle states. For the massless scalar field the Fourier modes are

$$e^{\mathrm{i} p_M x^M} = e^{\mathrm{i} p_\mu x^\mu} e^{\mathrm{i} p_\theta \theta} \tag{14.50}$$

corresponding to single-particle creating and annihilation operators of definite energy-momentum. But since $\theta \sim \theta + 2\pi$ single-valuedness of the field implies that $p_\theta = n \in \mathbb{Z}$ is an integer. But now the wave-equation implies that we have a dispersion relation:

$$E^2 - \vec{p}^2 = \frac{n^2}{R^2} \tag{14.51}$$

where $p_\mu = (E, \vec{p})$. From the viewpoint of a $d$-dimensional field theory, Fourier modes with $n \neq 0$ describe <u>massive</u> particles with $m^2 = n^2/R^2$.

Now consider the case that $R$ is very small compared to the scale of any observer. Then that observer will perceive only a $d$-dimensional spacetime. If $R$ is very small the single massless particle in $d+1$-dimensions is percieved as an infinite set of different massive particles with mass $|n|/R$ in $d$ dimensions. As $R \to 0$ the masses of the particles $\sim |n|/R$ goes to infinity. So, except for the $n = 0$ modes, the particles are very massive and therefore will not be created by low energy processes, and are hence in general unobservable. For example, if $R$ is on the order of the Planck scale then the nonzero Fourier modes are fields that represent particles of Planck-scale mass.

In a similar spirit, one finds that the Einstein-Hilbert action in $(d+1)$ dimensions describing gravity in $(d+1)$ dimensions is equivalent, upon keeping only the $n=0$ Fourier modes, to the action of $d$-dimensional general relativity together with the Maxwell action and the action for a scalar field. In a little more detail, suppose that $\mathcal{Y} = S^1$ and we use coordinates $X^M$, $M = 0, \dots, d+1$ on $\mathcal{X} \times S^1$ and coordinates $x^\mu$ with $\mu = 0, \dots, d$ on $\mathcal{X}$. So that $X^M = (x^\mu, \theta)$ where $\theta$ is an angular coordinate around $S^1$. Then we consider the metric:

$$ds^2 = g_{MN} dX^M dX^N = g_{\mu\nu}(x) dx^\mu dx^\nu + \Omega^2(x)(d\theta + A_\mu(x) dx^\mu)^2 \tag{14.52}$$

where the metric $g_{\mu\nu}$, the "warp factor" $\Omega^2$ and the "gauge connection" $A_\mu$ are only functions of $x^\mu$ (that is, we make the restriction to zero Fourier modes).

Note that this means the metric tensor looks like

$$g_{MN}(x^\mu, \theta) = \begin{pmatrix} g_{\mu\nu}(x) + \Omega^2(x) A_\mu(x) A_\nu(x) & \Omega^2(x) A_\mu(x) \\ \Omega^2(x) A_\nu(x) & \Omega^2(x) \end{pmatrix} \tag{14.53}$$

The general symmetric $(d+1) \times (d+1)$ matrix has

$$\frac{1}{2}(d+1)(d+2) = \frac{1}{2}d(d+1) + d + 1 \tag{14.54}$$

independent components, so we have not lost any generality in the form of the matrix, but writing it this way makes the connection to physical quantities (and to connections on a principal $U(1)$ bundle) clearer. By writing the fields on the RHS as functions of $x^\mu$ and not $(x^\mu, \theta)$ we have made a severe restriction - limiting attention to the massless modes in $d$ dimensions, as explained above.

Under these conditions one computes that the Riemann scalar for the $(d+1)$-dimensional metric is:

$$\mathcal{R}[g_{MN}] = \mathcal{R}[g_{\mu\nu}] - \frac{\Omega^2}{4} F_{\mu\nu} F^{\mu\nu} - 2(\nabla \log \Omega)^2 - 2\nabla^2 \log \Omega \tag{14.55}$$

so the Einstein-Hilbert action for GR in $(d+1)$ dimensions reduces to that of Einstein-Hilbert-Maxwell-Scalar in $d$ dimensions. This is a truly remarkable equation: Pure gravity in $(d+1)$ dimensions leads to both gravity and electricity and magnetism in $d$ dimensions!

**Remarks**:

1. The KK ansatz also leads to a scalar field $\log\Omega^2(x)$, known as the "dilaton" because it can dilate, in a space-time dependent way, the size of the "internal dimensions" $\mathcal{Y}$. Note that in electricity and magnetism the coupling constant enters via

$$S_{\text{Maxwell}} = \int \sqrt{g} \frac{1}{4e^2} F_{\mu\nu} F^{\mu\nu} \tag{14.56}$$

so the presence of the dilaton can lead to space-time variation of the fine structure constant. In physically viable models one must explain why the dilaton does not fluctuate wildly. The discovery of the naturally occurring nuclear reactor in Oklo, Africa, has led to the bound

$$|\frac{\dot{\alpha}}{\alpha}| < few \times 10^{-17} yr^{-1} \tag{14.57}$$

2. By considering internal spaces $\mathcal{Y}$ equipped with metric with isometry group $H$ similar considerations lead gauge theory with gauge group $H$ in $\mathcal{X}$.

It is interesting to understand how gauge symmetries in theories on $\mathcal{X}$ arise in this point of view. Suppose $\mathcal{D} \cong Diff(\mathcal{X})$ is a subgroup of diffeomorphisms of $\mathcal{X} \times \mathcal{Y}$ of the form

$$\psi_f : (x, y) \to (f(x), y) \qquad\qquad f \in Diff(\mathcal{X}). \tag{14.58}$$

We also consider a subgroup $\mathcal{G}$ of $Diff(\mathcal{X} \times \mathcal{Y})$ where $\mathcal{G}$ is isomorphic to a subgroup of $Map(\mathcal{X}, Diff(\mathcal{Y}))$. For the moment just take $\mathcal{G} = Map(\mathcal{X}, Diff(\mathcal{Y}))$, so an element $g \in \mathcal{G}$ is a family of diffeomorphisms of $\mathcal{Y}$ parametrized by $\mathcal{X}$: For each $x$ we have a diffeomorphism of $\mathcal{Y}$: $g_x : y \to g(y; x)$. Then we take $\mathcal{G}$ to be the subgroup of diffeomorphisms of $Diff(\mathcal{X} \times \mathcal{Y})$ of the form

$$\psi_g : (x, y) \to (x, g(y; x)) \qquad\qquad g \in \mathcal{G} \tag{14.59}$$

Note that within $Diff(\mathcal{X} \times \mathcal{Y})$ we can write the subgroup

$$\mathcal{G}\mathcal{D} \tag{14.60}$$

and $\mathcal{D}$ acts as a group of automorphisms of $\mathcal{G}$ via

$$\begin{aligned} \psi_f \psi_g \psi_f^{-1} : (x, y) &\to (f^{-1}(x), y) \\ &\to (f^{-1}(x), g(y; f^{-1}(x))) \\ &\to (x, g(y; f^{-1}(x))) \end{aligned} \tag{14.61}$$

so if $g \in \mathcal{G}$ and $f \in \mathcal{D}$ then $\psi_f \psi_g \psi_f^{-1} = \psi_{g'}$ with $g' \in \mathcal{G}$ and hence $\mathcal{G}\mathcal{D}$ is a semidirect product. In fact, it is an example of the generalized wreath product of the previous example.

This is a model for the group of gauge transformations in Kaluza-Klein theory. So $\mathcal{X}$ is the "large", possibly noncompact, spacetime where we have general relativity, while $\mathcal{Y}$ is the "small," possibly compact space giving rise to gauge symmetry. $\mathcal{D}$ is the diffeomorphism group of the large spacetime and is the gauge symmetry of general relativity on $\mathcal{X}$. Typically, $\mathcal{Y}$ is endowed with a fixed metric $ds^2_{\mathcal{Y}}$ and the diffeomorphism symmetry of $\mathcal{Y}$ is (spontaneously) broken down to the group of isometries of $\mathcal{Y}$, $Isom(\mathcal{Y}, ds^2_{\mathcal{Y}})$. So in the above construction we take $\mathcal{G}$ to be the unbroken subgroup $Map(\mathcal{X}, Isom(\mathcal{Y}, ds^2_{\mathcal{Y}})) \subset Map(\mathcal{X}, Diff(\mathcal{Y}))$. This subgroup $Map(\mathcal{X}, Isom(\mathcal{Y}, ds^2_{\mathcal{Y}}))$ is interpreted as a group of gauge transformations of a gauge theory on $\mathcal{X}$ coupled to general relativity on $\mathcal{X}$.

As a simple example of the remarks of the previous paragraph let us suppose that $\mathcal{Y} = S^1$ with coordinate $\theta$ and round metric $(d\theta)^2$. The isometries of the circle are just constant translations, $\theta \to \theta + \epsilon$. So if

$$g \in \mathrm{Map}(\mathcal{X}, Isom(S^1)) \tag{14.62}$$

then $g(x)$ will take $\theta \to \theta + \epsilon(x)$, so

$$\psi_g : (x^\mu, \theta) \to (x^\mu, \theta + \epsilon(x)) \tag{14.63}$$

so the metric in (14.52) transforms as

$$\psi_g^*(ds^2) = g_{\mu\nu}(x)dx^\mu dx^\nu + \Omega^2(x)(d\theta + dx^\mu \partial_\mu \epsilon + A_\mu(x)dx^\mu)^2 \tag{14.64}$$

meaning that the fields $g_{\mu\nu}$ and $\Omega$ are invariant, but

$$A_\mu \to A_\mu + \partial_\mu \epsilon \tag{14.65}$$

and thus, these special diffeomorphisms appear as gauge transformations of the Maxwell field!

---

**Exercise**

Show that $S_3 \cong \mathbb{Z}_3 \rtimes \mathbb{Z}_2$ where the generator of $\mathbb{Z}_2$ acts as the nontrivial outer automorphism of $\mathbb{Z}_3$.

---

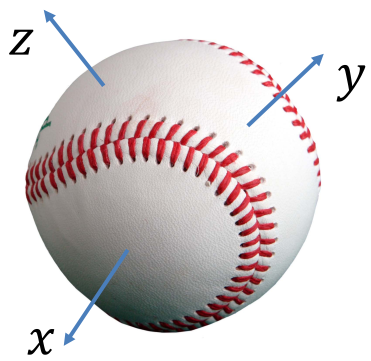



**Figure 40:** A baseball.

---

**Exercise** *Symmetries Of The Baseball*

Ignoring any writing, but taking into account the seams, find the symmetry group of a baseball. (See figure 40.) [214]

---

[214] *Answer* $D_4$.

**Exercise** *Symmetries Of The Cube*

a.) Consider a perfect cube. By considering the action of proper rotations on the four diagonal axes through vertices show that the symmetry group is isomorphic to $S_4$.

b.) Centering the cube on the origin with vertices $(\pm1, \pm1, \pm1)$ show that the symmetry group is $(\mathbb{Z}_2 \times \mathbb{Z}_2) \rtimes S_3$. Deduce that

$$S_4 \cong (\mathbb{Z}_2 \times \mathbb{Z}_2) \rtimes S_3 \tag{14.66}$$

---

---

**Exercise** *Centralizers in the symmetric group*

a.) Suppose that $g \in S_n$ has a conjugacy class given by $\prod_{i=1}^{n}(i)^{\ell_i}$. Show that the centralizer $Z(g)$ is isomorphic to

$$Z(g) \cong \prod_{i=1}^{n} \left(\mathbb{Z}_i^{\ell_i} \rtimes S_{\ell_i}\right) \tag{14.67}$$

where $\prod_i$ is a direct product.

b.) Use this to compute the order of a conjugacy class in the symmetric group, equation (7.159).

---

---

**Exercise** *The Lorentz Group As A Semidirect Product*

Let $\eta = Diag\{-1, \mathbf{1}_d\}$. The Lorentz group in $d+1$ dimensions, denoted $O(1,d)$ is the matrix group

$$O(1,d) = \{A | A\eta A^{tr} = \eta\} \tag{14.68}$$

a.) Considering the case $d = 1$ find the general solution and show that there are four connected components. [215]

b.) Show that, group-theoretically, we have

$$O(1,1) = SO_0(1,1) \rtimes (\mathbb{Z}_2 \times \mathbb{Z}_2) \tag{14.70}$$

where $SO_0(1,1)$ is the connected component of the identity.

**Remark**: In fact, more generally $O(1,d)$ has four connected components for $d \geq 1$ and

$$O(1,d) = SO_0(1,d) \rtimes (\mathbb{Z}_2 \times \mathbb{Z}_2) \tag{14.71}$$

[215] *Answer*: Writing out the four matrix elements of the defining equation easily arrives at the general solution:

$$A = \begin{pmatrix} \xi_1 \cosh\theta & \xi_2 \sinh\theta \\ \xi_4 \sinh\theta & \xi_3 \cosh\theta \end{pmatrix} \tag{14.69}$$

where $\theta \in \mathbb{R}$, $\xi_i \in \{\pm1\}$ and $\xi_4 = \xi_1\xi_2\xi_3$. By changing the sign of $\theta$ we can set $\xi_=2$. The sign of $\xi_1, \xi_2$ is meaningful, giving us four components.

You can easily see that there are at least four components since $\det A \in \{\pm 1\}$ and moreover $A_{00}^2 = 1 + \sum_{i=1}^d A_{0i}^2$ so that we can independently have $A_{00} \geq 1$ or $A_{00} \leq -1$.

---

---

**Exercise** *Holomorph*

Given a finite group $G$ a canonical semidirect product group is $G \rtimes \mathrm{Aut}(G)$ known as the holomorph of $G$.

a.) Show that this is the normalizer of the copy of $G$ in the symmetric group $S_{|G|}$ given by Cayley's theorem.

b.) Show that the affine Euclidean group $Euc(d)$ is the holomorph of the Abelian group $\mathbb{R}^d$.

---

---

**Exercise** *Equivalence of semidirect products*

A nontrivial homomorphism $\alpha : G \to \mathrm{Aut}(H)$ can lead to a semidirect product which is in fact isomorphic to a direct product. Show this as follows: Suppose $\phi : G \to H$ is a homomorphism. Define $\alpha : G \to \mathrm{Aut}(H)$ by $\alpha_g = I(\phi(g))$. Construct an isomorphism [216]

$$\Psi : H \rtimes_\alpha G \to H \times G \tag{14.72}$$

---

---

**Exercise** *Manipulating the Seitz notation*

a.) Show that:

$$\begin{aligned}
\{v|R\}^{-1} &= \{-R^{-1}v|R^{-1}\} \\
\{0|R\}\{v|1\} &= \{Rv|R\} \\
\{v|1\}\{0|R\} &= \{v|R\} \\
\{w|1\}\{v|R\} &= \{w+v|R\} \\
\{v_1|R_1\}\{v_2|R_2\}\{v_1|R_1\}^{-1} &= \{R_1 v_2 + (1 - R_1 R_2 R_1^{-1})v_1 | R_1 R_2 R_1^{-1}\} \\
[\{v_1|R_1\},\{v_2|R_2\}] &= \{(1 - R_1 R_2 R_1^{-1})v_1 - R_1 R_2 R_1^{-1} R_2^{-1}(1 - R_2 R_1 R_2^{-1})v_2 | R_1 R_2 R_1^{-1} R_2^{-1}\}
\end{aligned} \tag{14.73}$$

b.) Show that the subgroup of pure translations, that is, the subgroup of elements of the form $\{v|1\}$ with $v \in \mathbb{R}^d$ is a normal subgroup of Euc$(d)$.

---

[216] *Answer*: $\Psi(h,g) = (h\phi(g), g)$.

c.) Can you construct a homomorphism $O(d) \to \mathrm{Euc}(d)$?

d.) Consider the group $G = L \rtimes \mathbb{Z}_2$ where $L$ is a lattice and the nontrivial element in $\mathbb{Z}_2$ acts on $L$ by $v \to -v$. Compute the conjugacy classes in $G$. [217]

---

---

**Exercise** *Dependence on basepoint for isomorphism of Euclidean group*

Show that if $p' = p + v_0$ then

$$\Psi_{p'}(v', R) = \Psi_p(v, R) \in \mathrm{Euc}(d) \tag{14.74}$$

for

$$v' = v + (1 - R)v_0 \tag{14.75}$$

---

# 15. Group Extensions and Group Cohomology

## 15.1 Group Extensions

♣Add: Pushforward extensions ♣

Recall that an extension of $Q$ by a group $N$ is an exact sequence of the form:

$$1 \to N \stackrel{\iota}{\to} G \stackrel{\pi}{\to} Q \to 1 \tag{15.1}$$

There is a notion of *homomorphism of two group extensions*

$$1 \to N \stackrel{\iota_1}{\to} G_1 \stackrel{\pi_1}{\to} Q \to 1 \tag{15.2}$$

$$1 \to N \stackrel{\iota_2}{\to} G_2 \stackrel{\pi_2}{\to} Q \to 1 \tag{15.3}$$

This means that there is a group homomorphism $\varphi : G_1 \to G_2$ so that the following diagram commutes:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & N & \xrightarrow{\iota_1} & G_1 & \xrightarrow{\pi_1} & Q & \longrightarrow & 1 \\ & & \updownarrow \mathrm{Id} & & \downarrow \varphi & & \updownarrow \mathrm{Id} & & \\ 1 & \longrightarrow & N & \xrightarrow{\iota_2} & G_2 & \xrightarrow{\pi_2} & Q & \longrightarrow & 1 \end{array} \tag{15.4}$$

To say that a "diagram commutes" means that if one follows the maps around two paths with the same beginning and ending points then the compositions of the maps are the same. Thus (15.4) is completely equivalent to the pair of equations:

$$\begin{aligned} \pi_1 &= \pi_2 \circ \varphi \\ \iota_2 &= \varphi \circ \iota_1 \end{aligned} \tag{15.5}$$

[217] *Answer*: $C(\{v|1\}) = \{\{\pm v|1\}\}$ has two elements while $C(\{v|-1\}) = \{\{\pm v + 2v_1| - 1\}|v_1 \in L\}$ has infinitely many elements.

However, drawing a diagram makes the relations between maps, domains and codomains much more transparent. Sometimes a picture is worth a thousand equations. This is why mathematicians like commutative diagrams.

When there is a homomorphism of group extensions based on $\psi : G_2 \to G_1$ such that $\varphi \circ \psi$ and $\psi \circ \varphi$ are the identity then the group extensions are said to be isomorphic.

It can certainly happen that there is more than one nonisomorphic extension of $Q$ by $N$. Classifying all extensions of $Q$ by $N$ is a difficult problem. We will discuss it more in section 15.7 below.

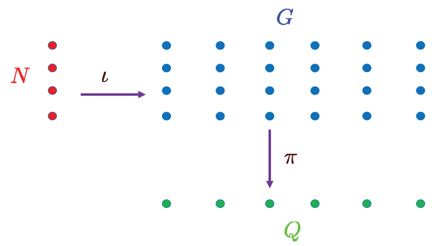



**Figure 41:** Illustration of a group extension $1 \to N \to G \to Q \to 1$ as an $N$-bundle over $Q$.

We would encourage the reader to think geometrically about this problem, even in the case when $Q$ and $N$ are finite groups, as in Figure 41. In particular we will use the important notion of a *section*, that is, a right-inverse to $\pi$: It is a map $s : Q \to G$ such that $\pi(s(q)) = q$ for all $q \in Q$. Such sections always exist.[218] Note that in general $s(\pi(g)) \neq g$. This is obvious from Figure 41. The set of pre-images, $\pi^{-1}(q)$, is called *the fiber of* $\pi$ *over* $q$. The map $\pi$ projects the entire fiber over $q$ to the single element $q$. A choice of section $s$ is a choice, for each and every $q \in Q$, of just one single point in the fiber above $q$.

In order to justify the picture of Figure 41 let us prove that, as a set, $G$ is just the product $N \times Q$. Note that for any $g \in G$ and any section $s$:

$$g(s(\pi(g)))^{-1} \tag{15.6}$$

maps to 1 under $\pi$ (check this). Therefore, since the sequence is exact

$$g(s(\pi(g)))^{-1} = \iota(n) \tag{15.7}$$

for some $n \in N$. That is, every $g \in G$ can be written as

$$g = \iota(n)s(q) \tag{15.8}$$

for some $n \in N$ and some $q \in Q$. In fact, this decomposition is *unique*: Suppose that:

$$\iota(n_1)s(q_1) = \iota(n_2)s(q_2) \tag{15.9}$$

[218]By the axiom of choice. For continuous groups such as Lie groups there might or might not be continuous sections.

Then we rewrite this as

$$\iota(n_2^{-1}n_1) = s(q_2)s(q_1)^{-1} \tag{15.10}$$

Now, applying $\pi$ we learn that $1 = q_2\pi(s(q_1)^{-1}) = q_2\left(\pi(s(q_1))\right)^{-1} = q_2q_1^{-1}$, so $q_1 = q_2$. But that implies $n_1 = n_2$. Therefore, as a set, $G$ can be identified with $N \times Q$.

**Remark**: As a nice corollary of the decomposition (15.8) note that if $\varphi$ defines a morphism of group extensions then $\varphi$ is in fact an isomorphism of $G_1$ to $G_2$. It is a homomorphism by definition. Now note that if $s_1 : Q \to G_1$ is a section of $\pi_1$ then $s_2 := \varphi \circ s_1 : Q \to G_2$ is a section of $\pi_2$ so

$$\begin{aligned}\varphi(g) &= \varphi(\iota_1(n)s_1(q)) \\ &= \varphi(\iota_1(n))\varphi(s_1(q)) \\ &= \iota_2(n)s_2(q)\end{aligned} \tag{15.11}$$

and since the decomposition is unique (given a choice of section) the map $\varphi$ is $1-1$.

Now, given an extension and a choice of section $s$ we define a map

$$\omega : Q \to \mathrm{Aut}(N) \tag{15.12}$$

denoted by

$$q \mapsto \omega_q \tag{15.13}$$

where the definition of $\omega_q$ is given by

$$\iota(\omega_q(n)) = s(q)\iota(n)s(q)^{-1} \tag{15.14}$$

Because $\iota(N)$ is normal the RHS is again in $\iota(N)$. Because $\iota$ is injective $\omega_q(n)$ is well-defined. Moreover, for each $q$ the reader should check that indeed $\omega_q(n_1n_2) = \omega_q(n_1)\omega_q(n_2)$, and $\omega_q$ is one-one on $N$. Therefore we really have a map of sets (15.12). Note carefully that we are not saying that $q \mapsto \omega_q$ is a group homomorphism. In general, it is not.

**Remark**: Clearly the $\iota$ is a bit of a nuisance and leads to clutter and it can be safely dropped if we consider $N$ simply to be a subgroup of $G$, for then $\iota$ is simply the inclusion map. The confident reader is encouraged to do this. The formulae will be a little cleaner. However, we will be pedantic and retain the $\iota$ in most of our formulae.

Let us stress that the map $\omega : Q \to \mathrm{Aut(N)}$ *in general is not a homomorphism* and *in general depends on the choice of section s*. We will discuss the dependence on the choice of section $s$ below when we have some more machinery and context. For now let us see how close $\omega$ comes to being a group homomorphism:

$$\begin{aligned}\iota\left(\omega_{q_1} \circ \omega_{q_2}(n)\right) &= s(q_1)\iota(\omega_{q_2}(n))s(q_1)^{-1} \\ &= s(q_1)s(q_2)\iota(n)(s(q_1)s(q_2))^{-1}\end{aligned} \tag{15.15}$$

We want to compare this to $\iota\left(\omega_{q_1q_2}(n)\right)$. In general they will be different unless $s(q_1q_2) = s(q_1)s(q_2)$, that is, unless $s : Q \to G$ is a homomorphism. In general the section is not a homomorphism, but clearly something nice happens when it is:

**Definition**: We say an extension *splits* if there exists a section $s : Q \to G$ which is *also* a *group homomorphism*. A choice of a section which is a group homomorphism is called a (choice of) *splitting*.

**Theorem**: An extension is isomorphic to a semidirect product iff it is a split extension.

*Proof*:

First suppose it splits. Choose a splitting $s$. Then from (15.15) we know that

$$\omega_{q_1} \circ \omega_{q_2} = \omega_{q_1q_2} \tag{15.16}$$

and hence $q \mapsto \omega_q$ defines a homomorphism $\omega : Q \to \mathrm{Aut}(N)$. Therefore, we can aim to prove that there is an isomorphism of $G$ with $N \rtimes_\omega Q$.

In general if $s$ is just a section the image $s(Q) \subset G$ is not a subgroup. But if the sequence splits, then it is a subgroup. The equation (15.8) implies that $G = \iota(N)s(Q)$ where $s(Q)$ is a subgroup, and by the internal characterization of semidirect products that means we have a semidirect product.

To give a more concrete proof, let us write the group law in the parametrization (15.8). Write

$$\iota(n)s(q)\iota(n')s(q') = \iota(n)\left(s(q)\iota(n')s(q)^{-1}\right)s(qq') \tag{15.17}$$

Note that

$$s(q)\iota(n')s(q)^{-1} = \iota(\omega_q(n')) \tag{15.18}$$

so

$$\iota(n_1)s(q_1)\iota(n_2)s(q_2) = \iota\left(n_1\omega_{q_1}(n_2)\right)s(q_1q_2) \tag{15.19}$$

But this just means that

$$\Psi(n,q) = \iota(n)s(q) \tag{15.20}$$

is in fact an isomorphism $\Psi : N \rtimes_\omega Q \to G$. Indeed equation (15.19) just says that:

$$\Psi(n_1,q_1)\Psi(n_2,q_2) = \Psi((n_1,q_1)\cdot_\omega(n_2,q_2)) \tag{15.21}$$

where $\cdot_\omega$ stresses that we are multiplying with the semidirect product rule.

Thus, we have shown that a split extension is isomorphic to a semidirect product $G \cong N \rtimes Q$. The converse is straightforward. ♠

In §15.7 below we will continue the general line of reasoning begun here. However, in order to appreciate the formulae better it is a good idea first to step back and consider a simple but important special case of extensions, namely, the *central extensions*. These are extensions such that $\iota(N)$ is a subgroup of the <u>center</u> of $G$. Here is the official definition: (Note the change of notation from the general situation above): ♣Do the general case first and then specialize? ♣

Let $A$ be an abelian group and $G$ any group.

**Definition** A *central extension* of $G$ by $A$, [219] is a group $\tilde{G}$ and an extension such that

$$1 \to A \overset{\iota}{\to} \tilde{G} \overset{\pi}{\to} G \to 1 \tag{15.22}$$

such that $\iota(A) \subset Z(\tilde{G})$.

We stress again that what we called $G$ in the previous discussion is here called $\tilde{G}$, and what we called $Q$ in the previous discussion is here called $G$.

**Example And Remark**:*Sections of group extensions vs. continuous sections of principal fiber bundles.* Let us return to the very important exact sequence (8.92):

$$1 \to \mathbb{Z}_2 \overset{\iota}{\to} SU(2) \overset{\pi}{\to} SO(3) \to 1 \tag{15.23}$$

The $\mathbb{Z}_2$ is embedded as the subgroup $\{\pm 1\} \subset SU(2)$, so this is a central extension. We said above that there is always a section, but when we said that we did not impose any properties of continuity in the case where $G$ and $Q$ are continuous groups. In this example while there is a section of $\pi$ there is, in fact, no continuous section. Such a continuous section $\pi s = Id$ would imply that $\pi_* s_* = 1$ on the first homotopy group of $SO(3)$. But that is impossible since it would have factor through $\pi_1(SU(2)) = 1$.

We are using a few facts here:

1. Every $SU(2)$ matrix can be written as

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \tag{15.24}$$

   where $\alpha, \beta$ are complex numbers with $|\alpha|^2 + |\beta|^2 = 1$. Writing this equation in terms of the real and imaginary parts of $\alpha, \beta$ we recognize the equation of the unit three dimensional sphere. Now recall that all the spheres of dimension $\geq 2$ are simply connected. Therefore $\pi_1(SU(2)) = 1$ is simply connected.

2. But $SO(3)$ is not simply connected! In fact, using a coffee cup you can informally demonstrate that $\pi_1(SO(3)) \cong \mathbb{Z}_2$. [Demonstrate].

3. If there were a continuous section then $s_* : \pi_1(SO(3)) \to \pi_1(SU(2))$ would be a well-defined group homomorphism and $s \circ \pi = Id$ would imply that on the fundamental groups $Id_* = s_* \pi_*$ in

$$\pi_1(SO(3)) \to \pi_1(SU(2)) = 1 \to \pi_1(SO(3)) \tag{15.25}$$

   But $Id_*$ takes the nontrivial element of $\mathbb{Z}_2$ to the nontrivial element, not to the trivial element. This is impossible if you factor through the trivial group.

In algebraic topology one introduces another kind of topological invariant known as homology. The homology groups of a manifold are Abelian groups that encode many important properties of the manifold. The homology group $H_1(M; \mathbb{Z}_2)$ tells us what are

[219] Some authors say an extension of $A$ by $G$.

the possible 2-fold covers of the manifold $M$. It turns out that $H_1(SO(3);\mathbb{Z}_2) \cong \mathbb{Z}_2$, (this is closely related to $\pi_1(SO(3)) \cong \mathbb{Z}_2$. so there are two double covers of $SO(3)$. One is $O(3) = \mathbb{Z}_2 \times SO(3)$ and the other is $SU(2)$, the nontrivial double cover.

The extension (15.1) generalizes to

$$1 \to \mathbb{Z}_2 \quad \stackrel{\iota}{\to} \quad \mathrm{Spin}(d) \quad \stackrel{\pi}{\to} \quad SO(d) \to 1 \tag{15.26}$$

as well as the two Pin groups which extend $O(d)$:

$$1 \to \mathbb{Z}_2 \quad \stackrel{\iota}{\to} \quad \mathrm{Pin}^{\pm}(d) \quad \stackrel{\pi}{\to} \quad O(d) \to 1 \tag{15.27}$$

we discuss these in Section *** below. Again, in these cases there is no continuous section. Thus, these examples are nontrivial as fiber bundles. Moreover, even if we allow ourselves to choose a discontinuous section, we cannot do so and make it a group homomorphism. In other words these examples are also nontrivial as group extensions.

---

**Exercise**

If $s : Q \to G$ is any section of $\pi$ show that for all $q \in Q$,

$$s(q^{-1}) = s(q)^{-1} n = n' s(q)^{-1} \tag{15.28}$$

for some $n, n' \in N$.

---

---

**Exercise** *The pullback construction*

There is one general construction with extensions which is useful when discussing symmetries in quantum mechanics. This is the notion of *pullback extension.* Suppose we are given both an extension

$$1 \longrightarrow H' \stackrel{\iota}{\longrightarrow} H \stackrel{\pi}{\longrightarrow} H'' \longrightarrow 1 \tag{15.29}$$

and a homomorphism

$$\rho : G \to H'' \tag{15.30}$$

one can define another extension of $G$ by $H'$ known as a *pullback extension.* We are trying to fill in the diagram:

$$\begin{array}{ccccccccc} & & & & & & G & & \\ & & & & & & \downarrow{\scriptstyle\rho} & & \\ 1 & \longrightarrow & H' & \stackrel{\iota}{\longrightarrow} & H & \stackrel{\pi}{\longrightarrow} & H'' & \longrightarrow & 1 \end{array} \tag{15.31}$$

with an extension on the first row of $G$ by $H'$.

We do this by defining a subgroup of the Cartesian product $\tilde{G} \subset H \times G$:

$$\tilde{G} := \{(h, g) | \pi(h) = \rho(g)\} \subset H \times G \tag{15.32}$$

We have an extension of the form

$$1 \longrightarrow H' \xrightarrow{\iota} \tilde{G} \xrightarrow{\tilde{\pi}} G \longrightarrow 1 \qquad (15.33)$$

where $\tilde{\pi}(h,g) := g$. Show that this extension fits in the commutative diagram

$$\begin{array}{ccccccccc} 1 & \longrightarrow & H' & \longrightarrow & \tilde{G} & \xrightarrow{\tilde{\pi}} & G'' & \longrightarrow & 1 \\ & & \| & & \downarrow{\scriptstyle \tilde{\rho}} & & \downarrow{\scriptstyle \rho} & & \\ 1 & \longrightarrow & H' & \longrightarrow & H & \xrightarrow{\pi} & H'' & \longrightarrow & 1 \end{array} \qquad (15.34)$$

(N.B. This is <u>not</u> a morphism of extensions.)

---

---

**Exercise** *The pushforward extension*

Under some circumstances one can complete the diagram

$$\begin{array}{ccccccccc} 1 & \longrightarrow & H' & \xrightarrow{\iota} & H & \xrightarrow{\pi} & H'' & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \rho} & & & & & & \\ & & \tilde{H} & & & & & & \end{array} \qquad (15.35)$$

to get an extension of $H''$ by $\tilde{H}$. This is not as universal as the pullback. But one can construct it if $\rho : H' \to \tilde{H}$ is surjective and $\iota(\ker(\rho)) \triangleleft H$. Give the construction. [220]

---

---

**Exercise** *Choice of splitting and the Euclidean group* Euc$(d)$

As we noted, the Euclidean group Euc$(d)$ is isomorphic to the semidirect product $\mathbb{R}^d \rtimes O(d)$, but to exhibit that we needed to choose an origin about which to define rotation-inversions. See equation (14.35) above.

a.) Show that a change of origin corresponds to a change of splitting.

b.) Using the Seitz notation show that another choice of origin leads to the splitting $R \mapsto \{(1-R)v|R\}$, and verify that this is also a splitting.

♣This is almost redundant with another exercise above. ♣

[220] *Answer*: Define the group $\tilde{G} := H/\iota(\ker(\rho))$. Then note that we can define $\tilde{\iota} : \tilde{H} \to \tilde{G}$ via $\tilde{\iota}(\rho(h)) = \iota(h) + \iota(\ker(\rho))$. Note that if $\rho(h) = \rho(h')$ then the RHS is the same so this does give a well-defined map $\tilde{\iota}$ on the image of $\rho$, but if $\rho$ is surjective that is enough. Now define $\tilde{\pi}(\tilde{e}) = \pi(e)$ where $\tilde{e} = e + \iota(\ker\rho)$. Then we have a commutative diagram:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & H' & \xrightarrow{\iota} & H & \xrightarrow{\pi} & H'' & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \rho} & & \downarrow{\scriptstyle \varphi} & & \updownarrow{\scriptstyle Id} & & \\ 1 & \longrightarrow & \tilde{H} & \xrightarrow{\tilde{\iota}} & \tilde{G} & \xrightarrow{\tilde{\pi}} & H'' & \longrightarrow & 1 \end{array} \qquad (15.36)$$

where $\varphi(h) = h + \iota(\ker(\rho))$.

**Exercise** *Another form of splitting*

Show that an equivalent definition of a split exact sequence for a central extension is that there is a homomorphism $t : \tilde{G} \to A$ which is a left-inverse to $\iota$, $t(\iota(a)) = a$.

(Hint: Define $s(\pi(\tilde{g})) = \iota t(\tilde{g}^{-1}))\tilde{g}$.)

---

**Exercise** *The exact sequence for a product of two cyclic groups*

Revisit the exact sequence discussed in equation (12.52). Show that this sequence splits. [221]

---

**Exercise** *Is A Restriction Of A Split Extension Split?*

Suppose

$$1 \longrightarrow N \xrightarrow{\iota} G \xrightarrow{\pi} Q \longrightarrow 1 \tag{15.37}$$

is a split extension and $N_1 \subset N$ and $Q_1 \subset Q$, so that there is an extension

$$1 \longrightarrow N_1 \xrightarrow{\iota} G_1 \xrightarrow{\pi} Q_1 \longrightarrow 1 \tag{15.38}$$

given by restriction to $G_1 \subset G$. Does it follow that this extension is a split extension?

---

**Exercise** *A Split Central Extension Is A Direct Product*

Suppose

$$1 \longrightarrow N \xrightarrow{\iota} G \xrightarrow{\pi} Q \longrightarrow 1 \tag{15.39}$$

is a split central extension. Show that $G \cong N \times Q$ and the extension is isomorphic to the trivial extension with $\iota$ inclusion into the first factor and $\pi$ projection onto the second factor. [222]

[221] *Answer*: Let $\omega_\ell$ generate $\mathbb{Z}_\ell$ then, using the notation above let $s : \omega_\ell \mapsto (\omega_\ell^{\mu_2 \nu_2}, \omega_\ell^{\mu_1 \nu_1})$.

[222] *Answer*: Choose a splitting $s$. Then use the parametrization $g = \iota(n)s(q)$. The group multiplication can be written

$$\begin{aligned} g_1 g_2 &= \iota(n_1)s(q_1)\iota(n_2)s(q_2) \\ &= \iota(n_1)\iota(n_2)s(q_1)s(q_2) \\ &= \iota(n_1 n_2)s(q_1 q_2) \end{aligned} \tag{15.40}$$

In the going from the first to second line we used that $\iota(n_2)$ is in the center. In going from the second to the third line we used that $\iota$ and $s$ are group homomorphisms. (Recall $s$ is a group homomorphism because it is a splitting.)

## 15.2 Projective Representations

We have already encountered the notion of a matrix representation of a group $G$. This is a homomorphism from $G$ into $GL(d,\kappa)$ for some field $\kappa$. In many contexts in mathematics and physics (especially in quantum physics) one encounters a generalization of the notion of matrix representation known as a *projective representation.* The theory of projective representations is closely related to the theory of central extensions.

Recall that a matrix representation of a group $G$ is a group homomorphism

$$\rho : G \to GL(d,\kappa) \tag{15.41}$$

A *projective representation* is a map

$$\rho : G \to GL(d,\kappa) \tag{15.42}$$

which is "almost a homomorphism" in the sense that

$$\rho(g_1)\rho(g_2) = f(g_1,g_2)\rho(g_1,g_2) \tag{15.43}$$

for some function $f : G \times G \to \kappa^*$. Of course $f(g_1,g_2)$ is "just a c-number" so you might think it is an unimportant nuisance. You might try to get rid of it by redefining

$$\tilde{\rho}(g) = b(g)\rho(g) \tag{15.44}$$

where $b(g) \in \kappa^*$ is a c-number. Then if there exists a function $b : G \to \kappa^*$ such that

$$f(g_1,g_2) \stackrel{?}{=} \frac{b(g_1g_2)}{b(g_1)b(g_2)} \tag{15.45}$$

then $\tilde{\rho}$ would be an honest representation.

The trouble is, in many important contexts, there is no function $b$ so that (15.45) holds. So we need to deal with it.

A simple example is the "spin representation of the rotation group $SO(3)$" where one attempts to define a map:

$$\rho : SO(3) \to SU(2) \subset GL(2,\mathbb{C}) \tag{15.46}$$

that attempts to describe the effects of a rotation on - say - a spin 1/2 particle. In fact, there is no such thing as the "spin representation of the rotation group $SO(3)$." There is a spin projective representation of $SO(3)$. [223]

We saw above that there is a very natural group homomorphism from $SU(2)$ to $SO(3)$, but there is no group homomorphism back from $SO(3)$ to $SU(2)$: There is no splitting. The so-called "spin representation of $SO(3)$" is usually presented by attempting to construct a splitting $\rho : SO(3) \to SU(2)$ using Euler angles. Indeed, under the standard

[223]However, there is a perfectly well-defined spin representation of the Lie algebra $\mathfrak{so}(3)$.

homomorphism $\pi : SU(2) \to SO(3)$ one recognizes that $\exp[\mathrm{i}\theta\sigma^i]$ maps to a rotation by angle $2\theta$ around the the $i^{th}$ axis. For example,

$$u = \exp[\mathrm{i}\theta\sigma^3] = \begin{pmatrix} e^{\mathrm{i}\theta} & 0 \\ 0 & e^{-\mathrm{i}\theta} \end{pmatrix} = \cos\theta + \mathrm{i}\sin\theta\sigma^3 \tag{15.47}$$

acts by

$$\begin{aligned} u\vec{x}\cdot\vec{\sigma}u^{-1} &= u \begin{pmatrix} z & x - \mathrm{i}y \\ x + \mathrm{i}y & -z \end{pmatrix} u^{-1} \\ &= \begin{pmatrix} z & e^{2\mathrm{i}\theta}(x - \mathrm{i}y) \\ e^{-2\mathrm{i}\theta}(x + \mathrm{i}y) & -z \end{pmatrix} \end{aligned} \tag{15.48}$$

One can represent any rotation in $SO(3)$ by a rotation around the $z$-axis, then around the $x$-axis, then around the $z$ axis. Call this $R(\phi, \theta, \psi)$. So one attempts to define $\rho$ by assigning

$$\rho : R(\phi, \theta, \psi) \to e^{\mathrm{i}\frac{\phi}{2}\sigma^3} e^{\mathrm{i}\frac{\theta}{2}\sigma^1} e^{\mathrm{i}\frac{\psi}{2}\sigma^3}. \tag{15.49}$$

Clearly, we are going to have problems making this mapping well-defined. For example, $R(\phi, 0, 0)$ would map to $e^{\mathrm{i}\frac{\phi}{2}\sigma^3}$, but this is not well-defined for all $\phi$ because $R(2\pi, 0, 0) = 1$ and $e^{\mathrm{i}\frac{2\pi}{2}\sigma^3} = -1$. The problem is that the Euler angle coordinates on $SO(3)$ are sometimes singular. So, we need to restrict the domain of $\phi, \theta, \psi$ so that (15.49) is well-defined for every $R \in SO(3)$. However, when we make such a restriction we will spoil the group homomorphism property, but only up to a phase. So, in this way, we get a two-dimensional projective representation of $SO(3)$.

As an exercise you can try the following: Every $SU(2)$ matrix can be written as $u = \cos(\chi) + \mathrm{i}\sin(\chi)\hat{n}\cdot\vec{\sigma}$ and this maps under $\pi$ to a rotation by $2\chi$ around the $\hat{n}$ axis. But again, you cannot smoothly identify every $SO(3)$ rotation by describing it as a rotation by $2\chi$ around an axis.

---

**Exercise** *Projective representations of $\mathbb{Z}_2 \times \mathbb{Z}_2$ and the Pauli group*
Consider the group $\mathbb{Z}_2 \times \mathbb{Z}_2$ multiplicatively:

$$\mathbb{Z}_2 \times \mathbb{Z}_2 = \{1, g_1, g_2, g_1 g_2\} \tag{15.50}$$

with relations $g_1^2 = g_2^2 = 1$ and $g_1 g_2 = g_2 g_1$. Consider the map

$$\rho : \mathbb{Z}_2 \times \mathbb{Z}_2 \to GL(2, \mathbb{C}) \tag{15.51}$$

defined by

$$\begin{aligned} \rho(1) &= 1 \\ \rho(g_1) &= \sigma^1 \\ \rho(g_2) &= \sigma^2 \\ \rho(g_1 g_2) &= \sigma^3 \end{aligned} \tag{15.52}$$

a.) Show that this defines a projective representation of $\mathbb{Z}_2 \times \mathbb{Z}_2$.
b.) Try to remove the phase to get a true representation.
c.) Show that $\rho$ defines a section of an exact sequence with $G$ given by the Pauli group.

---

### 15.2.1 How projective representations arise in quantum mechanics

********

REFER BACK TO WIGNER'S THEOREM SECTION 11.1 ABOVE AND EXPLAIN RELATION TO PROJECTIVE REPRESENTATION

********

We note that projective representations are quite pervasive in modern physics:

1. Projective representations appear naturally in quantization of bosons and fermions. The Heisenberg group is an extension of a translation group (on phase space). In addition, the symplectic group of linear canonical transformations gets quantum mechanically modified by a central extension to define something called the metaplectic group.

2. Projective representations are important in the theory of anomalies in quantum field theory.

3. Projective representations are very important in conformal field theory. The Virasoro group, and the Kac-Moody groups are both nontrivial central extensions of simpler objects.

Now, as we will explain near the end of section 15.3 below projective representations are very closely connected to central extensions. So in the next section we turn to a deeper investigation into the structure of central extensions.

**Remark**: The fact that the symmetry operators $U(g)$ should commute with the Hamiltonian has an important implication. Suppose that $H$ has a complete set of eigenvectors $\Psi_\lambda$ where $\lambda \in \mathrm{Spec}(H)$ is a discrete set of eigenvalues. Then we can restrict $U(g)$ to the different eigenspaces $\mathcal{H}_\lambda$ of $H$, for if

$$H\psi_\lambda = E_\lambda \psi_\lambda \tag{15.53}$$

and $U(g)$ commutes with $H$ then $U(g)\psi_\lambda$ also is an eigenvector of eigenvalue $E_\lambda$. This means that the eigenspaces $\mathcal{H}_\lambda$ are each projective representations of $G$. This can be extremely useful in diagonalizing $H$ and simplifying computations. ♣Need to work in some examples ♣

---

**Exercise** *Kernel Of* $\pi$ *In The Wigner Sequence*

Show that the kernel of $\pi$ in the exact sequence (11.26) is precisely the group of phases times the identity operator. [224]

---

### 15.3 How To Classify Central Extensions

There is an interesting way to classify central extensions of $G$ by $A$.

As before let $s : G \to \tilde{G}$ be a "section" of $\pi$. That is, a map such that

$$\pi(s(g)) = g \qquad \forall g \in G \tag{15.54}$$

As we have stressed, in general $s$ is not a homomorphism. In the case when the sequence splits, that is, when there exists a section which is a homomorphism, then we can say $\tilde{G}$ is isomorphic to a direct product $\tilde{G} \cong A \times G$ via

$$\iota(a)s(g) \to (a, g) \tag{15.55}$$

When the sequence splits the semidirect product of the previous section is a direct product because $A$ is central, so $\omega_g(a) = a$.

Now, let us allow that (15.22) does not necessarily split. Let us choose any section $s$ and measure by how much $s$ differs from being a homomorphism by considering the combination:

$$s(g_1)s(g_2)\left(s(g_1g_2)\right)^{-1}. \tag{15.56}$$

Now the quantity (15.56) is in the kernel of $\pi$ and hence in the image of $\iota$. Since $\iota$ is injective we can *define* a function $f_s : G \times G \to A$ by the equation

$$\iota(f_s(g_1, g_2)) := s(g_1)s(g_2)\left(s(g_1g_2)\right)^{-1}. \tag{15.57}$$

That is, we can write:

$$s(g_1)s(g_2) = \iota(f_s(g_1, g_2))s(g_1g_2) \tag{15.58}$$

The function $f_s$ satisfies the important *cocycle identity*

$$\boxed{f(g_2, g_3)f(g_1, g_2g_3) = f(g_1, g_2)f(g_1g_2, g_3)} \tag{15.59}$$

---

**Exercise** *Derivation Of The Cocycle Identity*

Verify (15.59) by using (15.57) to compute $s(g_1g_2g_3)$ in two different ways.

(Note that simply substituting (15.57) into (15.59) is not obviously going to work because $\tilde{G}$ need not be abelian.)

[224] *Answer*: Suppose $uP = Pu$ for every rank one projection operator. Consider the projection operator for the line through $\psi_1 + z\psi_2$ for any two nonzero vectors $\psi_1, \psi_2 \in \mathcal{H}$. Applying the condition for the cases $z = 1$ and $z = \sqrt{-1}$ deduce that $u$ commutes with $|\psi_1\rangle\langle\psi_2|$. Thus, choosing an ON basis $e_i$ it commutes with $|e_i\rangle\langle e_j|$ and therefore must be proportional to the identity matrix. On the other hand, it also must preserve norms, so it is the group of multiplication by a phase.

---

---

**Exercise** *Simple consequences of the cocycle identity*
a.) By putting $g_1 = 1$ and then $g_3 = 1$ show that any cocycle $f$ must satisfy:

$$f(g,1) = f(1,g) = f(1,1) \qquad \forall g \in G \tag{15.60}$$

b.) Show that [225]

$$f(g,g^{-1}) = f(g^{-1},g). \tag{15.61}$$

---

Now we introduce some fancy terminology:

**Definition:** In general
1. A *2-cochain on $G$ with values in $A$* is a function

$$f : G \times G \to A \tag{15.62}$$

We denote the set of all such 2-cochains by $C^2(G,A)$.
2. A *2-cocycle* is a 2-cochain $f : G \times G \to A$ satisfying (15.59). We denote the set of all such 2-cocycles by $Z^2(G,A)$.

**Remarks**:

1. The fancy terminology is introduced for a good reason because there is a topological space and a cohomology theory underlying this discussion. See Section §15.8 and Section §17.2 for further discussion.
2. Note that $C^2(G,A)$ is naturally an abelian group because $A$ is an abelian group. (Recall example 2.7 of Section §3.) $Z^2(G,A)$ inherits an abelian group structure from $C^2(G,A)$.

So, in this language, given a central extension of $G$ by $A$ and a section $s$ we naturally obtain a two-cocycle $f_s \in Z^2(G,A)$ via (15.57).

Now, if we choose a different section $\hat{s}$ then [226]

$$\hat{s}(g) = \iota(t(g))s(g) \tag{15.63}$$

for some function $t : G \to A$. It is easy to check that

[225] *Answer*: Consider the triple $g \cdot g^{-1} \cdot g$ and apply part (a).
[226] Since we are working with central extensions we could put the $\iota(t(g))$ on either side of the $s(g)$. However, when we discuss non-central extensions later the order will matter.

$$f_{\hat{s}}(g_1, g_2) = f_s(g_1, g_2)t(g_1)t(g_2)t(g_1g_2)^{-1} \tag{15.64}$$

where we have used that $\iota(A)$ is central in $\tilde{G}$.

**Definition:** In general two 2-cochains $f$ and $\hat{f}$ are said to *differ by a coboundary* if there exists a function $t : G \to A$ such that

$$\hat{f}(g_1, g_2) = f(g_1, g_2)t(g_1)t(g_2)t(g_1g_2)^{-1} \tag{15.65}$$

for all $g_1, g_2 \in G$.

One can readily check, using the condition that $A$ is Abelian, that if $f$ is a cocycle then any other $\hat{f}$ differing by a coboundary is also a cocycle. Moreover, being related by a cocycle defines an equivalence relation on the set of cocycles $f \sim \hat{f}$. Thus, we may define:

**Definition:** The *group cohomology* $H^2(G, A)$ is the set of equivalence classes of 2-cocycles modulo equivalence by coboundaries.

Now, the beautiful theorem states that group cohomology classifies central extensions:
227

**Theorem:** Isomorphism classes of central extensions of $G$ by an abelian group $A$ are in 1-1 correspondence with the second cohomology set $H^2(G, A)$.

*Proof*: Let $\mathcal{E}(G, A)$ denote the set of all extensions of $G$ by $A$, and let $\overline{\mathcal{E}}(G, A)$ denote the set of all isomorphism classes of extensions of $G$ by $A$.

We first construct a map:

$$\Psi_{\mathcal{E}\to H} : \mathcal{E}(G, A) \to H^2(G, A) \tag{15.66}$$

To do this, we choose a section, then from (15.57)(15.59)(15.64) we learn that we get a cocycle whose cohomology class does not depend on the section. So

$$\Psi_{\mathcal{E}\to H}\left(1 \to A \to \tilde{G} \to G \to 1\right) = [f_s] \tag{15.67}$$

is well-defined, because the RHS does not depend on the choice of section $s$.

Now we claim that this map descends to a map $\overline{\Psi}_{\mathcal{E}\to H} : \overline{\mathcal{E}}(G, A) \to H^2(G, A)$. Indeed, if we have an isomorphism of central extensions:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & A & \xrightarrow{\iota} & \widetilde{G} & \xrightarrow{\pi} & G & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle \mathrm{Id}} & & \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \mathrm{Id}} & & \\ 1 & \longrightarrow & A & \xrightarrow{\iota'} & \widetilde{G}' & \xrightarrow{\pi'} & G & \longrightarrow & 1 \end{array} \tag{15.68}$$

where $\psi : \tilde{G} \to \tilde{G}'$ is an isomorphism such that the inverse also leads to a commutative diagram, then $\psi$ can be used to map sections of $\pi : \tilde{G} \to G$ to sections of $\pi' : \tilde{G}' \to G$ by $s \mapsto s'$ where $s'(g) = \psi(s(g))$. Then

♣where do we use this condition in the proof? ♣

$$\begin{aligned} s'(g_1)s'(g_2) &= \psi(s(g_1))\psi(s(g_2)) \\ &= \psi(s(g_1)s(g_2)) \\ &= \psi(\iota(f_s(g_1,g_2))s(g_1g_2)) \\ &= \psi(\iota(f_s(g_1,g_2)))\psi(s(g_1g_2)) \\ &= \iota'(f_s(g_1,g_2))s'(g_1g_2) \end{aligned} \tag{15.69}$$

and hence we assign precisely the same 2-cocycle $f(g_1, g_2)$ to the section $s'$. Hence the map $\Psi_{\mathcal{E}\to H}$ only depends on the isomorphism class of the extension. This defines the map $\overline{\Psi}_{\mathcal{E}\to H}$.

Conversely, we can define a map $\Psi_{H\to\mathcal{E}} : Z^2(G, A) \to \mathcal{E}(G, A)$ as follows: Given a cocycle $f \in Z^2(G, A)$ we may define $\tilde{G} = A \times G$ as a set and we use $f$ to *define* the multiplication law:

$$(a_1, g_1)(a_2, g_2) := (a_1a_2f(g_1, g_2), g_1g_2) \tag{15.70}$$

You should check that this does define a valid group multiplication: The associativity follows from the cocycle relation. Note that if we use the *trivial cocycle*: $f(g_1, g_2) = 1$ for all $g_1, g_2 \in G$ then we just get the direct product of groups.

Now suppose that we use two 2-cocycles $f$ and $f'$ which are related by a coboundary as in (15.65) above. Then we claim that the map $\psi : \tilde{G} \to \tilde{G}'$ defined by

$$\psi : (a, g) \to (at(g)^{-1}, g) \tag{15.71}$$

is an isomorphism of central extensions as in (15.68). This means that the map $\Psi_{H\to\mathcal{E}} : Z^2(G, A) \to \mathcal{E}(G, A)$ actually descends to a well-defined map

$$\overline{\Psi}_{H\to\mathcal{E}} : H^2(G, A) \to \overline{\mathcal{E}}(G, A) \tag{15.72}$$

We leave it to the reader to check that $\overline{\Psi}_{H\to\mathcal{E}}$ and $\overline{\Psi}_{\mathcal{E}\to H}$ are inverse maps. ♠

**Remarks**:

1. *Central extensions and projective representations.* A very important consequence of the construction (15.70) is that, if we are given a projective representation of $G$ then we can associate a centrally extended group $\widetilde{G}$:

$$1 \to U(1) \to \widetilde{G} \to G \to 1 \tag{15.73}$$

and a <u>true</u> representation $\widetilde{\rho}$ of $\widetilde{G}$:

$$\widetilde{\rho}(z, g) := z\rho(g) \tag{15.74}$$

[227] In fact, both the set of isomorphism classes of extensions and $H^2(G, A)$ are Abelian groups and a stronger statement is that the 1-1 correspondence described here is an isomorphism of Abelian groups.

The evil failure of $\rho(g)$ to be a true representation of $G$ now becomes a virtuous fact that allows $\widetilde{\rho}$ to be a true representation of $\widetilde{G}$. This is the typical situation in quantum mechanics, where $G$ is a group of classical symmetries and $\widetilde{G}$ is the group that is implemented quantum-mechanically. A good example is the spin-1/2 system where $G = SO(3)$ is the classical group of rotations but for a quantum rotor the proper symmetry group is $\widetilde{G} = SU(2)$. There are many many other examples.

2. *Group Structure on* $\overline{\mathcal{E}}(G, A)$. The set $H^2(G, A)$ carries a natural structure of an Abelian group. Indeed, as we remarked above $C^2(G, A)$, being a set of maps with target space a group, $A$, is naturally a group. Then, because $A$ is Abelian, we can define a group structure on $Z^2(G, A)$ by the rule:

$$(f_1 \cdot f_2)(g, g') = f_1(g, g') \cdot f_2(g, g') \tag{15.75}$$

where we are writing the product in $A$ multiplicatively. Again using the fact that $A$ is abelian this descends to a well-defined multiplication on cohomology classes: $[f_1] \cdot [f_2] := [f_1 \cdot f_2]$. Therefore $H^2(G, A)$ itself is an abelian group. The identity element corresponds to the cohomology class of the trivializable cocycles, which in turn corresponds to the split extension $A \times G$.

♣It might be clearer to write $A$ additively... ♣

It is natural to ask whether one can give a more canonical description of the abelian group structure on the set of equivalence classes of central extensions of $G$ by $A$. Indeed we can: We pull back the Cartesian product to the diagonal of $G \times G$ and then push forward by the multiplication map $\mu : A \times A \to A$. That is, suppose we have two central extensions:

$$\mathcal{E}_1 : \qquad 1 \to A \overset{\iota_1}{\to} \tilde{G}_1 \overset{\pi_1}{\to} G \to 1 \tag{15.76}$$

$$\mathcal{E}_2 : \qquad 1 \to A \overset{\iota_2}{\to} \tilde{G}_2 \overset{\pi_2}{\to} G \to 1 \tag{15.77}$$

The Cartesian product $\mathcal{E}_1 \times \mathcal{E}_2$ is the extension of $G \times G$ by $A \times A$ using the group $\tilde{G}_1 \times \tilde{G}_2$ with the Cartesian product of the group homomorphisms. We want an extension of $G$ by $A$, corresponding, under the 1-1 correspondence of the above theorem to the natural group structure on $H^2(G, A)$. To construct it, let

$$\Delta : G \to G \times G \tag{15.78}$$

be the diagonal homomorphism: $\Delta : g \mapsto (g, g)$. Then we claim that the product extension $\mathcal{E}_1 \cdot \mathcal{E}_2$ can be identified as

$$\mathcal{E}_1 \cdot \mathcal{E}_2 = \mu_* \Delta^*(\mathcal{E}_1 \times \mathcal{E}_2) \tag{15.79}$$

where $\Delta^*(\mathcal{E}_1 \times \mathcal{E}_2)$ is the pull-back extension under $\Delta$ (see equation (15.34)), an extension of $G$ by $A \times A$, and $\mu_*$ is the pushforward extension. In concrete terms the pullback extension under $\Delta^*$ is:

$$1 \to A \times A \overset{(\iota_1, \iota_2)}{\to} \widehat{G_{12}} \overset{\pi_{12}}{\to} G \to 1 \tag{15.80}$$

where

$$\widehat{G_{12}} := \{(\tilde{g}_1, \tilde{g}_2) | \pi_1(\tilde{g}_1) = \pi_2(\tilde{g}_2)\} \subset \tilde{G}_1 \times \tilde{G}_2 \tag{15.81}$$

We can define $\pi_{12}(\tilde{g}_1, \tilde{g}_2) := \pi_1(\tilde{g}_1) = \pi_2(\tilde{g}_2)$. Now consider the "anti-diagonal"

$$A^{\rm anti} := \ker(\mu) = \{(a, a^{-1})\} \subset A \times A \tag{15.82}$$

and its image:

$$N := \{(\iota_1(a), \iota_2(a^{-1})) | a \in A\} \subset \widehat{G_{12}} \tag{15.83}$$

Because we are working with central extensions this will be a normal subgroup. Then we let

$$\tilde{G}_{12} := \widehat{G_{12}}/N \tag{15.84}$$

Since $N$ is in the kernel of $\pi_{12}$ and since it is central the homomorphism $\pi_{12}$ descends to a surjective homomorphism which we will also call $\pi_{12} : \tilde{G}_{12} \to G$. Now we have an exact sequence

$$1 \to A \stackrel{\iota_{12}}{\to} \tilde{G}_{12} \stackrel{\pi_{12}}{\to} G \to 1 \tag{15.85}$$

where $\iota_{12}(a) := [(\iota_1(a), \iota_2(1))] = [(\iota_1(1), \iota_2(a))]$. Given sections $s_1, s_2$ of $\pi_1, \pi_2$ respectively we can define a section $s_{12}(g) := [(s_1(g), s_2(g))]$ and one can check that the resulting cocycle is indeed in the cohomology class of $f_{s_1} \cdot f_{s_2}$. The extension (15.85) represents the product of extensions (15.76) and (15.77). The point of this construction is that it is canonical: We did not make any choices of sections to define the product extension.

3. *Trivial vs. Trivializable.* Above we defined the trivial cocycle to be the one with $f(g_1, g_2) = 1_A$ for all $g_1, g_2$. We define a cocycle to be *trivializable* if it is cohomologous to the trivial cocycle. Note that a trivializable cocycle $f$ could be trivialized in multiple ways. Suppose both $b$ and $\tilde{b}$ trivialize $f$. Then you should show that $\tilde{b}$ and $b$ "differ" by a group homomorphism $\phi : G \to A$ in the sense that

$$\tilde{b}(g) = \phi(g) b(g) \tag{15.86}$$

There are situations where a cohomological obstruction vanishes and the choice of trivialization has physical significance.

♣Should give some examples: $H^3$ is obstruction to orbifolding CFT and choice of trivialization is $H^2$ - hence discrete torsion. There are bundle examples. Find example where class in $H^2$ is zero but trivialization has physical meaning. ♣

4. *An analogy to gauge theory*: Changing a cocycle by a coboundary is strongly analogous to making a gauge transformation in a gauge theory. In Maxwell's theory we can make a change of gauge of the vector potential $A_\mu$ by

$$A'_\mu(x) = A_\mu(x) - {\rm i} g^{-1}(x) \partial_\mu g(x) \tag{15.87}$$

where $g(x) = e^{{\rm i}\epsilon(x)}$ is a function on spacetime valued in $U(1)$. In the case of electromagnetism we would say that $A_\mu$ is trivializable if there is a gauge transformation $g(x)$ that simplifies it to 0. (For valid gauge transformations $g(x)$ must be a single-valued function on spacetime.) If we are presented with $A_\mu(x)$ and we want to know if it is trivializable then we should check whether gauge invariant quantities vanish.

One such quantity is the fieldstrength tensor $F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu$, but this is not a complete gauge invariant. The isomorphism class of a field is completely specified by the holonomies $\exp i \oint_\gamma A$ around all the closed cycles $\gamma$ in spacetime. Even when $A_\mu(x)$ is not trivializable, it is very often useful to use gauge transformations to try to simplify $A_\mu$. In the next remark we do the same for cocycles.

5. *Simplifying Cocycles Using Coboundaries.* Using a coboundary one can usefully simplify cocycles. Since this topic will be unfamiliar to some readers we explain this in excruciating detail. Those who are familiar with cohomology can safely skip the rest of this remark. To begin, note that a coboundary modification takes a cochain $f$ to $f^{(1)}$ satisfying:

$$f^{(1)}(1,1) = f(1,1)\frac{t(1)t(1)}{t(1 \cdot 1)} = f(1,1)t(1) \tag{15.88}$$

so by choosing any function $t$ such that $t(1) = f(1,1)^{-1}$ we get a new cochain satisfying $f^{(1)}(1,1) = 1$. Choose any such function. (The simplest thing to do is set $t(g) = 1$ for all other $g \neq 1$. We will make this choice, but it is really not necessary.) Now recall that if $f$ is a cocycle then a modification of $f$ by any coboundary produces a new cochain $f^{(1)}$ that is also a cocycle. So now, if $f$ is a cocycle and we have set $f^{(1)}(1,1) = 1$ then, by (15.60) we have $f^{(1)}(g,1) = f^{(1)}(1,g) = 1$ for all $g$. Now, we can continue to make modifications by coboundaries to simplify further our cocycle $f^{(1)}$. In order not to undo what we have done we require that the new coboundaries we use, say, $\tilde{t}$ satisfy $\tilde{t}(1) = 1$. We may say that we are "partially choosing a gauge" by choosing representatives so that $f^{(1)}(1,1) = 1$ and then the further coboundaries $\tilde{t}$ must "preserve that gauge." Now suppose that $g \neq 1$. Then (using our particular choice of $t$ above):

$$f^{(1)}(g,g^{-1}) = f(g,g^{-1})\frac{1}{t(1)} = f(g,g^{-1})f(1,1) \tag{15.89}$$

is not particularly special. (Remember that we are making the somewhat arbitrary choice that $t(g) = 1$ for $g \neq 1$.) So we have not simplified these quantities. However, we still have plenty of gauge freedom left and we can try to simplify the values as follows: Suppose, first, that $g \neq g^{-1}$, equivalently, suppose $g^2 \neq 1$ so $g$ is not an involution. Then we can make another "gauge transformation" by a coboundary function $\tilde{t}$ to produce:

$$f^{(2)}(g,g^{-1}) = f^{(1)}(g,g^{-1})\frac{\tilde{t}(g)\tilde{t}(g^{-1})}{\tilde{t}(g \cdot g^{-1})} = \hat{f}(g,g^{-1})\tilde{t}(g)\tilde{t}(g^{-1}) \tag{15.90}$$

where in the second equality we used the "gauge-preserving" property that $\tilde{t}(1) = 1$. Now, in any way you like, divide the non-involution elements of $G$ into two disjoint sets $S_1 \amalg S_2$ so that no two group elements in $S_1$ are related by $g \to g^{-1}$. Then, if $g \in S_2$ we have $g^{-1} \in S_1$ and vice versa. Then we can choose a function $\tilde{t}$ so that for every $g \in S_2$ we have

$$\tilde{t}(g) = (\tilde{t}(g^{-1}))^{-1}(f^{(1)}(g,g^{-1}))^{-1} \tag{15.91}$$

Consequently:

$$f^{(2)}(g, g^{-1}) = 1 \qquad \forall g \in S_2 \tag{15.92}$$

It doesn't really matter what we choose for $\tilde{t}$ on $S_1$. For definiteness we choose it to be $= 1$. But if we had made another choice the above procedure would still lead to equation (15.92). Now recall from (15.61) that any cocycle $f$ satisfies $f(g, g^{-1}) = f(g^{-1}, g)$ for all $g$. Since $f^{(2)}$ is a cocycle (if we started with a cocycle $f$) then we conclude that for all the non-involutions:

$$f^{(2)}(g, g^{-1}) = f^{(2)}(g^{-1}, g) = 1 \qquad \forall g \in S_1 \amalg S_2 \tag{15.93}$$

Note that there is still a lot of "gauge freedom": We have not yet constrained $\tilde{t}(g)$ for $g \in S_1$, nor have we constrained $\tilde{t}(g)$ for the involutions, that is, the group elements $g$ with $g^2 = 1$. What can we say about $f^{(2)}(g, g)$ for $g$ an involution? we have

$$f^{(2)}(g, g) = f^{(1)}(g, g)\frac{\tilde{t}(g)^2}{\tilde{t}(g^2)} = f^{(1)}(g, g)(\tilde{t}(g))^2 \tag{15.94}$$

Now, it might, or might not be the case that $f^{(1)}(g, g)$ is a perfect square in the group. If it is not a perfect square then we are out of luck: We cannot make any further gauge transformations to set $f^{(2)}(g, g) = 1$. Now one can indeed check that the property of $f(g, g)$ being a perfect square, or not, for an involution $g$ is a truly "gauge invariant" condition. Therefore we have proven: *If $f(g, g)$ is not a perfect square for some nontrivial involution $g$ then we know that $f$ is not "gauge equivalent" - that is, is not cohomologous to - the trivial cocycle. That is, $[f]$ is a nontrivial cohomology class.* Such cocycles will define nontrivial central extensions.

**Example 1** . *Extensions of $\mathbb{Z}_2$ by $\mathbb{Z}_2$.* WLOG we can take $f(1, 1) = f(1, \sigma) = f(\sigma, 1) = 1$. Then we have two choices: $f(\sigma, \sigma) = 1$ or $f(\sigma, \sigma) = \sigma$. Each of these choices satisfies the cocycle identity and they are not related by a coboundary. Indeed $\sigma$ is an involution and also $\sigma$ is not a perfect square, so by our discussion above a cocycle with $f(\sigma, \sigma) = \sigma$ cannot be gauged to the trivial cocycle. In other words $H^2(\mathbb{Z}_2, \mathbb{Z}_2) = \mathbb{Z}_2$. For the choice $f = 1$ we obtain $\tilde{G} = \mathbb{Z}_2 \times \mathbb{Z}_2$. For the nontrivial choice $f(\sigma, \sigma) = \sigma$ we obtain $\tilde{G} \cong \mathbb{Z}_4$. Let us see this in detail. We'll let $\sigma_1 \in A \cong \mathbb{Z}_2$ and $\sigma_2 \in G \cong \mathbb{Z}_2$ be the nontrivial elements so we should write $f(\sigma_2, \sigma_2) = \sigma_1$. Note that $(\sigma_1, 1)$ has order 2, but then

$$(1, \sigma_2) \cdot (1, \sigma_2) = (f(\sigma_2, \sigma_2), 1) = (\sigma_1, 1) \tag{15.95}$$

shows that $(1, \sigma_2)$ has order 4. Moreover $(\sigma_1, \sigma_2) = (\sigma_1, 1)(1, \sigma_2) = (1, \sigma_2)(\sigma_1, 1)$. Thus,

$$\begin{aligned} \Psi &: (\sigma_1, 1) \to \omega^2 = -1 \\ \Psi &: (1, \sigma_2) \to \omega \end{aligned} \tag{15.96}$$

where $\omega$ is a primitive $4^{th}$ root of 1 defines an isomorphism with the group of fourth roots of unity. In conclusion, the nontrivial central extension of $\mathbb{Z}_2$ by $\mathbb{Z}_2$ is:

$$1 \to \mathbb{Z}_2 \to \mathbb{Z}_4 \to \mathbb{Z}_2 \to 1 \tag{15.97}$$

Recall that $\mathbb{Z}_4$ is not isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. The square of this extension is the trivial extension.

**Example 2.** *Extensions of $\mathbb{Z}_p$ by $\mathbb{Z}_p$.* The generalization of the previous example to the extension of $\mathbb{Z}_p$ by $\mathbb{Z}_p$ for an odd prime $p$ is extremely instructive. So, let us study in detail the extensions

$$1 \to \mathbb{Z}_p \to G \to \mathbb{Z}_p \to 1 \tag{15.98}$$

In this example we will write our cyclic groups multiplicatively. Now, using methods of topology one can show that [228]

$$H^2(\mathbb{Z}_p, \mathbb{Z}_p) \cong \mathbb{Z}_p. \tag{15.99}$$

The result (15.99) should puzzle you. After all we know that $G$ must be a group of order $p^2$, and we know from the class equation and Sylow's theorems that there are exactly two groups of order $p^2$, up to isomorphism! How is that compatible with the $p$ distinct extensions predicted by equation (15.99) !? The answer is that there can be nonisomorphic extensions (15.22) involving the same group $\tilde{G}$. Let us see how this works in the present example by examining in detail the possible extensions:

$$1 \to \mathbb{Z}_p \xrightarrow{\iota} \mathbb{Z}_{p^2} \xrightarrow{\pi} \mathbb{Z}_p \to 1 \tag{15.100}$$

We write the first, second and third groups in this sequence as

$$\begin{aligned} \mathbb{Z}_p &= \langle \sigma_1 | \sigma_1^p = 1 \rangle \\ \mathbb{Z}_{p^2} &= \langle \alpha | \alpha^{p^2} = 1 \rangle \\ \mathbb{Z}_p &= \langle \sigma_2 | \sigma_2^p = 1 \rangle \end{aligned} \tag{15.101}$$

respectively.

For the injection $\iota$ we have

$$\iota(\sigma_1) = \alpha^x \tag{15.102}$$

for some $x$. For this to be a well-defined homomorphism we must have

$$\iota(1) = \iota(\sigma_1^p) = \alpha^{px} = 1 \tag{15.103}$$

and therefore $px = 0 \bmod p^2$ and therefore $x = 0 \bmod p$. But since $\iota$ must be an injection it must be of the form

$$\iota_k(\sigma_1) := \alpha^{kp} \tag{15.104}$$

[228]You can also show it by examining the cocycle equation directly. We will write down the nontrivial cocycles presently.

where $k$ is relatively prime to $p$. We can take

$$1 \leq k \leq p-1 \tag{15.105}$$

or (preferably) we can regard $k \in \mathbb{Z}_p^*$.

Similarly, for $\pi$ we must have $\pi(\alpha) = \sigma_2^y$ for some $y$. Now, since $\pi$ has to be surjective, $\sigma_2^y$ must be a generator and hence $\pi$ must be of the form

$$\pi_r(\alpha) = \sigma_2^r \qquad\qquad 1 \leq r \leq p-1 \tag{15.106}$$

where again we should really regard $r$ as an element of $\mathbb{Z}_p^*$.

Note that the kernel of $\pi_r$ is the set of elements $\alpha^\ell$ with $\sigma_2^{\ell r} = 1$. This implies $\ell r = 0 \bmod p$ and therefore $\ell = 0 \bmod p$ so

$$\ker(\pi_r) = \{1, \alpha^p, \alpha^{2p}, \dots, \alpha^{(p-1)p}\} \tag{15.107}$$

Since $k \in \mathbb{Z}_p^*$ we have

$$\ker(\pi_r) = \mathrm{im}(\iota_k) \tag{15.108}$$

so our sequence is exact for any choice of $r, k \in \mathbb{Z}_p^*$. We have now described all the extensions of $\mathbb{Z}_p$ by $\mathbb{Z}_p$. Let us find a representative cocycle $f_{k,r}$ for each of these extensions.

To find the cocycle we choose a section of $\pi_r$. It is instructive to try to make it a homomorphism. Therefore we must take $s(1) = 1$. What about $s(\sigma_2)$? It must be of the form $s(\sigma_2) = \alpha^x$ for some $x$, and since $\pi_r(s(\sigma_2)) = \sigma_2$ we must have

$$\sigma_2^{xr} = \sigma_2 \tag{15.109}$$

so that

$$xr = 1 \mathrm{mod}\ p \tag{15.110}$$

Recall that $r \in \mathbb{Z}_p^*$ and let $r^*$ be the integer $1 \leq r^* \leq p-1$ such that

$$rr^* = 1 \mathrm{mod}\ p \tag{15.111}$$

Then we have that $x = r^* + \ell p$ for any $\ell$. That is, $s(\sigma_2)$ could be any of

$$\alpha^{r^*}, \alpha^{r^*+p}, \alpha^{r^*+2p}, \dots, \alpha^{r^*+(p-1)p} \tag{15.112}$$

Here we will make the simplest choice $s(\sigma_2) = \alpha^{r^*}$. The reader can check that the discussion is not essentially changed if we make one of the other choices. (After all, this will just change our cocycle by a coboundary!)

Now that we have chosen $s(\sigma_2) = \alpha^{r^*}$, if $s$ were a homomorphism then we would be forced to take:

$$\begin{aligned} s(1) &= 1 \\ s(\sigma_2) &= \alpha^{r^*} \\ s(\sigma_2^2) &= \alpha^{2r^*} \\ \vdots \quad & \quad \vdots \\ s(\sigma_2^{p-1}) &= \alpha^{(p-1)r^*} \end{aligned} \tag{15.113}$$

But now we are stuck! The property that $s$ is a homomorphism requires two contradictory things. On the one hand, we must have $s(1) = 1$ for any homomorphism. On the other hand, from the above equations we also must have $s(\sigma_2^p) = \alpha^{pr^*}$. But because $1 \leq r^* \leq p-1$ we know that $\alpha^{pr^*} \neq 1$. So the conditions for $s$ being a homomorphism are impossible to meet. Therefore, with this choice of section we find a nontrivial cocycle as follows:

$$s(\sigma_2^x)s(\sigma_2^y)s(\sigma_2^{x+y})^{-1} = \begin{cases} 1 & x+y \leq p-1 \\ \alpha^{r^* p} & p \leq x+y \end{cases} \tag{15.114}$$

Here we computed:

$$\alpha^{r^* x}\alpha^{r^* y}\alpha^{-r^*(x+y-p)} = \alpha^{r^* p} \tag{15.115}$$

where you might note that if $p \leq x+y \leq 2p-2$ then $0 \leq x+y-p \leq p-2$. Therefore, our cocycle is $f_{k,r}$ where

$$f_{k,r}(\sigma_2^x, \sigma_2^y) := \begin{cases} 1 & x+y \leq p-1 \\ \sigma_1^{k^* r^*} & p \leq x+y \end{cases} \tag{15.116}$$

since

$$\iota_k(\sigma_1^{k^* r^*}) = \alpha^{k^* r^* kp} = \alpha^{r^* p} \tag{15.117}$$

and here we have introduced an integer $1 \leq k^* \leq p-1$ so that

$$kk^* = 1 \bmod p \tag{15.118}$$

Although it is not obvious from the above formula for $f_{k,r}$, we know that $f_{k,r}$ will satisfy the cocycle equation because we constructed it from a section of a group extension.

Now, we know the cocycle is nontrivial because $\mathbb{Z}_p \times \mathbb{Z}_p$ is not isomorphic to $\mathbb{Z}_{p^2}$. But let us try to trivialize our cocycle by a coboundary. So we modify our section to

$$\tilde{s}(\sigma_2^x) = \iota(t(\sigma_2^x))s(\sigma_2^x) \tag{15.119}$$

We can always write our function $t$ in the form

$$t(\sigma_2^x) = \sigma_1^{\tau(x)} \tag{15.120}$$

for some function $\tau(x)$ valued in $\mathbb{Z}/p\mathbb{Z}$. We are trying to find a function $\tau(x)$ so that the new cocycle $f_{\tilde{s}}$ is identically 1. We certainly need $\tilde{s}(1) = 1$ and hence $\tau(\bar{0}) = \bar{0}$. But now, because $f(\sigma_2^x, \sigma_2^y) = 1$ already holds for $x+y \leq p-1$ don't want to undo that so we learn that

$$\tau(x) + \tau(y) - \tau(x+y) = 0 \text{mod} p \tag{15.121}$$

for $x+y \leq p-1$. This means we must take

$$\tau(x) = x\tau(1) \qquad\qquad 1 \leq x \leq p-1 \tag{15.122}$$

So, our coboundary is completely fixed up to a choice of $\tau(1)$. But now let us compute for $x+y \geq p-1$:

$$\tilde{s}(\sigma_2^x)\tilde{s}(\sigma_2^y)\tilde{s}(\sigma_2^{x+y})^{-1} = \alpha^{r^* p}\iota(\sigma_1^{\tau(x)+\tau(y)-\tau(x+y)}) = \alpha^{r^* p} \tag{15.123}$$

So, we cannot gauge the cocycle to one, confirming what we already knew: The cocycle is nontrivial.

Now let us see when the different extensions defined by $k, r \in \mathbb{Z}_p^*$ are actually equivalent. To see this let us try to construct $\varphi$ so that

$$\begin{array}{ccccccc} & & & \langle\alpha\rangle & & & \\ & & \nearrow^{\iota_{k_1}} & & \searrow^{\pi_{r_1}} & & \\ 1 \longrightarrow & \langle\sigma_1\rangle & & \downarrow \varphi & & \langle\sigma_2\rangle & \longrightarrow 1 \\ & & \searrow_{\iota_{k_2}} & & \nearrow_{\pi_{r_2}} & & \\ & & & \langle\alpha\rangle & & & \end{array} \tag{15.124}$$

Now $\varphi$, being a homomorphism, must be of the form

$$\varphi(\alpha) = \alpha^y \tag{15.125}$$

for some $y$. We know this must be an isomorphism so $y$ must be relatively prime to $p$. Moreover commutativity of the diagram implies

$$\pi_{r_2}(\varphi(\alpha)) = \pi_{r_1}(\alpha) \qquad \Rightarrow \qquad r_2 y = r_1 \ \mathrm{mod} p \tag{15.126}$$

$$\varphi(\iota_{k_1}(\sigma_1)) = \iota_{k_2}(\sigma_1) \qquad \Rightarrow \qquad k_1 p y = k_2 p \ \mathrm{mod} p^2 \qquad \Rightarrow \qquad k_1 y = k_2 \bmod p \tag{15.127}$$

Putting these equations together, and remembering that $y$ is multiplicatively invertible modulo $p$ we find that there exists a morphism of extensions iff

$$k_1 r_1 = k_2 r_2 \mathrm{mod} p \tag{15.128}$$

Note that the cocycles $f_{k,r}$ constructed in (15.116) indeed only depend on $kr\mathrm{mod}p$. Equivalently, we can label their cohomology class by $(kr)^* = k^* r^* \mathrm{mod} p$.

The conclusion is that $kr \in \mathbb{Z}_p^*$ is the invariant quantity. Extensions with the same group $\tilde{G} = \mathbb{Z}_{p^2}$ in the middle, but with different $kr \in \mathbb{Z}_p^*$, define inequivalent extensions of $\mathbb{Z}_p$ by $\mathbb{Z}_p$.

Now let us examine the group structure on the group cohomology. Just multiplying the cocycles we get:

$$(f_{k_1,r_1} \cdot f_{k_2,r_2})\,(\sigma_2^x, \sigma_2^y) = \begin{cases} 1 & x + y \le p - 1 \\ \sigma_1^{(k_1 r_1)^* + (k_2 r_2)^*} & p \le x + y \end{cases} \tag{15.129}$$

Thus if we map

$$[f_{k,r}] \mapsto (kr)^* \mathrm{mod} \mathbb{Z}_p \tag{15.130}$$

we have a homomorphism of $H^2(G, A)$ to the <u>additive</u> group $\mathbb{Z}/p\mathbb{Z}$, with the trivializable cocycle representing the direct product and mapping to $\bar{0} \in \mathbb{Z}/p\mathbb{Z}$.

In conclusion, we describe the *group* of isomorphism classes of central extensions of $\mathbb{Z}_p$ by $\mathbb{Z}_p$ as follows: The identity element is the trivial extension

$$1 \to \mathbb{Z}_p \to \mathbb{Z}_p \times \mathbb{Z}_p \to \mathbb{Z}_p \to 1 \tag{15.131}$$

and then there is an orbit of $(p-1)$ nontrivial extensions of the form

$$1 \to \mathbb{Z}_p \to \mathbb{Z}_{p^2} \to \mathbb{Z}_p \to 1 \tag{15.132}$$

acted on by $\mathrm{Aut}(\mathbb{Z}_p) = \mathbb{Z}_p^*$.

**Example 3**:*Prime Powers.* Once we start to look at prime powers things start to get more complicated. We will content ourselves with extensions of $\mathbb{Z}_4$ by $\mathbb{Z}_2$. Here it can be shown that

$$H^2(\mathbb{Z}_4, \mathbb{Z}_2) \cong \mathbb{Z}_2 \tag{15.133}$$

so there should be two inequivalent extensions. One is the direct product and the other is

♣Do this more systematically so show that there are precisely two extensions of $\mathbb{Z}_4$ by $\mathbb{Z}_2$. ♣

$$1 \to \mathbb{Z}_2 \to \mathbb{Z}_8 \to \mathbb{Z}_4 \to 1 \tag{15.134}$$

We will think of these as multiplicative groups of roots of unity, with generators $\sigma = -1$ for $\mathbb{Z}_2$, $\alpha = \exp[2\pi i/8]$ for $\mathbb{Z}_8$, and $\omega = \exp[2\pi i/4]$ for $\mathbb{Z}_4$.

The inclusion map $\iota : \sigma \to \alpha^4$, while the projection map takes $\pi : \alpha \to \alpha^2 = \omega$.

Let us try to find a section. Since we want a normalized cocycle we must choose $s(1) = 1$. Now, $\pi(s(\omega)) = \omega$ implies $s(\omega)^2 = \omega$, and this equation has two solutions: $s(\omega) = \alpha$ or $s(\omega) = \alpha^5$. Let us choose $s(\omega) = \alpha$. (The following analysis for $\alpha^5$ is similar.) If we try to make $s$ into a homomorphism then we are forced to choose

$$\begin{aligned} s(\omega) &= \alpha \\ s(\omega^2) &= \alpha^2 \\ s(\omega^3) &= \alpha^3 \end{aligned} \tag{15.135}$$

but now we have no choice - we *must* set $s(\omega^4) = s(1) = 1$. On the other hand, if $s$ *were* to have been a homomorphism we would have wanted to set $s(\omega^4) = s(\omega)^4 = \alpha^4$, but, as we just said, we cannot do this. With the above choice of section we get the symmetric cocycle whose nontrivial entries are

$$f(\omega, \omega^3) = f(\omega^2, \omega^2) = f(\omega^2, \omega^3) = f(\omega^3, \omega^3) = \alpha^4 = \sigma. \tag{15.136}$$

♣Probably should just describe all extensions with $Q = \mathbb{Z}_n$ ♣

**Example 4.**. *Products Of Cyclic Groups.* Another natural generalization is to consider products of cyclic groups. For simplicity we will only consider the case

$$G = \mathbb{Z}_p \oplus \cdots \oplus \mathbb{Z}_p \tag{15.137}$$

where there are $k$ summands, but $p$ is prime. We will think of our group additively and moreover we will think of $\mathbb{Z}_p$ as a ring in this example. If we write elements as $\vec{x} = (x_1, \dots, x_k)$ with $x_i \in \mathbb{Z}_p$ and our cocycle $f(\vec{x}, \vec{y})$ is also valued in $\mathbb{Z}_p$, so that we are considering central extensions:

$$0 \to \mathbb{Z}_p \to \widetilde{G} \to \mathbb{Z}_p^{\oplus k} \to 0 \tag{15.138}$$

then the cocycle condition becomes:

$$f(\vec{x},\vec{y}) + f(\vec{x}+\vec{y},\vec{z}) = f(\vec{x},\vec{y}+\vec{z}) + f(\vec{y},\vec{z}) \tag{15.139}$$

An obvious way to satisfy this condition is to use a bilinear form:

$$f(\vec{x},\vec{y}) = A_{ij}x_i y_j \tag{15.140}$$

where the matrix elements $A_{ij} \in \mathbb{Z}_n$. We can modify by a coboundary:

$$f(\vec{x},\vec{y}) \to f(\vec{x},\vec{y}) + q(\vec{x}+\vec{y}) - q(\vec{x}) - q(\vec{y}) \tag{15.141}$$

Notice a linear term cancels out. It we want to restrict attention to expressions which are quadratic then we can modify

$$A_{ij} \to A_{ij} - (q_{ij} + q_{ji}) \tag{15.142}$$

where $q_{ij}$ is any matrix with values in $\mathbb{Z}_p$.

Now we must distinguish the case $p = 2$ from $p$ an odd prime. If $p = 2$ we can use the coboundary to make the off-diagonal part asymmetric, and WLOG we can agree that for each $i < j$ either $A_{ij} = A_{ji} = 0$ or $A_{ij} = 0$ and $A_{ji} = 1$. Note that the diagonal matrix elements are gauge invariant since $q_{ii} + q_{ii} = 2q_{ii} = 0$. Therefore we can produce in this way $\frac{1}{2}k(k+1)$ independent cocycles.

If $p$ is an odd prime we can require that the matrix is "anti-symmetric" in the sense that $A_{ij} + A_{ji} = 0$ for all $i, j$, because 2 is invertible. In this way we only produce $\frac{1}{2}k(k-1)$ independent cocycles.

On the other hand, using methods of topology (See section **** below for hints) one can prove that

$$H^2(\mathbb{Z}_p^{\oplus k}, \mathbb{Z}_p) \cong \mathbb{Z}_p^{\frac{1}{2}k(k+1)} \tag{15.143}$$

for any prime $p$. What are the $k$ "missing" cocycles for $p$ an odd prime? They are exactly the extensions we discussed in detail in Example 2 above!

**Example 5.**. As a special case of the above, consider extensions of $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ by $\mathbb{Z}_2$. This will be a group of order 8. As we will see, there are five groups of order 8 up to isomorphism:

$$\mathbb{Z}_8, \quad \mathbb{Z}_2 \times \mathbb{Z}_4, \quad \mathbb{Z}_2^3, \quad Q, \quad D_4 \tag{15.144}$$

where $Q$ and $D_4$ are the quaternion and dihedral groups, respectively. Now $\mathbb{Z}_8$ cannot sit in an extension of $\mathbb{Z}_2 \times \mathbb{Z}_2$. (Why not? [229] ) This leaves 4 isomorphisms classes of groups which do fit in extensions of $\mathbb{Z}_2 \times \mathbb{Z}_2$ by $\mathbb{Z}_2$ and it happens they are all central extensions. They are:

$$\begin{aligned} &1 \to \mathbb{Z}_2 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \to 1 \\ &1 \to \mathbb{Z}_2 \to \mathbb{Z}_2 \times \mathbb{Z}_4 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \to 1 \\ &1 \to \mathbb{Z}_2 \to Q \to \mathbb{Z}_2 \times \mathbb{Z}_2 \to 1 \\ &1 \to \mathbb{Z}_2 \to D_4 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \to 1 \end{aligned} \tag{15.145}$$

[229] *Answer*: Because $\mathbb{Z}_2 \times \mathbb{Z}_2$ would have to be a quotient of $\mathbb{Z}_8$. But we can easily list the subgroups of $\mathbb{Z}_8$ and no quotient is of this form.

where $Q$ is the quaternion group and $D_4$ the dihedral group. We have already met $Q$ and $D_4$ above. One can define a homomorphism $\pi : Q \to \mathbb{Z}_2 \oplus \mathbb{Z}_2$ by

$$\begin{aligned}
\pi(\pm 1) &= 0 = (0,0) \\
\pi(\pm \mathrm{i}\sigma^1) &= v_1 = (1,0) \\
\pi(\pm \mathrm{i}\sigma^2) &= v_2 = (0,1) \\
\pi(\pm \mathrm{i}\sigma^3) &= v_1 + v_2 = (1,1)
\end{aligned} \tag{15.146}$$

where we are thinking of $\mathbb{Z}_2 \cong \mathbb{Z}/2\mathbb{Z}$ additively to make contact with the previous example. We can choose a section:

$$\begin{aligned}
s(v_1) &= \mathrm{i}\sigma^1 \\
s(v_2) &= \mathrm{i}\sigma^2 \\
s(v_1 + v_2) &= \mathrm{i}\sigma^3
\end{aligned} \tag{15.147}$$

and, computing the cocycle we find that it is given by the bilinear form (see the previous exercise):

$$A_Q = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \tag{15.148}$$

Similarly, we can define a homomorphism $\pi : D_4 \to \mathbb{Z}_2 \times \mathbb{Z}_2$ by

$$\begin{aligned}
\pi(\pm 1) &= 0 = (0,0) \\
\pi(\pm R(\pi/2)) &= v_1 = (1,0) \\
\pi(\pm P) &= v_2 = (0,1) \\
\pi(\pm P R(\pi/2)) &= v_1 + v_2 = (1,1)
\end{aligned} \tag{15.149}$$

We can choose a section:

$$\begin{aligned}
s(v_1) &= R(\pi/2) \\
s(v_2) &= P \\
s(v_1 + v_2) &= P R(\pi/2)
\end{aligned} \tag{15.150}$$

and, computing the cocycle we find that it is given by the bilinear form (see the previous exercise):

$$A_{D_4} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \tag{15.151}$$

Now, on the other hand, using methods of topology one can prove that

$$H^2(\mathbb{Z}_2 \times \mathbb{Z}_2, \mathbb{Z}_2) = \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \tag{15.152}$$

We can understand this group in terms of the bilinear quadratic forms mentioned in the previous example. Under $A_{ij} \to A_{ij} - (q_{ij} + q_{ji})$. Note that $A_{11}$ and $A_{22}$ are invariant, but we can modify the off-diagonal part of $A_{ij}$ by a symmetric matrix. Thus, there are 8 possible values for $A_{11}, A_{22}, A_{12} \in \mathbb{Z}_2$.

In a way analogous to our discussion of extensions of $\mathbb{Z}_p$ by $\mathbb{Z}_p$, while there are only four different isomorphism classes of groups, there can be different extensions. An extension with group cocycle $A_{ij}x_1^i x_2^j$ defines a group of elements $(z, \vec{x})$. If we only care about the isomorphism class of the group we are free to consider an isomorphism

$$(z, \vec{x}) \mapsto (z, S\vec{x}) \tag{15.153}$$

where $S \in GL(2, \mathbb{Z}_2)$. This maps $A \to SAS^{tr}$. In general that will produce an isomorphic group, but a different extension.

In all our examples up to now the group $\tilde{G}$ has been Abelian, but in this example we have produced two nonisomorphic nonabelian groups $Q$ and $D_4$ of order 8.

♣Explain in detail how this takes us from 8 extensions to four isomorphism classes of groups using the explicit transforms by elements of $GL(2, \mathbb{Z}_2) \cong S_3$. ♣

♣Should add an exercise showing that $f(g, g)$ for $g$ an involution determines the entire cocycle in this case. There are three nontrivial involutions, again giving 8 possible nonisomorphic cocycles. ♣

♣Should also consider central extensions of $\mathbb{Z}_2$ by $\mathbb{Z}_2 \oplus \mathbb{Z}_2$. ♣

**Example 5.**. Nonabelian groups can also have central extensions. Indeed, we already saw this for $G = SO(3)$. Here is an example with $G$ a nonabelian finite group. We take $G$ to be the symmetric group $S_n$. It turns out that it has one nontrivial central extension by $\mathbb{Z}_2$:

$$H^2(S_n; \mathbb{Z}_2) \cong \mathbb{Z}_2 \tag{15.154}$$

To define it we let $\sigma_i = (i, i+1)$, $1 \le i \le n-1$ be the transpositions generating $S_n$. Then $\hat{S}_n$ is generated by $\hat{\sigma}_i$ and a central element $z$ satisfying the relations:

$$\begin{aligned} z^2 &= 1 \\ \hat{\sigma}_i^2 &= z \\ \hat{\sigma}_i\hat{\sigma}_{i+1}\hat{\sigma}_i &= \hat{\sigma}_{i+1}\hat{\sigma}_i\hat{\sigma}_{i+1} \\ \hat{\sigma}_i\hat{\sigma}_j &= z\hat{\sigma}_j\hat{\sigma}_i \qquad\qquad j > i+1 \end{aligned} \tag{15.155}$$

When restricted to the alternating group $A_n$ we get an extension of $A_n$ that can be elegantly described using spin groups.

**Remarks**:

1. One generally associates cohomology with the subject of topology. There is indeed a beautiful topological interpretation of group cohomology in terms of "classifying spaces."

2. In the case where $G$ is itself abelian we can use more powerful methods of homological algebra to classify central extensions.

3. The special case $H^2(G, U(1))$ (or sometimes $H^2(G, \mathbb{C}^*)$, they are the same) is known as the *Schur multiplier*. It plays an important role in the study of projective representations of $G$. We will return to this important point.

4. We mentioned that a general extension (15.1) can be viewed as a principal $N$ bundle over $Q$. Let us stress that trivialization of $\pi : G \to Q$ as a principal bundle is completely different from trivialization of the extension (by choosing a splitting).

These are different mathematical structures! For example, for finite groups the bundle is of course trivial because any global section is also continuous. However, as we have just seen the extensions might be nontrivial. It is true, quite generally, that if a central extension is trivial as a group extension then $\tilde{G} = A \times G$ and hence $\pi : \tilde{G} \to G$ is trivializable as an $A$-bundle.

♣In general a central extension by $U(1)$ is equivalent to a line bundle over the group and you should explain that here. ♣

---

**Exercise**

Suppose that the central extension (15.22) is equivalent to the trivial extension with $\tilde{G} = A \times G$, the direct product. Show that the possible splittings are in one-one correspondence with the set of group homomorphisms $\phi : G \to A$.

---

---

**Exercise**

Construct cocycles corresponding to each of the central extensions in (15.145) and show how the automorphisms of $\mathbb{Z}_2 \times \mathbb{Z}_2$ account for the the fact that there are only four entries in (15.145) while (15.152) is order 8.

---

---

**Exercise** $D_4$ *vs.* $Q$

a.) Show that $D_4$ and $Q$ both fit in exact sequences

$$1 \to \mathbb{Z}_4 \to D_4 \to \mathbb{Z}_2 \to 1 \tag{15.156}$$

$$1 \to \mathbb{Z}_4 \to Q \to \mathbb{Z}_2 \to 1 \tag{15.157}$$

b.) Are these <u>central</u> extensions?

c.) Are $D_4$ and $Q$ isomorphic? [230]

---

---

**Exercise**

Choosing the natural section $s : \sigma_i \to \hat{\sigma}_i$ in (15.155) and find the corresponding cocycle $f_s$.

---

[230] *Answer*: No. $D_4$ has 5 nontrivial involutions: The reflections in the four symmetry axes of the square and the rotation by $\pi$, while $Q$ has only one nontrivial involution, namely $-1$.

**Exercise** *Due Diligence*

Show that the associative law for the twisted product (15.70) is equivalent to the cocycle condition on the 2-cochain $f$.

---

---

**Exercise** *Involution Criterion For A Nontrivial Cocycle*

Let $g$ be a nontrvial involution. Show that the condition that $f(g,g)$ is, or is not, a perfect square is independent of which cocycle we use within a cohomology class.

---

---

**Exercise** *Group Commutator Criterion For A Nontrivial Cocycle*

a.) Show that if a central extension is defined by a cocycle $f$ then the group commutator is:

$$[(a_1, g_1), (a_2, g_2)] = \left( \frac{f(g_1 g_2, g_1^{-1} g_2^{-1})}{f(g_2 g_1, g_1^{-1} g_2^{-1})} \frac{f(g_1, g_2)}{f(g_2, g_1)}, g_1 g_2 g_1^{-1} g_2^{-1} \right) \tag{15.158}$$

b.) Suppose $G$ is abelian. Show that $\tilde{G}$ is abelian iff $f(g_1, g_2)$ is symmetric.

c.) In general the condition that $f$ is symmetric: $f(g_1, g_2) = f(g_2, g_1)$ would not be preserved by a coboundary transformation. Show that it does make sense in this setting.

d.) Suppose $\tilde{G}$ is a central extension of a not-necessarily-Abelian group $G$ by an Abelian group $A$. Show that if $(g_1, g_2)$ is a commuting pair of elements in $G$ and if $f(g_1, g_2)/f(g_2, g_1)$ is not the identity then the extension is nontrivial. [231]

---

## 15.4 Extended Example: Charged Particle On A Circle Surrounding A Solenoid

In the following extended example we will illustrate how classical symmetries can be centrally extended in the context of a very interesting quantum system. Along the way we will take the opportunity to introduce many ideas about quantum field theory in a very simple context.

### 15.4.1 Hamiltonian Analysis

Consider a particle of mass $m$ confined to a ring of radius $r$ in the $xy$ plane. The position of the particle is described by an angle $\phi$, so we identify $\phi \sim \phi + 2\pi$, and the action is

$$S = \int \frac{1}{2} m r^2 \dot{\phi}^2 = \int \frac{1}{2} I \dot{\phi}^2 \tag{15.159}$$

[231] *Answer*: This follows because, as we saw above, a split central extension is a direct product. But the group commutator of $(1, g_1)$ and $(1, g_2)$ must then be the identity. On the other hand, $f(g_1, g_2)/f(g_2, g_1)$ is gauge invariant, so if it is nontrivial then the group commutator cannot be the identity.

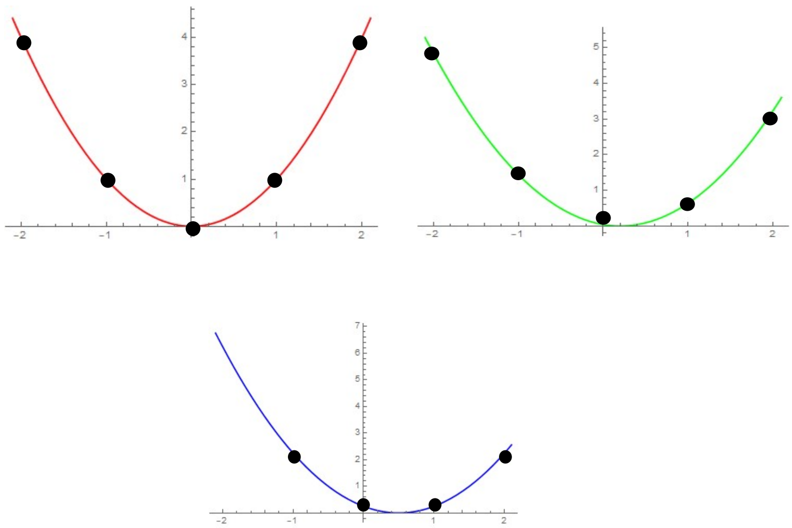


**Figure 42:** Spectrum of a particle on a circle as a function of $\mathcal{B} = eB/2\pi$. The upper left shows the low-lying spectrum for $\mathcal{B} = 0$. It is symmetric under $m \to -m$. The upper right shows the spectrum for $\mathcal{B} = 0.2$. There is no symmetry in the spectrum. The lower figure shows the spectrum for $\mathcal{B} = 1/2$. There is again a symmetry, but under $m \to 2\mathcal{B} - m = 1 - m$. In general there will be no symmetry unless $2\mathcal{B} \in \mathbb{Z}$. If $2\mathcal{B} \in \mathbb{Z}$ the spectrum is symmetric under $m \to 2\mathcal{B} - m$.

with $I = mr^2$ the moment of inertia.

Let us also suppose that our particle has electric charge $e$ and that the ring is threaded by a solenoid with magnetic field $B$, so the particle moves in a zero $B$ field, but there is a nonzero gauge potential [232]

$$A = \frac{B}{2\pi} d\phi \tag{15.160}$$

The action is therefore:

$$\begin{aligned} S &= \int \frac{1}{2} I \dot{\phi}^2 dt + \oint eA \\ &= \int \frac{1}{2} I \dot{\phi}^2 dt + \frac{eB}{2\pi} \dot{\phi} dt \end{aligned} \tag{15.161}$$

The second term is an example of a "topological term" or a "$\theta$-term." Classically, the second term does not affect physical predictions, since it is a total derivative. However as we will soon see, quantum mechanically, it will have an important effect on physical predictions.

[232]For readers not familiar with differential form notation this means, in cylindrical coordinates that $A_z = 0$, $A_r = 0$ and $A_\phi = B/2\pi$.

We are going to analyze the symmetries of this system and compare their realization in the classical and quantum theories.

Classical Symmetries:

We begin by analyzing the classical symmetries. Because the $\theta$-term does not affect the classical dynamics the classical system has $O(2)$ symmetry. We can rotate: $R(\alpha) : e^{i\phi} \to e^{i\alpha} e^{i\phi}$, or, if you prefer, translate $\phi \to \phi + \alpha$ (always bearing in mind that $\alpha$ and $\phi$ are only defined modulo addition of an integral multiple of $2\pi$). If we think of the circle in the $x - y$ plane centered on the origin, with the solenoid along the $z$-axis then we could also take as usual:

$$R(\alpha) = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} . \tag{15.162}$$

Also we can make a "parity" or "charge conjugation" transformation $P : \phi \to -\phi$. The second term in the Lagrangian is not invariant but this "doesn't matter" because it is a total derivative. Put differently: $\phi \to -\phi$ is a symmetry of the equations of motion, and hence it is a classical symmetry.

Note that these group elements in $O(2)$ satisfy

$$\begin{aligned} R(\alpha)R(\beta) &= R(\alpha + \beta) \\ P^2 &= 1 \\ PR(\alpha)P &= R(-\alpha) \end{aligned} \tag{15.163}$$

and indeed, as we have seen, $O(2)$ is a semidirect product:

$$O(2) = SO(2) \rtimes \mathbb{Z}_2 \tag{15.164}$$

with $\omega : \langle P \rangle \cong \mathbb{Z}_2 \to \mathrm{Aut}(SO(2)) \cong \mathbb{Z}_2$ acting by taking the nontrivial element of $\mathbb{Z}_2$ to the outer automorphism that sends $R(\alpha) \to R(-\alpha)$.

Diagonalizing The Hamiltonian

Now let us consider the quantum mechanics with the "$\theta$-term" added to the Lagrangian. Our goal is to see how that term affects the quantum theory.

We will first analyze the quantum mechanics in the Hamiltonian approach. See the remark below for some remarks on the path integral approach. The conjugate momentum is

$$L = I\dot{\phi} + \frac{eB}{2\pi} \tag{15.165}$$

We denote it by $L$ because it can be thought of as angular momentum.

Note that the coupling to the flat gauge field has altered the usual relation of angular momentum and velocity. Now we obtain the Hamiltonian from the Legendre transform:

$$\int L\dot{\phi}dt - S = \int \frac{1}{2I}(L - \frac{eB}{2\pi})^2 dt \tag{15.166}$$

Upon quantization $L \to -\mathrm{i}\hbar\frac{\partial}{\partial\phi}$, so the Hamiltonian is

$$H_{\mathcal{B}} := \frac{\hbar^2}{2I}\left(-\mathrm{i}\frac{\partial}{\partial\phi} - \mathcal{B}\right)^2 \tag{15.167}$$

where $\mathcal{B} := \frac{eB}{2\pi\hbar}$.

The eigenfunctions of the Hamiltonian $H_{\mathcal{B}}$ are just

$$\Psi_m(\phi) = \frac{1}{\sqrt{2\pi}}e^{\mathrm{i}m\phi} \qquad m \in \mathbb{Z} \tag{15.168}$$

They give energy eigenstates with energy

$$E_m = \frac{\hbar^2}{2I}(m - \mathcal{B})^2 \tag{15.169}$$

There is just one energy eigenstate for each $m \in \mathbb{Z}$.

Before moving on with the analysis of the symmetries in this quantum mechanical problem let us take the opportunity to make a long list of:

**Remarks**:

1. The action (15.161) makes good sense for $\phi$ valued in the real line or for $\phi \sim \phi + 2\pi$, valued in the circle. Making this choice is important in the choice of what theory we are describing. Where - in the above analysis - did we make the choice that the target space is a circle? [233]

2. Taking $\phi \sim \phi + 2\pi$, even though the $\theta$-term is a total derivative it has a nontrivial effect on the quantum physics as we can see since $B$ has shifted the spectrum of the quantum Hamiltonian in a physically observable fashion: *This is how we see that topological terms matter.*

3. Note that when $2\mathcal{B}$ is even the energy eigenspaces are two-fold degenerate, except for the ground state at $m = \mathcal{B}$. On the other hand, when $2\mathcal{B}$ is odd all the energy eigenspaces are two-fold degenerate, including the ground state. If $2\mathcal{B}$ is not an integer all the energy eigenspaces are one-dimensional. See Figure 42.

4. The total spectrum is *periodic* in $\mathcal{B}$, and shifting $\mathcal{B} \to \mathcal{B}+1$ is equivalent to $m \to m+1$. To be more precise, we can define a unitary operator on the Hilbert space by its action on a basis:

$$U\Psi_m = \Psi_{m+1} \tag{15.170}$$

and

$$UH_{\mathcal{B}}U^{-1} = H_{\mathcal{B}+1} \tag{15.171}$$

[233] *Answer*: If we took the case where $\phi$ is valued in $\mathbb{R}$ and not the circle then there would be no quantization on $m$ and the spectrum of the Hamiltonian would be continuous. In this case the Chern-Simons term would not affect the physics in the quantum mechanical version as well.

5. The quantum mechanics problem (15.161) and the spectrum (15.169) arise in the discussion of the "Coulomb blockade" in physics of quantum dots. See Yoshimasa Murayama, *Mesoscopic Systems*, Section 10.10.

6. *Viewing the system as a field theory.* We have introduced this system as describing the quantum mechanics of a particle. However, it is important to note that it can also be viewed as a special case of a quantum field theory. In general, in a field theory [234] we have a spacetime $M$ and the fields $\phi$ are functions on $M$ valued in some *target space* $\mathcal{X}$. (So the term "target space" means nothing more or less than the codomain of the fields.) An important example is that of a nonrelativistic particle of mass $m$ moving on a Riemannian manifold $X$ with metric $ds^2 = g_{\mu\nu}(x)dx^\mu \otimes dx^\nu$. The action would be

$$S = \int dt \frac{m}{2} g_{\mu\nu}(x(t))\dot{x}^\mu \dot{x}^\nu dt \tag{15.172}$$

If, in addition, the particle has charge $e$ and there is an electromagnetic potential $A_\mu(x)dx^\mu$ on $X$ then the action is

$$S = \int dt \frac{m}{2} g_{\mu\nu}(x(t))\dot{x}^\mu \dot{x}^\nu dt + \int eA_\mu(x(t))\dot{x}^\mu dt \tag{15.173}$$

Here $M$ is the manifold of time. It could be $M = \mathbb{R}$ if we describe the entire history of the particle, or $M = [t_{in}, t_{fin}]$ if we describe only the motion in a finite time interval. As we will soon see, it can also be interesting to let $M = S^1$. The "field" is a suitably differentiable map

$$x : M \to X \tag{15.174}$$

describing the position of the particle as a function of time. This is an example of a "$0+1$ dimensional field theory." A generalization would be a theory of maps from a $d$-dimensional spacetime with metric $h_{ab}d\sigma^a d\sigma^b$ and action

$$S = \int d^{d+1}\sigma \sqrt{|\det h|} h^{ab}(\sigma) \frac{1}{2} m g_{\mu\nu}(x(t)) \partial_a x^\mu \partial_b x^\nu \tag{15.175}$$

and the "field" would be a suitably differentiable map:

$$x : M \to X \tag{15.176}$$

Equations (15.173) and (15.175) are examples of what is known as a "nonlinear sigma model." [235] In our case our fields are maps

$$e^{i\phi} : M \to S^1 \tag{15.180}$$

[234] As traditionally conceived. The topic of topological field theory generalizes the next few lines considerably.

[235] For the mathematically sophisticated reader we note that, for general nonlinear sigma models

$$dx : T_\sigma M \to T_{x(\sigma)} X \tag{15.177}$$

is a linear map between two inner-product spaces. We can use the inner products to define $(dx)^\dagger$ and then the kinetic term is

$$\int_M \mathrm{Tr}((dx)(dx)^\dagger) \mathrm{vol}\,(h) \tag{15.178}$$

We have been referring to $\phi \to -\phi$ as "parity" because that is the appropriate term in the context of the quantum mechanics of a particle constrained to a circle in the plane. The parity operation is just reflection around some line in the plane. However, if we take the point of view that we are discussing a $0+1$ dimensional "field theory" then it would be better to refer to the operation as "charge conjugation" because it complex conjugates the $U(1)$-valued field $e^{\mathrm{i}\phi}$.

In addition there are (in the field theory interpretation) "worldvolume symmetries" of time translation invariance and time reversal. These form the group $\mathbb{R} \rtimes \mathbb{Z}_2$. We will put those aside. (Note that time reversal is not a symmetry of the second term in the Lagrangian but is a symmetry of the space of solutions of the equations of motion.)

7. *Relations to higher dimensional field theories and string theory.* The $\theta$-term we have added has a very interesting analog in $1+1$ dimensional field theory, where it is known as a coupling to the $B$-field. It can also be obtained from a Kaluza-Klein reduction of $1+1$ dimensional Maxwell theory:

$$\begin{aligned} S &= \frac{1}{e^2}\int dx^0 dx^1 F_{01}^2 + \int \frac{\theta}{2\pi} F_{01} dx^0 \wedge dx^1 \\ &= \frac{1}{e^2}\int F * F + \int \frac{\theta}{2\pi} F \end{aligned} \tag{15.181}$$

In $1+1$ dimensional theory we can choose $A_0 = 0$ gauge and gauge away the $x^1$ dependence so that on $S^1 \times \mathbb{R}$ the only gauge invariant quantity is

$$e^{\mathrm{i}\phi(t)} = e^{\mathrm{i}\oint_{S^1} A} = e^{\mathrm{i}\oint_{S^1} A_1 dx^1} \tag{15.182}$$

With this in mind we can say

$$\theta = 2\pi\mathcal{B} \tag{15.183}$$

**Remark**: More generally, in $1+1$ dimensional Yang-Mills theory on $S^1 \times \mathbb{R}$ we can always go to $A_0 = 0$ gauge and then the only gauge invariant observable is the conjugacy class of the holonomy around the circle.

The theta term also has a close analog in $3+1$-dimensional gauge theory. In the case of $3+1$ dimensional Maxwell theory we can write

$$\begin{aligned} S &= \int d^4x \frac{1}{4e^2} F_{\mu\nu}F^{\mu\nu} + \int \frac{\theta}{8\pi}\epsilon^{\mu\nu\lambda\rho} F_{\mu\nu}F_{\lambda\rho} d^4x \\ &= \int \frac{1}{2e^2} F * F + \int \frac{\theta}{(4\pi)} F \wedge F \end{aligned} \tag{15.184}$$

More to the point $dx$ is a section of $(TM)^{\vee} \otimes f^*(TX)$ and we can use the metric on this bundle to write $\int_M \parallel dx \parallel^2$. In the case of the charged particle moving on the Riemannian manifold $X$, there is also the data of a principal $U(1)$ bundle with connection $d+A$ and the topological term is based on the holonomy of the pulled back connection:

$$S = \int \frac{m}{2} \parallel dx \parallel^2 dt + \oint ex^*(A) \tag{15.179}$$

There are similar topological terms for the $d > 0$ sigma models.

In fact, in the effective theory of electromagnetism in the presence of an insulator a very similar action arises with a $\theta$ term. If a parity- and/or time-reversal symmetry is present then $\theta$ is zero or $\pi$, corresponding to our case $2\mathcal{B} \in \mathbb{Z}$. The difference between a normal and a topological insulator is then, literally, the difference between $2\mathcal{B}$ being even (normal) and odd (topological), respectively. Finally, in the 3+1-dimensional Yang-Mills theories that describe the standard model of electro-weak and strong interactions one can add an analogous $\theta$-term. Topological terms matter, and in this case the topological term for the strong gauge field leads to the prediction of an intrinsic electric dipole moment of the neutron. However, to excellent accuracy it is known that if the neutron dipole moment it is very small and

$$|\theta| < 10^{-9} \tag{15.185}$$

One of the great unsolved mysteries about nature is why the (effective) theta angle for the strong interactions in the standard model is so small. [236]

Now let us get back to the symmetries of the particle on the ring. We have seen that the classical "internal" symmetry group - the "internal" symmetry group of the equations of motion - is $O(2)$. Now let us analyze how the symmetries are implemented in the quantum theory:

In quantum mechanics the $SO(2)$ shift symmetry $\phi \to \phi+\alpha$ is realized by a translation operator $\rho(R(\alpha)) = \mathcal{R}(\alpha)$ and acting on $\Psi_m$ we have

$$(\mathcal{R}(\alpha) \cdot \Psi_m) = e^{\mathrm{i}m\alpha}\Psi_m \tag{15.186}$$

Can we also represent $\rho(P) = \mathcal{P}$ on the Hilbert space? Classically, parity symmetry $P$ just takes $\phi \to -\phi$. If we make this substitution in the Hamiltonian $H_{\mathcal{B}}$ we see that the naive parity operation takes

$$\mathcal{P}H_{\mathcal{B}}\mathcal{P}^{-1} = H_{-\mathcal{B}} \tag{15.187}$$

For general values of $\mathcal{B}$ the operator $H_{\mathcal{B}}$ is not unitarily equivalent to $H_{-\mathcal{B}}$. However, thanks to (15.171) it is clear that when $2\mathcal{B} \in \mathbb{Z}$ they are unitarily equivalent and the naive operation of taking $\phi \to -\phi$, which takes $m \to -m$ on eigenvectors of $H_{\mathcal{B}}$, should be accompanied by $U^{2\mathcal{B}}$. Therefore $\mathcal{P}$ should map the eigenspace associated with $m$ to that associated with with $2\mathcal{B} - m$. As a sanity check note that indeed $E_m = E_{2\mathcal{B}-m}$. Therefore we should define a parity operation:

$$\mathcal{P} \cdot \Psi_m = \xi_m \Psi_{2\mathcal{B}-m} \tag{15.188}$$

where $\xi_m$ is a phase which we can take to be 1. Note that the operator $\mathcal{P}$ so defined commutes with the Hamiltonian: Indeed, it takes eigenvectors to eigenvectors with the same eigenvalue.

♣You should allow the possibility of a phase in the definition of $\mathcal{P}$ and show in detail it doesn't matter. ♣

If $2\mathcal{B}$ is not an integer the parity symmetry is broken and the quantum symmetry group is just $SO(2)$.

[236] For much more about this see M. Dine's TASI lectures https://arxiv.org/pdf/hep-ph/0011376.pdf.

Now consider the case when $2\mathcal{B} \in \mathbb{Z}$ and let us study the relations obeyed by the operators $\mathcal{R}(\alpha)$ and $\mathcal{P}$ and compare them with the classical relations (15.163). We still have $\mathcal{R}(\alpha)\mathcal{R}(\beta) = \mathcal{R}(\alpha+\beta)$ and $\mathcal{P}^2 = 1$ but now the third line of (15.163) is modified to:

$$\mathcal{P}\mathcal{R}(\alpha)\mathcal{P} = e^{\mathrm{i}2\mathcal{B}\alpha}\mathcal{R}(-\alpha) \tag{15.189}$$

---

**Exercise** *Due Diligence*

a.) Check that (15.189) is well-defined, even though $\alpha$ is only defined up to a shift by an integral multiple of $2\pi$.

b.) Check the operator relation (15.189)!

---

We now consider the group of operators generated by the operators $\mathcal{P}$, $\mathcal{R}(\alpha)$, and $z1_{\mathcal{H}}$ where $z \in U(1)$. (Do not forget that we identify $\alpha \sim \alpha + 2\pi$. This will be quite important in what follows.) Denote this group of operators by $\mathcal{G}_{\mathcal{B}}$. Naively we might have expected this group of operators on Hilbert space to be isomorphic to $U(1) \times O(2)$ where $O(2)$ is our classical symmetry group and $U(1)$ is just the group of phases acting on wavefunctions by scalar multiplication. However, equation (15.189) is not satisfied by a direct product. So, how is $\mathcal{G}_{\mathcal{B}}$ related to $U(1)$ and $O(2)$? General principles tell us it will be an extension

$$1 \longrightarrow U(1) \longrightarrow \mathcal{G}_B \longrightarrow O(2) \longrightarrow 1 \tag{15.190}$$

But what extension?

Now, when $\mathcal{B}$ is an integer we can indeed define an isomorphism of $\mathcal{G}_{\mathcal{B}}$ with $U(1)\times O(2)$ by setting

$$\tilde{\mathcal{R}}(\alpha) := e^{-\mathrm{i}\mathcal{B}\alpha}\mathcal{R}(\alpha) \tag{15.191}$$

We now recover the standard relations of $O(2)$, so the classical $O(2)$ symmetry is not modified quantum mechanically. However, when $\mathcal{B}$ is a half-integer, $\tilde{\mathcal{R}}$ is not well-defined since we must identify $\alpha \sim \alpha + 2\pi$. In this case the group $\mathcal{G}_{\mathcal{B}}$ is really different.

To understand what happens when $\mathcal{B}$ is half-integral we introduce a new group called Spin(2). As an abstract group it is isomorphic to $SO(2)$, and $U(1)$, and $\mathbb{R}/\mathbb{Z}$. The groups are all isomorphic. What makes Spin(2) nontrivial is its relation to $SO(2)$. The group elements in Spin(2) can be parametrized by $\hat{\alpha}$ with $\hat{\alpha} \sim \hat{\alpha} + 2\pi$. Let us call the elements of the spin group $\hat{R}(\hat{\alpha})$. You can think of it in terms of Pauli matrices as

$$\hat{R}(\hat{\alpha}) = \exp[\hat{\alpha}\sigma^1\sigma^2] = \cos(\hat{\alpha}) + \mathrm{i}\sin(\hat{\alpha})\sigma^3 \tag{15.192}$$

But it is called the *spin group* because it comes with a nontrivial double cover:

$$\pi : \mathrm{Spin}(2) \to SO(2) \tag{15.193}$$

the double covering is given by restricting our standard projection $\pi : SU(2) \to SO(3)$ to the subgroup of $SU(2)$ in (15.192). In this way we get a double cover of the rotation group around the $z$ axis:

$$\pi : \hat{R}(\hat{\alpha}) \mapsto R(2\hat{\alpha}) \tag{15.194}$$

See equation (15.48) above.

Now, taking $\mathbb{Z}_2$ to act on Spin(2) by the nontrivial outer automorphism. So, denoting the nontrivial element of $\mathbb{Z}_2$ by $\hat{P}$ we use the homomorphism $\alpha : \mathbb{Z}_2 \to \mathrm{Aut}(\mathrm{Spin}(2))$ defined by

$$\alpha(\hat{P}) : \hat{R}(\hat{\alpha}) \to (R(\hat{\alpha}))^{-1} = \hat{R}(-\hat{\alpha}) \tag{15.195}$$

Then, one definition of the group $\mathrm{Pin}^+(2)$ is that it is the semidirect product:

$$\mathrm{Pin}^+(2) \cong \mathrm{Spin}(2) \rtimes_\alpha \mathbb{Z}_2 \tag{15.196}$$

(We will give a slightly different definition below.) There is a generalization of Spin and Pin groups to higher dimensions. They double cover $SO(d)$ and $O(d)$, respectively. See the remark below for a brief description and Chapters *** and **** for full details.

Now, when $2\mathcal{B}$ is an odd integer the group $\mathcal{G}_B$ is generated by

$$\begin{aligned} & z\mathbf{1}_{\mathcal{H}} \\ \rho(\hat{R}(\hat{\alpha})) &:= e^{-\mathrm{i}(2\mathcal{B})\hat{\alpha}}\mathcal{R}(2\hat{\alpha}) \qquad\qquad 0 \le \hat{\alpha} < 2\pi \\ \rho(\hat{P}) &:= \mathcal{P} \end{aligned} \tag{15.197}$$

where we take $\hat{P}$ to be the nontrivial element in $\mathbb{Z}_2$ in the semidirect product that defines $\mathrm{Pin}^+(2)$, so that $\hat{R}(\hat{\alpha})$ and $\hat{P}$ generate $\mathrm{Pin}^+(2)$. One checks that $\rho$ is a homomorphism and the image under $\rho$ is an isomorphic copy of $\mathrm{Pin}^+(2)$ inside $\mathcal{G}_B$, and we have:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & \mathrm{Pin}^+(2) & \longrightarrow & O(2) & \longrightarrow & 1 \\ & & \downarrow & & \downarrow \rho & & \downarrow Id & & \\ 1 & \longrightarrow & U(1) & \longrightarrow & \mathcal{G}_B & \longrightarrow & O(2) & \longrightarrow & 1 \end{array} \tag{15.198}$$

where $\mathbb{Z}_2 \cong \{\pm 1\} \subset U(1)$. When $2\mathcal{B}$ is odd $\rho$ has no kernel. (When $2\mathcal{B}$ is even there is a kernel.)

In conclusion:

1. The classical theory has an $O(2)$ symmetry.

2. In the quantum theory when $2\mathcal{B}$ is not an integer the symmetry is broken to $SO(2)$.

3. In the quantum theory, when $2\mathcal{B}$ is an even integer the theory still has $O(2)$ symmetry. The sequence

$$1 \longrightarrow U(1) \longrightarrow \mathcal{G}_B \longrightarrow O(2) \longrightarrow 1 \tag{15.199}$$

splits and $\mathcal{G}_B \cong U(1) \times O(2)$.

4. In the quantum theory, when $2\mathcal{B}$ is an odd integer,

$$1 \longrightarrow U(1) \longrightarrow \mathcal{G}_B \longrightarrow O(2) \longrightarrow 1 \tag{15.200}$$

does not split. When we try to realize the classical $O(2)$ symmetry on the Hilbert space we are forced to implement the pin double cover $\mathrm{Pin}^+(2)$, a central extension of $O(2)$ by $\mathbb{Z}_2$. It is related to $\mathcal{G}_{\mathcal{B}}$ as in (15.198).

We conclude with some remarks:

1. We stress that the particle we put on the ring did NOT have any intrinsic spin!! Having said that, if we define an angular momentum $\mathcal{L}$ so that $H = \frac{\mathcal{L}^2}{2I}$ then indeed when $\mathcal{B}$ is half-integral the angular momentum has half-integral eigenvalues, as one expects for a spin representation. So, what we are finding is that the half flux quantum is inducing a half-integral spin of the system so that the classical $O(2)$ symmetry of the classical system is implemented as a $\mathrm{Pin}^+(2)$ symmetry in the quantum theory. This is an intriguing phenomenon appearing in quantum symmetries with nontrivial gauge fields and topological terms: The statistics and spins of particles can be shifted from their classical values, often in ways that involve curious fractions.

2. *Spin And Pin Groups.* Enquiring minds will wonder about the definition of $\mathrm{Pin}^{\pm}(d)$. These groups are defined using Clifford algebras. For much more detail and motivation see the two chapters on Clifford algebras and Spin groups. In brief, consider the Clifford algebra generated by $\{\gamma_i, \gamma_j\} = 2Q_{ij}$ where $Q_{ij}$ is an invertible $d \times d$ symmetric matrix. For a vector $v^i$ define $\gamma(v) := v^i \gamma_i$. Assume that $Q_{ij} = \delta_{ij}$. Then $\mathrm{Pin}^+(d)$ is the group of expressions of the form
$$\pm\gamma(v_1)\cdots\gamma(v_r) \tag{15.201}$$
for any set of $r$ vectors of norm-squared one. The group $\mathrm{Pin}^-(d)$ is similarly defined with $Q_{ij} = -\delta_{ij}$. The group $\mathrm{Spin}(d)$ is the subgroup of such expressions where $r$ is even. The projection $\pi : \mathrm{Pin}^{\pm}(d) \to O(d)$ is defined by the equation:
$$\gamma(\pi(g)\cdot w) = (-1)^r g\gamma(w)g^{-1} \tag{15.202}$$
The key idea here is that
$$-\gamma(v)\gamma(w)\gamma(v)^{-1} = \gamma(R_v(w)) \tag{15.203}$$
where $R_v(w)$ is the reflection of $w$ through the plane orthogonal to $v$, as the reader can easily check in an exercise below. Then use the fact that all elements of $O(d)$ are products of reflections. The restriction to $\mathrm{Spin}(d)$ defines $\pi : \mathrm{Spin}(d) \to SO(d)$. This is a generalization of our standard double-covering $\pi : SU(2) \to SO(3)$. Although $\mathrm{Spin}(3) \cong SU(2)$ for $d > 3$ $\mathrm{Spin}(d)$ is not isomorphic to a unitary or orthogonal group. The difference between $\mathrm{Pin}^+(d)$ and $\mathrm{Pin}^-(d)$ is whether the lift of a reflection will square to $+1$ or $-1$ respectively. As a group $\mathrm{Pin}^-(d)$ is isomorphic to $(\mathrm{Spin}(2) \rtimes \mathbb{Z}_4)/\mathbb{Z}_2$.

3. It is instructive to study the representation of $\mathrm{Pin}^+(2)$ on the two-dimensional space of ground states, $\mathcal{H}_{\mathrm{grnd}}$, when $\mathcal{B} = 1/2$. In this case we can choose the ordered basis $\{\Psi_0, \Psi_1\}$ for $\mathcal{H}_{\mathrm{grnd}}$, and, relative to this basis we have a matrix representation:
$$\begin{aligned} \rho(\widehat{R}(\hat\alpha))|_{\mathcal{H}_{\mathrm{grnd}}} &= \begin{pmatrix} e^{-\mathrm{i}\hat\alpha} & 0 \\ 0 & e^{\mathrm{i}\hat\alpha} \end{pmatrix} \\ \rho(\widehat{P})|_{\mathcal{H}_{\mathrm{grnd}}} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned} \tag{15.204}$$

We stress the appearance of $\hat{\alpha}$ in the representation matrix. This transformation corresponds to a translation of $\phi$ by $\alpha = 2\hat{\alpha}$. Had we tried to express the representation in terms of $\alpha$ we would encounter the phase $e^{\pm i\alpha/2}$ which is not well-defined because $\alpha$ is only defined modulo $\alpha \sim \alpha + 2\pi$.

4. As pointed out in a recent paper [237] this extension of the symmetry group at half-integral $\theta$ is an excellent baby model for how one can learn about nontrivial dynamics of quantum systems (in particular, QCD) by thinking carefully about group extensions. For example, if we were to add a potential $U(\phi)$ to the problem we just discussed we could no longer solve exactly for the eigenstates. Also, a generic potential would be of the form

$$U^{generic}(\phi) = \sum_{n\in\mathbb{Z}} c_n \cos(n\phi) + \sum_{n\in\mathbb{Z}} s_n \sin(n\phi) \tag{15.205}$$

Potentials with generic coefficients will explicitly break all of the $O(2)$ symmetry. Suppose however, that we can restrict attention to a special class of potentials with only cosine Fourier coefficients that are 0 mod 2:

$$U^{special}(\phi) = \sum_{n} u_n \cos(2n\phi) \tag{15.206}$$

Then, even though we cannot solve the spectrum of the Hamiltonian exactly we can make an interesting statement about it. For such potentials the classical $O(2)$ symmetry is explicitly broken to $\mathbb{Z}_2 \times \mathbb{Z}_2$ generated by $P : \phi \to -\phi$ and $r : \phi \to \phi + \pi$. We have shown that when $2\mathcal{B}$ is odd and the potential is zero the $O(2)$ symmetry is centrally extended and realized as the double-cover $\mathrm{Pin}^+(2)$ on the Hilbert space. The double cover of the subgroup group $\langle P, r\rangle \subset O(2)$ acting on Hilbert space is described by the pullback diagram:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & \mathbb{Z}_2 & \xrightarrow{\iota} & D_4 & \xrightarrow{\pi_1} & \mathbb{Z}_2 \times \mathbb{Z}_2 & \longrightarrow & 1 \\ & & \downarrow \mathrm{Id} & & \downarrow \iota & & \downarrow \iota & & \\ 1 & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & \mathrm{Pin}^+(2) & \xrightarrow{\pi_2} & O(2) & \longrightarrow & 1 \end{array} \tag{15.207}$$

To check this note that $P$ lifts to the operators $\pm\mathcal{P}$ and $r = R(\pi)$ lifts to the operators

$$\pm\hat{r} = e^{i\mathcal{B}\pi}\mathcal{R}(\pi) \tag{15.208}$$

One checks that $\langle \mathcal{P}, \hat{r}\rangle$ generates a group isomorphic to $D_4$. Indeed, these operators satisfy the defining relations of $D_4$: $\mathcal{P}^2 = 1$, $\hat{r}^4 = 1$, and $\mathcal{P}\hat{r}\mathcal{P} = \hat{r}^{-1}$. Note that, in addition $\hat{r}^2 = -1$ on the entire Hilbert space. The representation on the Qbit groundstate in (15.204) (with $\hat{\alpha} = \pm\pi/2$) is a two-dimensional irrep of $D_4$. In fact all the doubly degenerate energy eigenspaces are two-dimensional irreps of $D_4$.

[237] D. Gaiotto, A. Kapustin, Z. Komargodski and N. Seiberg, "Theta, Time Reversal, and Temperature," https://arxiv.org/pdf/1703.00501.pdf

It is reasonable to assume that when we turn on a weak potential of the form (15.206) the classically preserved $\mathbb{Z}_2 \times \mathbb{Z}_2$ subgroup again lifts to a $D_4$ action on the Hilbert space, even though we can no longer construct the operators in the $D_4$ group explicitly. The cocycle is discrete: So if it is a continuous function of the parameters $u_n$ at $u_n = 0$ (this is an assumption) then the classical $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry must be realized by $D_4$ on the Hilbert space.

Now, we saw that, in the absence of the potential, there is a Qbit giving a two-dimensional representation of $\mathrm{Pin}^+(2)$. This representation restricts to an irreducible two-dimensional representation of $D_4$. (See equation (15.204) with $\hat{\alpha} = \pm\pi/2$.) Now, $D_4$ has four one-dimensional irreducible representations $\mathbf{1}_{\pm,\pm}$ and, we will show later, exactly one two-dimensional irreducible representation. In particular, the set of representations of $D_4$ is discrete. Again, it is reasonable to suppose that the representation is a continuous function of $u_n$. Again, this is an assumption. But granting this, turning on a weak potential cannot change the decomposition of the energy eigenspaces into irreducible representations. This leads to a striking prediction: The two-fold groundstate degeneracy is not broken by potentials of the form (15.206) when $2\mathcal{B}$ is odd! This is remarkable when one compares to the standard discussion of the double-well potential of one-dimensional quantum mechanics. In that standard case one has a two-fold classical degeneracy broken by tunneling (instanton) effects so that there is a unique ground state. For potentials of the form (15.206) there are (generically) four stationary points of the potential, at $\phi = 0, \pm\pi/2, \pi$. Generically, two will be maxima and two will be minima. So, classically, and perturbatively in quantum mechanics, for a generic potential of the form (15.206) there will be a two-fold degenerate groundstate. However, unlike the textbook discussion of the double-well potential, the degeneracy will not be lifted by nonperturbative tunneling effects.

---

**Exercise**
Show that the ground state energy is

$$E_{\text{ground}} = Min_{m\in\mathbb{Z}} \frac{\hbar^2}{2I}(m - \mathcal{B})^2 \tag{15.209}$$

and, using the floor function, give a formula for $E_{\text{ground}}$ directly in terms of $\mathcal{B}$ (without requiring minimization).

---

---

**Exercise** *Pin Action*
a.) Show that (15.202) defines a homomorphism to $O(d)$. [238]

[238] *Answer*: The key equation to check is (15.203). To check this consider the two cases that $w$ is parallel to $v$ and that $w$ is perpendicular to $v$. Then note that every element in $O(d)$ is a product of reflections.

b.) Show that the general definition of $\mathrm{Pin}^+(d)$ specializes to the definition of $\mathrm{Pin}^+(2)$ as a semidirect product. [239]

---

---

**Exercise** *One-Dimensional Representations Of $D_4$*

Show that there are four distinct one-dimensional representations of $D_4$. [240]

---

---

**Exercise** *A Cocycle Puzzle*

Note that had we defined $\mathcal{P}$ with an extra factor of i we would have concluded that it is order 4, not order 2. Now, we know that the sequence $1 \to \mathbb{Z}_2 \to \mathbb{Z}_4 \to \mathbb{Z}_2 \to 1$ is not split and has a nontrivial cocycle. When then, can we define a parity operation of order two? [241]

---

### 15.4.2 Remarks About The Quantum Statistical Mechanics Of The Particle On The Ring

In quantum statistical mechanics a central object of study is the partition function:

$$Z := \mathrm{Tr}_{\mathcal{H}} e^{-\beta H} \tag{15.211}$$

Here $\beta = 1/(kT)$ where $k$ is Boltzmann's constant and $T$ is the absolute temperature.

For simplicity we will henceforth set $\hbar = k = 1$ (as can always be done by a suitable choice of units).

Since we have diagonalized the Hamiltonian exactly we can immediately say that

$$Z = \sum_{m \in \mathbb{Z}} e^{-\frac{\beta}{2I}(m-\mathcal{B})^2} \tag{15.212}$$

This is in fact a very interesting function of $\frac{\beta}{2I}$ and $\mathcal{B}$. Some immediate facts we can note are

---

[239] *Answer*: First reproduce Spin(2) using the general definition. We can represent the $d = 2$ Clifford algebra by $\sigma^1, \sigma^2$ and then the product of two vectors of the form $\gamma(v)$ with $v^2 = 1$ is a matrix of the form $(x^1\sigma^1 + x^2\sigma^2)(y^1\sigma^1 + y^2\sigma^2)$ where $v_1 = (x^1, x^2)$ and $v_2 = (y^1, y^2)$ are unit vectors in $\mathbb{R}^2$. Multiplying this out we get

$$\cos\theta + \sin\theta\sigma^1\sigma^2 \tag{15.210}$$

where $\theta$ is the angle between $v_1$ and $v_2$. Now let $\hat{P}$ be represented by $\sigma^1$ (or $\gamma(w)$ for any unit vector $w$).

[240] *Answer*: $D_4$ has generators $x, y$ with $x^2 = 1$ and $y^4 = 1$ and $xyx = y^{-1}$. In a one-dimensional representation $x, y$ will be represented by complex numbers. So we solve the above equations with $x, y \in \mathbb{C}$. Clearly $x \in \{\pm 1\}$ and then $y^2 = 1$ so $y \in \{\pm 1\}$ and there is no correlation between the choice of sign for $x$ and the choice of sign for $y$.

[241] *Hint*:It is important to think about which group the cocycle and coboundaries take values in.

1. The expression is manifestly periodic under integer shifts of $\mathcal{B}$, illustrating the general claim above that the theory is invariant under integral shifts of the "theta angle" $\mathcal{B}$.
2. Moreover, at low temperature, $\beta \to \infty$ there is a single dominant term from the sum, unless $\mathcal{B}$ is a half-integer, in which there are two equally dominant terms - this reflects the double degeneracy of the ground state when $2\mathcal{B}$ is odd: The ground state is a Qbit. A standard technique in field theory is to study the IR behavior of a partition function to learn about the ground states of the system.

We are going to see that this system in fact has a very interesting high/low temperature duality and use this to understand better the $\theta$-dependence of the previous example in terms of path integrals.

To relate $Z$ to a path integral we observe that we can write:

$$Z = \int_0^{2\pi} d\phi \langle \phi | e^{-\beta H} | \phi \rangle \tag{15.213}$$

Now, we can interpret $\langle \phi | e^{-\beta H} | \phi \rangle$ as a specialization of the matrix elements of the Euclidean time propagator

$$\langle \phi_2 | e^{-t_E H_{\mathcal{B}}} | \phi_1 \rangle = \frac{1}{2\pi} \sum_{m \in \mathbb{Z}} e^{-\frac{t_E}{2I}(m-\mathcal{B})^2 + \mathrm{i} m(\phi_1 - \phi_2)} \tag{15.214}$$

Indeed, the usual propagator in quantum mechanics is

$$U(t) = e^{-\mathrm{i}tH/\hbar} \tag{15.215}$$

Under good conditions this family of operators for $t \in \mathbb{R}$ has an "analytic continuation" to part of the complex plane. What this means is that there is a well-defined family of operators

$$e^{-\mathrm{i}zH/\hbar} \tag{15.216}$$

where $z$ takes values in a region $\mathcal{R} \subset \mathbb{C}$. The region $\mathcal{R}$ should, at least contain the real axis of time on its boundary (or closure). To see that there might be restrictions on $\mathcal{R}$ suppose that $\mathcal{R} = \mathbb{C}$. Then we can consider the restriction to the imaginary axis, setting

$$z = -\mathrm{i}t_E \qquad t_E \in \mathbb{R} \tag{15.217}$$

Here $t_E$ is called *Euclidean time* because if we were to make a substitution $t \to -\mathrm{i}t_E$ in the Lorentz metric then we would get a metric of definite signature. (If we take signature $(-, +^d)$ we would get the Euclidean metric.) If the Hamiltonian is bounded below $\exp[-t_E H]$ should make sense for $t_E$ positive, but if the spectrum of $H$ grows rapidly the operator will be unbounded and certainly not traceclass for $t_E$ negative. So, for Hamiltonians, such as the one we are considering the region $\mathcal{R}$ can be taken to be the negative half-plane. Defining such an analytic family of operators and restricting to the negative imaginary axis (or some other part of the imaginary axis) is called *Wick rotation.*

Now, it is well-known that the propagator: $\langle \phi_2 | e^{-\frac{\mathrm{i}tH}{\hbar}} | \phi_1 \rangle$ can be represented by a Feynman path integral. After Wick rotation we still have a path integral representation.

Feynman's argument proceeds just as well with $e^{-t_E H/\hbar}$. In fact, formally, it is better since the integral has better (formal) convergence properties for Euclidean actions whose real part is bounded below.

Now, by setting $\phi_1 = \phi_2 = \phi$ and integrating over $\phi$ we are making Euclidean time periodic, with period $\beta$ and computing the path integral on a compact spacetime, namely, the circle. The path integral for $\phi_1 = \phi_2$ is done with boundary conditions on the fields so that $\phi(0) = \phi(\beta)$. This is precisely the kind of boundary condition that says that $\phi(t)$ is defined on a circle. More details on path integrals are available in many textbooks. See, for examples:

1. Feynman and Hibbs, *Quantum Mechanics and Integrals*
2. Feynman, *Statistical Mechanics*
3. C. Itzykson and J.B. Zuber, *Quantum Field Theory*,
4. J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*,

In this subsubsection we will henceforth drop the subscript $E$ on $t_E$ and just use $t$ for the real Euclidean time coordinate.

In the Wick rotation to Euclidean space the "$\theta$-angle" $\oint eA$ remains real so the matrix elements of the Euclidean time propagator have the path integral representation:

$$\langle\phi_2|e^{-\beta H}|\phi_1\rangle = Z(\phi_2, \phi_1|\beta) := \int [d\phi(t)]_{\phi(0)=\phi_1}^{\phi(\beta)=\phi_2} e^{-\frac{1}{\hbar}\int_0^\beta \frac{1}{2} I\dot{\phi}^2 dt - \mathrm{i}\int \mathcal{B}\dot{\phi} dt} \tag{15.218}$$

(One must be careful with the sign of the imaginary term, and it matters.)

Viewed as a field theory, this is a free field theory and the path integral can be done exactly by semiclassical techniques:

The equation of motion is simply $\ddot{\phi} = 0$. Again, the $\theta$-term has not changed it.

Thus, the classical solutions to the equations of motion with boundary conditions $\phi(0) = \phi_1, \phi(\beta) = \phi_2$ are:

$$\phi_{cl}(t) = \phi_1 + \left(\frac{\phi_2 - \phi_1 + 2\pi w}{\beta}\right)t \qquad w \in \mathbb{Z} \tag{15.219}$$

or more to the point:

$$e^{\mathrm{i}\phi_{cl}(t)} = e^{\mathrm{i}\left((1-\frac{t}{\beta})\phi_1 + \frac{t}{\beta}\phi_2\right) + \frac{2\pi \mathrm{i} t w}{\beta}} \tag{15.220}$$

These are solutions of the Euclidean equations of motion, and are known as "instantons" for historical reasons. Notice that because of the compact nature of the spacetime on which we define our $0+1$ dimensional field theory there are infinitely many solutions labeled by $w \in \mathbb{Z}$. There are two circles in the game: The spacetime of this $0+1$-dimensional field theory is the Euclidean time circle. Then the target space of the field theory is also a circle. The quantum field $e^{\mathrm{i}\phi(t)}$ is a map $M \to \mathcal{X}$. $M$, which is spacetime is $S^1$ and $\mathcal{X}$, which is the target space is also $\mathcal{X}$. As we saw: $\pi_1(S^1) \cong \mathbb{Z}$. There can be topologically inequivalent field configurations. That is the space of maps $Map(M \to \mathcal{X})$ has different connected components. The different topological sectors are uniquely labelled by the winding number of the map (15.220). In the path integral we sum over all field configurations so we should sum over all these instanton configurations.

We now write

$$\phi = \phi_{cl} + \phi_q \tag{15.221}$$

where $\phi_{cl}$ is an instanton solution, as in (15.219), and $\phi_q$ is the quantum fluctuation with $\phi_q(0) = \phi_q(\beta) = 0$, and, moreover, $\phi_q(t)$ is in the topologically trivial component of $Map(S^1, \mathcal{X})$. The action is a sum $S(\phi_{cl}) + S(\phi_q)$, precisely because $\phi_{cl}$ solves the equation of motion. Indeed the integral factorizes and we just get:

$$Z(\phi_2, \phi_1|\beta) = Z_q \sum_{w\in\mathbb{Z}} e^{-\frac{2\pi^2 I}{\beta}(w+\frac{\phi_2-\phi_1}{2\pi})^2+2\pi i\mathcal{B}(w+\frac{\phi_2-\phi_1}{2\pi})} \tag{15.222}$$

The summation runs over classical solutions, weighted by the value of the classical action on that solution.

$Z_q$ is the path integral over $\phi_q$:

$$Z_q = \int [d\phi_q(t)]_{\phi_q(0)=0}^{\phi(\beta)=0} e^{-\int_0^\beta \frac{1}{2} I\dot{\phi}^2 dt} \tag{15.223}$$

We are integrating over the space of "all" maps $\phi_q : [0,1] \to U(1)$ with $\phi_q(0) = \phi_q(1) = 0$ that are homotopically trivial. We can do it by noticing that this is a Gaussian integral.

Now in finite dimensions we have the integral

$$\int \prod_{i=1}^n \frac{dx^i}{\sqrt{2\pi}} e^{-\frac{1}{2}x^i A_{ij} x^j + b_i x^i} = \frac{1}{\sqrt{\det A}} e^{\frac{1}{2} b_i (A^{-1})^{ij} b_j} \tag{15.224}$$

where $Re(A) > 0$ is a symmetric matrix. When $A$ can be diagonalized by a real orthogonal transformation we can replace

$$\det A = \prod_{i=1}^n \lambda_i \tag{15.225}$$

where the product runs over the eigenvalues of $A$. Thus, we need to generalize this expression to the determinant of an infinite-dimensional "matrix"

$$\int [d\phi_q] \exp[-\int_0^1 \phi_q (-\frac{I}{2\beta}\frac{d^2}{d\tau^2})\phi_q] = (2\pi) \mathrm{Det}'^{-1/2}(\mathcal{O}) \tag{15.226}$$

Here the prime on the determinant means that we have omitted the zero-mode and the analog of $A$ is the operator $\mathcal{O} = -\frac{I}{2\hbar\beta}\frac{d^2}{d\tau^2}$.

One way to make sense of $\mathrm{Det}\mathcal{O}$ for an operator $\mathcal{O}$ on Hilbert space is known as "$\zeta$-function regularization." (It will only work for a suitable class of operators.) Note that

$$\frac{d}{ds}|_{s=0} \lambda^{-s} = -\log\lambda \tag{15.227}$$

So if we define

$$\zeta_{\mathcal{O}}(s) := \sum_\lambda \lambda^{-s} \tag{15.228}$$

where we take the sum over the spectrum of $\mathcal{O}$ (and we assume $\mathcal{O}$ is diagonalizable with discrete spectrum) then, formally:

$$\prod_\lambda \lambda = \exp[-\zeta'_{\mathcal{O}}(0)] \tag{15.229}$$

For good operators $\mathcal{O}$ the spectrum goes to infinity sufficiently fast that $\zeta_{\mathcal{O}}(s)$ exists as an analytic function of $s$ in a half plane $Re(s) > N$ for some $N$. Moreover, $\zeta_{\mathcal{O}}(s)$ also admits an analytic continuation in $s$ to an open region around $s = 0$. In this case, we can define the determinant by the RHS of (15.229).

For $\mathcal{O} = -\frac{I}{2\hbar\beta}\frac{d^2}{d\tau^2}$ we have $\zeta_{\mathcal{O}}(s) = 2\left(\frac{I\pi^2}{2\hbar\beta}\right)^s\zeta(2s)$ where $\zeta(s)$ is the standard Riemann $\zeta$-function, and since

$$\zeta(s) = -\frac{1}{2} + s\log\left(\frac{1}{\sqrt{2\pi}}\right) + \mathcal{O}(s^2) \tag{15.230}$$

we have

$$\mathrm{Det}'(\mathcal{O}) := \exp[-\zeta'_{\mathcal{O}}(0)] = \frac{\beta}{I} \tag{15.231}$$

(There are some factors of 2 and $\pi$ that need to be fixed in this equation.)

We can understand this result nicely as follows. Let us study the $\beta \to 0$ behavior of the path integral. Then for $|\phi_1 - \phi_2| < \pi$,

$$Z \to Z_q e^{-\frac{I}{2\beta}(\phi_2-\phi_1)^2 + i\mathcal{B}(\phi_2-\phi_1)}\left(1 + \mathcal{O}(e^{-\kappa/\beta})\right) \tag{15.232}$$

where $\kappa > 0$. In plain English: the instantons are only important at *large* $\beta$. This is intuitively very satisfying: At very small times $\beta$ it must cost a lot of action for $\phi(t)$ to make a nonzero number of circuites around the circle because the velocity must then be large, and large velocity means large action. So for physical quantities based on such small fluctuations the topologically nontrivial field configurations must contribute subleading effects. Let us therefore compare $Z(\phi_2, \phi_1|\beta)$ as $\beta \to 0$ with the standard quantum mechanical propagator. For small $\phi$ we can remove the phase from the $B$-field via $\psi(\phi) \to e^{-i\mathcal{B}\phi}\psi(\phi)$, so $Z_q$ should not depend on $\mathcal{B}$. (Note this transformation is not globally defined in $\phi$ for generic $\mathcal{B}$ so we cannot use it to remove $\mathcal{B}$ from the problem when we treat the full quantity $Z$ exactly.) After this transformation we expect to recover the standard propagator of a particle of mass $M = I$ on the line. Rotated to Euclidean space this would be:

$$\sqrt{\frac{M}{2\pi\hbar\beta}}e^{-\frac{M(\phi_2-\phi_1)^2}{2\hbar\beta}} \tag{15.233}$$

so

$$Z_q = \sqrt{\frac{I}{2\pi\hbar\beta}} \tag{15.234}$$

The net result is that

$$Z(\phi_2, \phi_1|\beta) = \sqrt{\frac{I}{2\pi\beta}}\sum_{w\in\mathbb{Z}} e^{-\frac{2\pi^2 I}{\beta}(w+\frac{\phi_2-\phi_1}{2\pi})^2 - 2\pi \mathrm{i}\mathcal{B}(w+\frac{\phi_2-\phi_1}{2\pi})} \tag{15.235}$$

Now compare (15.214) (with $t_E = \beta$) with (15.235). These expressions look very different! One involves a sum of exponentials with $\beta$ in the numerator and the other with $\beta$ in the demoninator. One is well-suited to discussing the asymptotic behavior for $\beta \to \infty$ (low temperature) and the other for $\beta \to 0$ (high temperature), respectively. Nevertheless, we have computed the same physical quantity, just using two different methods. So they must be the same. But the mathematical identity that says they are the same appears somewhat miraculous. We now explain how to verify the two expressions are indeed identical using a direct mathematical argument.

The essential fact is the Poisson summation formula discussed in section

In our case, the expression computed directly from the diagonalization of the Hamiltonian is

$$
\begin{aligned}
Z(\phi_2, \phi_1|\beta) &= \frac{1}{2\pi} \sum_{m\in\mathbb{Z}} e^{-\frac{\beta}{2I}(m-\mathcal{B})^2 + \mathrm{i}m(\phi_1-\phi_2)} \\
&= \frac{1}{2\pi} e^{\mathrm{i}\mathcal{B}(\phi_1-\phi_2)} \vartheta[{\theta \atop \phi}](0|\tau)
\end{aligned}
\tag{15.236}
$$

where we have written it in terms of the Riemann theta function (11.503) with:

$$
\begin{aligned}
\tau &= \mathrm{i}\frac{\beta}{2\pi I} \\
\theta &= -\mathcal{B} \\
\phi &= \frac{\phi_1 - \phi_2}{2\pi}
\end{aligned}
\tag{15.237}
$$

On the other hand, the expression that emerges naturally from the semiclassical evaluation of the Euclidean path integral is

$$
\begin{aligned}
Z(\phi_2, \phi_1|\beta) &= \sqrt{\frac{I}{2\pi\beta}} \sum_{w\in\mathbb{Z}} e^{-\frac{2\pi^2 I}{\beta}(w+\frac{\phi_2-\phi_1}{2\pi})^2 - 2\pi\mathrm{i}\mathcal{B}(w+\frac{\phi_2-\phi_1}{2\pi})} \\
&= \sqrt{\frac{I}{2\pi\beta}} \vartheta[{\theta' \atop \phi'}](0|\tau')
\end{aligned}
\tag{15.238}
$$

where we have written it in terms of the Riemann theta function (**??**):

$$
\begin{aligned}
\tau' &= \mathrm{i}\frac{2\pi I}{\beta} \\
\theta' &= -\frac{\phi_1 - \phi_2}{2\pi} \\
\phi' &= -\mathcal{B}
\end{aligned}
\tag{15.239}
$$

Note that the modular transformation law of the Riemann theta function, equation (**??**) is equivalent to the relation between the expressions naturally arising from the Hamiltonian and Lagrangian approaches to evaluation of the matrix elements of the Euclidean time propagator!

Note in particular that for the partition function proper we have, as $\beta \to 0$:

$$
\begin{aligned}
Z(S^1) &= \left(\frac{2\pi I}{\beta\hbar}\right)^{1/2} \sum_{n\in\mathbb{Z}} e^{-2\pi^2 n^2 \frac{I}{\beta\hbar} + 2\pi\mathrm{i}n\mathcal{B}} \\
&\sim_{\beta\to 0} \left(\frac{2\pi I}{\beta\hbar}\right)^{1/2} \left(1 + 2e^{-2\pi^2\frac{I}{\beta\hbar}}\cos(2\pi\mathcal{B}) + \cdots\right)
\end{aligned}
\tag{15.240}
$$

The overall factor of $\beta^{-1/2}$ gives the expected divergence. The first correction term to the factor in parentheses is an instanton effect.

Note that in the Hamiltonian version the only thing that is manifest about the high-temperature, $\beta \to 0$, limit is that $Z$ diverges. Note that for $\beta \to 0$ all the terms in the sum contribute about equally and the sum diverges. The modular transformation reveals an interesting duality: Once we factor out this multiplicative divergence we discover another theta function.

---

**Exercise** *A Parity Puzzle*
Suppose $2\mathcal{B}$ is an odd integer.
a.) Show that as $\beta \to \infty$ we have

$$\langle \phi_2 | e^{-\frac{\beta H}{\hbar}} | \phi_1 \rangle \sim 2 e^{\frac{\mathrm{i}}{2}(\phi_1 - \phi_2)} \cos\left(\frac{\phi_1 - \phi_2}{2}\right) e^{-\beta E_{\mathrm{ground}}} + \cdots \tag{15.241}$$

b.) Note that this expression is not invariant under $\phi \to -\phi$. But in $\mathrm{Pin}^+(2)$ there is an element $P$ which corresponds to $\phi \to -\phi$. How is this compatible with our argument that $\mathrm{Pin}^+(2)$ is a valid symmetry of the quantum theory? [242]

---

### 15.4.3 Gauging The Global $SO(2)$ Symmetry, Chern-Simons Terms, And Anomalies

When a theory has a symmetry one can implement a procedure called "gauging the symmetry." This is a two-step process:

1. Make the symmetry local and couple to a gauge field.
2. Integrate over "all possible" gauge fields consistent with the symmetry.

It is not necessary to proceed to step (2) after completing step (1). In this case, we say that we are coupling to nondynamical external gauge fields. It makes perfectly good sense to introduce nondynamical, external gauge fields for a symmetry. We do this all the time in quantum mechanics courses where we couple our quantum system to an electromagnetic field, but do not try to quantize the electromagnetic field.

For the more mathematically sophisticated reader the two-step process can be summarized, somewhat more concisely and precisely, as saying that we:

1. Identify the symmetry group with the structure group of a principal bundle and we change the bordism category in the domain of the field theory functor to include $G$-bundles with connection (where $G$ is the symmetry we are gauging).

[242] *Answer*: Show that

$$\mathcal{P} \cdot |\phi\rangle = e^{-\mathrm{i}\phi} | -\phi\rangle \tag{15.242}$$

You can prove this by expanding $|\phi\rangle = \sum_{m\in\mathbb{Z}} \langle \Psi_m | \phi \rangle \Psi_m$. Now, using this expression check that the propagator indeed transforms correctly.

2. Sum over isomorphism classes of principal bundles and integrate over the isomorphism classes of connections on those bundles.

In the present simple example of the charged particle on a ring surrounding a solenoid we can "gauge" the global $SO(2)$ symmetry $\phi \to \phi + \alpha$ that is present for all values of $\mathcal{B}$. It is then interesting to see how coupling to the external gauge field tells us about the subtleties of combining $SO(2)$ symmetry with charge conjugation symmetry that we studied above. (The following discussion was inspired by Appendix D of. [243] )

So in our simple example we implement Step 1 above as follows: We seek to make the shift symmetry local, that is, we attempt to make

$$\phi(t) \to \phi(t) + \alpha(t) \tag{15.243}$$

into a symmetry where $\alpha(t)$ is not a constant but an "arbitrary" function of time. When $\alpha(t)$ is time dependent the action $\sim \int \dot{\phi}^2$ is not invariant under such transformations. To compensate for this we introduce an extra function of time into the problem, call it $A^{(e)}(t)$. Here the superscript $e$ - for "external" - reminds us that this is an "external" or "background" field: We will <u>not</u> do a path integral over these functions (unless we proceed to Step 2 above). By contrast, we <u>will</u> do a path integral over the "dynamical" field $\phi(t)$ or, equivalently, over $\Phi(t) = e^{i\phi(t)}$.

The gauged action is

$$\begin{aligned} S &= \int \frac{1}{2} I(\dot{\phi} + A^{(e)})^2 dt + \oint \mathcal{B}(\dot{\phi} + A^{(e)}) dt \\ &= \int \frac{1}{2} I \left( \Phi(t)^{-1}(-i\frac{d}{dt} + A^{(e)})\Phi(t) \right)^2 dt + \oint \mathcal{B}\Phi(t)^{-1}(-i\frac{d}{dt} + A^{(e)})\Phi(t) dt \end{aligned} \tag{15.244}$$

This action is a functional of both the nondynamical field $A^{(e)}(t)$ and the dynamical field $\phi(t)$. Note that the action is invariant under the gauge transformation:

$$\begin{aligned} \phi(t) &\to \phi(t) + \alpha(t) \\ A^{(e)}(t) &\to A^{(e)}(t) - \partial_t \alpha(t) \end{aligned} \tag{15.245}$$

where, for the moment, we ignore boundary terms.

It turns out that it is better to regard the function $A^{(e)}(t)$ as a component of a 1-form:

$$A^{(e)} := A^{(e)}(t) dt \tag{15.246}$$

and, better still, $A^{(e)}$ is the local one-form associated to a connection on a (locally trivialized) principal $SO(2)$ bundle over the time manifold $M$. We stress that the gauge field $A^{(e)}$ is <u>NOT</u> the gauge field of electromagnetism. (That field has already produced our theta term.) Rather, it is a new field in our system: It is a gauge field for the shift symmetry of the field $\phi(t)$. [244]

[243] D. Gaiotto, A. Kapustin, Z. Komargodski and N. Seiberg, "Theta, Time Reversal, and Temperature," https://arxiv.org/pdf/1703.00501.pdf

[244] The physical interpretation of the gauging process in terms of our original charged particle on a ring is not completely clear to the author. But the mathematical structure makes sense and is a good toy model for other field theoretic systems where the physical interpretation of the gauging is clear.

Note that we could also write the gauge transformation in the form:

$$
\begin{aligned}
e^{\mathrm{i}\phi(t)} &\to e^{\mathrm{i}\phi(t)}e^{\mathrm{i}\alpha(t)} \\
d - \mathrm{i}A^{(e)} &\to e^{\mathrm{i}\alpha(t)}(d + \mathrm{i}A^{(e)})e^{-\mathrm{i}\alpha(t)}
\end{aligned}
\tag{15.247}
$$

This is better because is captures better the geometric content, Consequently, it makes more sense when working on topologically nontrivial spacetimes, such as the Euclidean time circle. One can make choices of gauge group so that it becomes important that $e^{\mathrm{i}\alpha(t)}$ can be single-valued even when $\alpha(t)$ is not.

What about charge conjugation symmetry?

First consider the classical theory. In the absence of the external gauge field we noted that there is an $O(2)$ symmetry of the equations of motion, even though under $\phi(t) \to -\phi(t)$ the theta term in the action flips sign. In the presence of the external gauge field the equations of motion are modified, however, as we will see below, we can gauge $A^{(e)}(t)$ to be a constant, and in this case they are not modified. So we still have an $O(2)$ symmetry.

Now consider the quantum theory. One can show that, appropriately defined, the quantum Hamiltonian is still $H_{\mathcal{B}}$. Under charge conjugation we must flip $\mathcal{B}$ and then we change the Hamiltonian (unless $\mathcal{B} = 0$). But, as noted above, if $2\mathcal{B} \in \mathbb{Z}$ in that case $H_{\mathcal{B}}$ is unitarily equivalent to $H_{-\mathcal{B}}$ and we can implement a unitary operator $\mathcal{P}$ corresponding to the charge conjugation operation. In the path integral the value of the action matters. The action (15.244) is invariant under the charge conjugation transformation if we take

$$
\begin{aligned}
\phi(t) &\to -\phi(t) \\
A^{(e)} &\to -A^{(e)} \\
\mathcal{B} &\to -\mathcal{B}
\end{aligned}
\tag{15.248}
$$

and consequently if we change $\mathcal{B} \to -\mathcal{B}$ we must also take $A^{(e)} \to -A^{(e)}$ as noted above. We will return to the quantum implementation of charge conjugation symmetry.

Now let us re-examine the periodicity of the physics as a function of $\mathcal{B}$. In the absence of the external gauge field $A^{(e)}$ we found that physical quantities are periodic functions of $\mathcal{B}$ with period one. However, in the presence of a nonzero $A^{(e)}$, the term $\int \mathcal{B}A^{(e)}(t)dt$ spoils the periodicity in $\mathcal{B}$, because the value of the action matters in the quantum theory.

We can restore a kind of periodicity in $\mathcal{B}$ by adding a <u>Chern-Simons term</u> to the action. We will comment in detail on the Chern-Simons term below. In Euclidean space the new action is:

$$
e^{-S} = e^{-\int \frac{1}{2}I(\dot{\phi}+A_t^{(e)})^2 dt - \mathrm{i}\oint \mathcal{B}(\dot{\phi}+A_t^{(e)})dt} e^{\mathrm{i}k\int A_t^{(e)}dt} \tag{15.249}
$$

and the last factor is the Chern-Simons term. By introducing the Chern-Simons term we have introduced yet another parameter, the level $k$, into our theory. Classically the action with $(\mathcal{B}, k)$ is equivalent to the action with $(\mathcal{B} + r, k + r)$ where $r \in \mathbb{R}$ is any real number.

Now that we have restored some kind of periodicity we can ask about quantum implementation of charge conjugation symmetry. We must take $\mathcal{B} \to -\mathcal{B}$, but quantum mechanically the theory with $\mathcal{B}$ is only equivalent to that with $-\mathcal{B}$ when $2\mathcal{B} \in \mathbb{Z}$. So, in the quantum theory we can only hope to have charge conjugation invariance if there is an integer $N$ so that

$$(\mathcal{B} + N, k + N) = (-\mathcal{B}, -k) \tag{15.250}$$

In other words $k = \mathcal{B} = N/2 \in \mathbb{Z}/2$.

The introduction of the Chern-Simons term raises a new issue: When one sees gauge potentials in an action that do not enter through fieldstrengths or covariant derivatives it is important to ask about gauge invariance. In order to discuss the gauge invariance of the Chern-Simons term properly we need first to discuss more carefully the space of gauge fields and the group of gauge transformations.

## The Space Of Gauge Fields And The Group Of Gauge Transformations

In our simple setting the space $\mathcal{A}$ of gauge fields can be identified with the space of single-valued, continuous real-valued functions $A^{(e)}(t)$ on $M$.

We now must choose a *gauge group* $G$. This will be a Lie group - typically finite-dimensional, although not necessarily connected. In our case, there are two natural choices: We could take $G = \mathbb{R}$ or we could take $G = U(1)$. Then the *group of gauge transformations* is a group of maps

$$\mathcal{G} = Map[M \to G] \tag{15.251}$$

If $M$ and $G$ have positive dimension the group of gauge transformations will be an infinite-dimensional Lie group.

Of particular interest in gauge theory is the quotient space $\mathcal{A}/\mathcal{G}$, the space of gauge orbits, or, equivalently, the space of gauge-inequivalent field configurations.

Let us examine a few examples of $\mathcal{A}/\mathcal{G}$:

1. Let us first consider what happens when $G = \mathbb{R}$ and $M$ is an interval or the real line. Then the space $\mathcal{A}$ of gauge fields can be identified with the space of real-valued continuous functions on $M$. The group $\mathcal{G}$ is the space of real-valued $C^1$ functions on $M$, $t \mapsto \alpha(t) \in \mathbb{R}$. The group $\mathcal{G}$ acts on $\mathcal{A}$ via

$$A^{(e)}(t) \to A^{(e)}(t) - \partial_t \alpha(t) \tag{15.252}$$

If $M = [t_1, t_2]$ with free boundary conditions on $\mathcal{A}$ and $\mathcal{G}$ then we can always solve

$$\partial_t \alpha(t) = A^{(e)}(t) \tag{15.253}$$

for some $\alpha(t)$ and hence we can always gauge $A^{(e)}(t)$ to zero. So $\mathcal{A}/\mathcal{G}$ is just a point.[245]

For the discussion below it is useful to note here that the expression

$$\exp[\mathrm{i} \int_{t_1}^{t_2} A_t^{(e)}(t')dt'] \tag{15.254}$$

[245] Actually, this is too naive: $\mathcal{A}/\mathcal{G}$ is more properly thought of as a stack. See the section below on groupoids.

is in general not gauge invariant. Rather:

$$\exp[\mathrm{i}\int_{t_1}^{t_2} A_t^{(e)}(t')dt'] \to e^{\mathrm{i}\alpha(t_1)}\exp[\mathrm{i}\int_{t_1}^{t_2} A_t^{(e)}(t')dt']e^{-\mathrm{i}\alpha(t_2)} \tag{15.255}$$

2. Now, continuing to take $G = \mathbb{R}$ let us consider what happens if $M = \mathbb{R}$ and we impose boundary conditions that $\alpha(t) \to 0$ at $t \to \pm\infty$. In this case

$$\int_{-\infty}^{+\infty} A^{(e)}(t)dt \tag{15.256}$$

is gauge invariant. There are no "local" invariants. From the previous discussion we see that we can gauge $A^{(e)}(t)$ to zero in any compact region. In this case

$$\mathcal{A}/\mathcal{G} \cong \mathbb{R} \tag{15.257}$$

and the integral (15.256) fully determines the gauge equivalence class.

3. Now let us consider the case $G = U(1)$, and let us also take $M$ to be the Euclidean time circle so

$$\mathcal{G} = \mathrm{Map}(S^1_{s.t.} \to U(1)) \tag{15.258}$$

Now, just viewing the gauge transformation as a set of continuous maps $S^1 \to S^1$ there is a winding number. If this winding number is nonzero there is an obstruction to finding a single-valued function $\alpha(t)$ so that $g(t) = e^{\mathrm{i}\alpha(t)}$.

There is a normal subgroup $\mathcal{G}_0$ of *small gauge transformations* for which $g(t)$ admits a well-defined logarithm. That is, the gauge transformations $g(t) \in \mathcal{G}_0$ are of the form $g(t) = e^{\mathrm{i}\alpha(t)}$ where $\alpha(t)$ is single-valued. Then

$$1 \to \mathcal{G}_0 \to \mathcal{G} \quad \stackrel{\pi}{\to} \quad \mathbb{Z} \to 1 \tag{15.259}$$

where the map $\pi$ can be viewed as the winding number. Gauge transformations in $\mathcal{G}_0$ are known as *small gauge transformations*. Those which have nonzero winding numbers are known as *large gauge transformations*. It is worth noting that the above sequence splits: For $w \in \mathbb{Z}$ we can take $s(w) = g_w$ to be the gauge transformation:

$$g_w(t) = \exp[2\pi \mathrm{i} wt/\beta] \tag{15.260}$$

and $g_w(t)g_{w'}(t) = g_{w+w'}(t)$. Note that for these transformations if we tried to define $\alpha(t)$ it would be $\alpha(t) = 2\pi wt/\beta$ and would not be single-valued when $M$ is the circle.

Now, referring to (15.263) it is clear that if $\alpha(t)$ is single valued then

$$\exp[\mathrm{i}\oint_{S^1} A_t^{(e)}(t')dt'] \tag{15.261}$$

is gauge invariant. However when we have a large gauge transformation $g_w(t)$ we can cut the circle say at $t = 0$ and $t = \beta$ and then, with $\alpha(t) = 2\pi wt/\beta$ with $w \in \mathbb{Z}$ the holonomy is still gauge invariant. Note that the large gauge transformations $g_w(t)$ take

$$A_t^{(e)}(t') \to A_t^{(e)}(t') + w/\beta \tag{15.262}$$

but preserves the holonomy (15.261). Put differently, (15.263) is generalized to

$$\exp[\mathrm{i}\int_{t_1}^{t_2} A_t^{(e)}(t')dt'] \to g(t_2)^{-1}\exp[\mathrm{i}\int_{t_1}^{t_2} A_t^{(e)}(t')dt']g(t_1) \tag{15.263}$$

so that on the circle the *holonomy* (15.261) is gauge invariant. Equation (15.263) generalizes nicely to the case of nonabelian groups on arbitrary spacetimes.

We can ask if there are other independent gauge invariant functions of $A^{(e)}(t)$ besides the holonomy. Since $A^{(e)}(t)$ is periodic we can decompose $A^{(e)}(t)$ in a Fourier expansion. Write $\tilde{A}_t^{(e)}(t)$ for the sum of the nonzero frequency modes. Then we can solve the differential equation

$$\partial_t \alpha(t) = \tilde{A}_t^{(e)}(t) \tag{15.264}$$

with a single-valued $\alpha(t)$ to choose a gauge so that

$$A^{(e)}(t) = \mu/\beta \tag{15.265}$$

is constant. Put differently, $\mathcal{A}/\mathcal{G}_0$ can be identified with the space of real numbers, given by the constant $\mu$. We will denote $\mathcal{A}^{\mathrm{red}} = \mathcal{A}/\mathcal{G}_0$. Then the "large" gauge transformations $g_w(t)$ shift $\mu \to \mu + 2\pi w$, with $w \in \mathbb{Z}$. Therefore we have

$$\mathcal{A}/\mathcal{G} \cong \mathcal{A}^{\mathrm{red}}/\mathbb{Z} \cong \{[\mu] = [\mu + 2\pi w] \qquad w \in \mathbb{Z}\} \cong U(1) \tag{15.266}$$

and the holonomy $e^{\mathrm{i}\mu}$ around the circle is a complete gauge invariant.

4. Finally, let us take the gauge group to be $G = O(2) = SO(2) \rtimes \mathbb{Z}_2$. Let $\mathcal{G}_{SO(2)}$ be the group of gauge transformations when the gauge group is $SO(2)$ and $\mathcal{G}_{O(2)}$ be the group of gauge transformations when the gauge group is $O(2)$. Then $\mathcal{G}_{O(2)} = \mathcal{G}_{SO(2)} \rtimes \mathbb{Z}_2$ and we have the exact sequence

$$1 \to \mathcal{G}_0 \to \mathcal{G}_{O(2)} \to \mathbb{Z} \rtimes \mathbb{Z}_2 \cong D_\infty \to 1 \tag{15.267}$$

Again the sequence splits and the infinite dihedral group $D_\infty \cong \mathbb{Z} \rtimes \mathbb{Z}_2$ preserves the space $\mathcal{A}^{\mathrm{red}}$ of constant gauge fields with generators acting as

$$\begin{aligned} \sigma :& \mu \to -\mu \\ s :& \mu \to \mu + 2\pi \end{aligned} \tag{15.268}$$

Comments On The Chern-Simons Term

Now we are ready to discuss the gauge invariance of the Chern-Simons term. The Chern-Simons term on $M = S^1$ is invariant under $\mathcal{G}_0$ for any value of $k$. However, under the large gauge transformations $g_w(t)$ with $w \neq 0$:

$$\exp[\mathrm{i}k \oint_{S^1} A_t^{(e)}(t')dt'] \to e^{2\pi \mathrm{i} wk} \exp[\mathrm{i}k \oint_{S^1} A_t^{(e)}(t')dt'] \tag{15.269}$$

and therefore, if we are going to allow our theory to make sense on a circle with the gauge group $SO(2) \cong U(1)$ then $k$ should be quantized to be an integer. Note there would be no such quantization of $k$ if the gauge group is taken to be $\mathbb{R}$.

The above observation is related to two extremely important conceptual points that are essential to all discussions of the use of Chern-Simons terms in quantum physics:

> *It is not necessary for the action to be invariant. All that is necessary for a well-defined path integral is that the exponentiated action must be invariant.*

> *Gauge invariance of the "Chern-Simons term" under large gauge transformations implies that the level $k$, one of the couplings of the theory, is quantized: $k \in \mathbb{Z}$.*

In our case the action in equation (15.249) is <u>not</u> gauge invariant! Under large gauge transformations with winding number $w$ we have $S \to S + 2\pi \mathrm{i} kw$ with $w \in \mathbb{Z}$. However, having a well-defined measure in the path integral only requires $e^{-S}$ to be well-defined, and this will be the case if, and only if, $k \in \mathbb{Z}$.

Note that

$$\exp[\mathrm{i}k \oint_{S^1} A_t^{(e)}(t')dt'] = e^{\mathrm{i}k\mu} \tag{15.270}$$

and, since $k \in \mathbb{Z}$ is quantized this is properly periodic under $\mu \to \mu + 2\pi$ and the exponentiated Chern-Simons term descends to a well-defined function on $\mathcal{A}/\mathcal{G}$. For the group $O(2)$ we would have to consider $\cos(\mu)$. (Of course, $\cos(n\mu)$ is a Tchebyshev polynomial of the basic invariant $\cos(\mu)$.)

<u>Anomalies</u>

We can now discuss, very generally, the notion of anomalies. In quantum systems we typically have both "dyanmical variables" such as dynamical fields, degrees of freedom, etc. as well as "external" or "background" or "control" variables. We will denote generic "background fields" by $\phi^{bck}$ and generic "dynamical fields" by $\phi^{dyn}$. Any parameter of the theory should be considered a "field." The space of all fields is then fibered:

$$\begin{array}{ccc} \mathcal{F}^{dyn} & \longrightarrow & \mathcal{F} \\ & & \downarrow \\ & & \mathcal{F}^{bck} \end{array} \tag{15.271}$$

In the very simple situation we are discussing here the fibration is just a Cartesian product.

In the computation of physical quantities we will typically integrate over $\mathcal{F}^{dyn}$ thus producing a function (or, more generally, a section of a bundle) on $\mathcal{F}^{bck}$. We then study the physical quantity as a function on $\mathcal{F}^{bck}$.

In our example $\mathcal{F}^{dyn}$ can be taken to be the set of functions $\Phi(t) : M \to U(1)$ and $\mathcal{F}^{bck}$ can be taken to be the set of functions $A^{(e)}(t)$, or better, the connections on a principal $G$-bundle over $M$ where in our present examples $G = \mathbb{R}, SO(2)$ or $O(2)$. [246]

Now suppose that there is a group $\mathcal{G}$ acting on $\mathcal{F}$ so that physical quantities are formally invariant. For example, if we have an invariant action $S[\phi^{dyn}; \phi^{bck}]$ and a formally invariant measure, then the path integral will be formally invariant. Then, physical quantities such as the partition function:

$$Z[\phi^{bck}] = \int_{\mathcal{F}^{dyn}} e^{-S[\phi^{dyn};\phi^{bck}]} \mathrm{vol}\,(\phi^{dyn}) \tag{15.272}$$

will, formally define a $\mathcal{G}$-invariant function on $\mathcal{F}^{bck}$. However, it can happen that when one defines the path integral carefully the partition function fails to be $\mathcal{G}$-invariant. In that case we say that there is a *potential anomaly.* Sometimes potential anomalies can be removed by physically unimportant redefinitions. When this cannot be done we say there is an *anomaly.*

If we tried to consider the Chern-Simons term for $k \notin \mathbb{Z}$ we would say it has an *anomaly.* For any value of $k$ it descends to a function on $\mathcal{A}/\mathcal{G}_0$. However, only when $k \in \mathbb{Z}$ does it desend to a well-defined function on $\mathcal{A}/\mathcal{G}$.

It can happen that there can be different subgroups $\mathcal{H}_1 \subset \mathcal{G}$ and $\mathcal{H}_2 \subset \mathcal{G}$ such that there are different definitions of the path integral so that it is invariant either under $\mathcal{H}_1$ or under $\mathcal{H}_2$ but there is no definition so that it is invariant simultaneously under both $\mathcal{H}_1$ and $\mathcal{H}_2$. In this case we say there is a *mixed anomaly.*

The Partition Function As Function On $\mathcal{A}$ And Its Behavior Under The Action Of $\mathcal{G}$

Let us now illustrate some of the above ideas about anomalies by examining the partition function in our example of the gauged particle on a ring.

There will not be any interesting anomalies under $\mathcal{G}_0$. As we have explained we can always use $\mathcal{G}_0$ to gauge $A^{(e)}$ to be a constant 1-form, and we will henceforth take our gauge field to be constant. Then the equation of motion is the same as before, and performing the path integral just as in the previous section we find

$$\begin{aligned} Z(\mu) &= e^{\mathrm{i}k\mu} Z_q \sum_{w\in\mathbb{Z}} e^{-\frac{2\pi^2 I}{\beta}(w+\frac{\mu}{2\pi})^2 - 2\pi \mathrm{i}\mathcal{B}(w+\frac{\mu}{2\pi})} \\ &= e^{\mathrm{i}k\mu} Z_q \vartheta[{\mu/2\pi \atop -\mathcal{B}}](0|\tau) \end{aligned} \tag{15.273}$$

[246] One can also promote $\mathcal{B}$ to be a field for fun and profit. This has been discussed in many places. See e-Print: 1905.09315 for a recent discussion.

with $\tau = \mathrm{i}\frac{2\pi I}{\beta}$. All we need to do here is replace the value of the classical action for solutions with $\dot{\phi} = 2\pi w/\beta$ by making the substitution $w \to w + \mu/2\pi$.

As in the case without the external gauge field there is a Hamiltonian interpretation. Performing the Poisson summation (or using the modular transformation law of the theta function) we get:

$$Z(\mu) = e^{\mathrm{i}(k-\mathcal{B})\mu} \frac{1}{2\pi} \sum_{m\in\mathbb{Z}} e^{-\frac{\beta}{2I}(m-\mathcal{B})^2 - \mathrm{i}(m-\mathcal{B})\mu} \tag{15.274}$$

It can be shown that the Euclidean path integral with action (15.161) is in fact equal to

♣NEED TO RESOLVE SIGN PROBLEM. SHOULD WE HAVE $e^{-\mathrm{i}\mu Q}$ IN THE TRACE? Maybe change the sign in the definition of $Q$? ♣

$$Z(\mu) = e^{\mathrm{i}k\mu}\mathrm{Tr}e^{-\beta H_{\mathcal{B}}} e^{\mathrm{i}\mu Q} \tag{15.275}$$

where $Q$ is the operator measuring the charge of the $SO(2)$ symmetry we gauged. In our case $Q\Psi_m = m\Psi_m$. [247] As noted above we still have

$$H_{\mathcal{B}} = \frac{1}{2I}\left(-\mathrm{i}\frac{\partial}{\partial\phi} - \mathcal{B}\right)^2 \tag{15.276}$$

acting on $L^2(\mathcal{X}) = L^2(S^1)$. One easily checks the equality of (15.274) and (15.275).

From either point of view $Z(\mu)$ is a periodic function of $\mu$ and there is no anomaly under the group $\mathbb{Z}$ of large gauge transformations, so long as $k \in \mathbb{Z}$.

## The Gauge Group $O(2)$ and Mixed Anomalies

What happens if we try to extend the gauge group to gauge the full $O(2)$? Then, as we have seen, the quotient group $\mathcal{G}/\mathcal{G}_0$ is the infinite dihedral group generated by $\sigma$ and $s$ defined in equation (15.268) above.

If $2\mathcal{B}$ is even then we can take $k = \mathcal{B} \in \mathbb{Z}$. The partition function is invariant under $\mu \to -\mu$ and has the expected periodicity $\mu \sim \mu + 2\pi$. In other words $Z(\mu)$ is invariant under the group of large gauge transformations isomorphic to $D_\infty$ and generated by $\sigma$ and $s$, so it descends to a function on $\mathcal{A}/\mathcal{G}$ and there is no anomaly.

Things are much more subtle when $2\mathcal{B}$ is odd. As we saw, we can only expect charge conjugation symmetry when

$$k = \mathcal{B} \in \mathbb{Z} + \frac{1}{2} \tag{15.277}$$

But this clashes with the constraint $k \in \mathbb{Z}$. So we see an example of a mixed anomaly.

It is interesting to see how the mixed anomaly is manifested in the partition function. The main point can be seen most easily by considering the leading term in the $\beta \to \infty$ expansion which is (taking $\mathcal{B} = 1/2$ for simplicity):

$$Z \to \frac{e^{-\beta E_{ground}}}{2\pi} e^{\mathrm{i}(k-\frac{1}{2})\mu}\left(e^{\mathrm{i}\mu/2} + e^{-\mathrm{i}\mu/2}\right) + \cdots \tag{15.278}$$

[247]To give a first-principles proof of why this should be so we gauge away $A^{(e)}$ in the path integral. The result is an identification of the fields at $t = 0$ with the fields at $t = \beta$ accompanied by a gauge transformation by the holonomy $e^{\mathrm{i}\mu}$ which takes $e^{\mathrm{i}\phi} \to e^{\mathrm{i}(\phi+\mu)}$. So $\Psi_m \to e^{\mathrm{i}m\mu}\Psi_m = e^{\mathrm{i}\mu Q}\Psi_m$.

If $k = 0$ then

$$Z \to \frac{e^{-\beta E_{ground}}}{2\pi}\left(1 + e^{-\mathrm{i}\mu}\right) + \cdots \tag{15.279}$$

the expression is properly periodic in $\mu$, but not invariant under the analog of charge conjugation: $\mu \to -\mu$. This is not surprising since $k \neq \mathcal{B}$.

As we will discuss below, by changing the physical system (yet again!) there is a way to make sense of the half-integral level Chern-Simons term. If we just go ahead and recklessly substitute $k = \mathcal{B} = 1/2$ in the above formula for $Z(\mu)$ we get:

$$Z \to \frac{e^{-\beta E_{ground}}}{2\pi}\left(e^{\mathrm{i}\mu/2} + e^{-\mathrm{i}\mu/2}\right) + \cdots \tag{15.280}$$

The action is now invariant under the generator $\sigma : \mu \to -\mu$ of $D_\infty$ but it is no longer invariant under the generator $s : \mu \to \mu + 2\pi$.

Provided we view the different choices of Chern-Simons terms as different definitions of the theory, we can define the theory to be invariant under the group generated by $s$, but with that definition $\sigma$ is anomalous, or, by making a suspicious choice of $k$, we can define the theory to be invariant under the group generated by $\sigma$, but then with that definition $s$ is anomalous. So in this sense there is a mixed anomaly of $\mathbb{Z}$ and $\mathbb{Z}_2$ in the $D_\infty$ subgroup of global gauge transformations.

Making sense of Chern-Simons terms with half-integer level

There is a way to make sense of the half-integer quantized Chern-Simons term by viewing the $0+1$ dimensional theory as the boundary of a well-defined $1+1$ dimensional theory. By Stokes' theorem we have:

$$\exp[ik \oint_{S^1} A_t^{(e)} dt] = \exp[\mathrm{i}k \int_\Sigma F^{(e)}] \tag{15.281}$$

where $F^{(e)} = dA^{(e)}$. The RHS makes sense even if $k$ is not an integer, but now the expression depends on details of the gauge field in the "bulk" of the $1+1$ dimensional spacetime $\Sigma$.

A very analogous phenomenon is observed in real condensed matter systems where the boundary theory of a 3+1 dimensional topological insulator is described by a Chern-Simons theory with half-integral level. (That is, half the level allowed by naive gauge invariance.)

Such half-integral level Chern-Simons terms come up in many interesting physical systems. For example, half-integral (spin) Chern-Simons theory is needed to describe the topological features of the fractional quantum Hall effect. In supersymmetric field theories and string theories many of the supergravity effective actions and brane effective actions involve half-integrally quantized Chern-Simons terms.

---

**Exercise** *Puzzle*

Warning: This exercise requires some knowledge of topology.

Resolve the following paradox:

We first argued that, if $k \notin \mathbb{Z}$ then the LHS of (15.281) is not invariant under large gauge transformations. Then we proceeded to define the LHS by the expression on the RHS which is manifestly gauge invariant.

How can these two statements be compatible? [248]

---

### 15.5 Heisenberg Extensions

Consider again a central extension

$$1 \to A \to \widetilde{G} \to G \to 1 \tag{15.282}$$

In many of the examples above we had $G$ Abelian and $\widetilde{G}$ was also Abelian. However, as our examples with $Q$ and $D_4$ have shown, in general $\widetilde{G}$ need not be Abelian. (See equation (15.145).) In this section we focus on an important class of examples where $G$ is Abelian and $\widetilde{G}$ is non-Abelian. They are known as *Heisenberg groups* and *Heisenberg extensions*. In fact in the literature closely related but slightly different things are meant by "Heisenberg extensions" and "Heisenberg groups." These kinds of extensions show up all the time in physics, in many different ways. They are very basic in quantum field theory and other areas of physics, so we are going to dwell upon them a bit.

#### 15.5.1 Heisenberg Groups: The Basic Motivating Example

Those who have taken quantum mechanics will be familiar with the relation between position and momentum operators for the quantum mechanics of a particle on the real line:

$$[\hat{q}, \hat{p}] = \mathrm{i}\hbar \tag{15.283}$$

One realization of these operator relations is in terms of normalizable wavefunctions $\psi(q)$ where we write:

$$\begin{aligned} (\hat{q} \cdot \psi)(q) &= q\psi(q) \\ (\hat{p} \cdot \psi)(q) &= -\mathrm{i}\hbar \frac{d}{dq}\psi(q) \end{aligned} \tag{15.284}$$

Now, let us consider the operators

$$\begin{aligned} U(\alpha) &:= \exp[\mathrm{i}\alpha\hat{p}] \\ V(\beta) &:= \exp[\mathrm{i}\beta\hat{q}] \end{aligned} \tag{15.285}$$

These are unitary when $\alpha, \beta$ are real. When $\alpha$ is real $U(\alpha)$ implements translation in position space by $\hbar\alpha$. When $\beta$ is real $V(\beta)$ implements translation in momentum space by $-\hbar\beta$.

---

[248] *Answer*: The gauge transformation $e^{\mathrm{i}\alpha(t)}$ must extend to a continuous map $\Sigma \to U(1)$. If $\Sigma$ is a smooth manifold whose only boundary is $S^1$, as we have tacitly assumed in writing equation (15.281), then such maps always restrict to small gauge transformations on the bounding $S^1$.

The group of operators $\{U(\alpha)|\alpha \in \mathbb{R}\}$ is isomorphic to $\mathbb{R}$ because $U(\alpha_1)U(\alpha_2) = U(\alpha_1+\alpha_2)$. A similar statement holds for the group of operators $V(\beta)$. But when we take products of both $U(\alpha)$ and $V(\beta)$ operators we do not get the group $\mathbb{R}\oplus\mathbb{R}$ of translations in position and momentum, separately. Rather, one can show in a number of ways that:

$$U(\alpha)V(\beta) = e^{\mathrm{i}\hbar\alpha\beta}V(\beta)U(\alpha) \tag{15.286}$$

This is an extremely important equation. We can understand it in many different ways. We will explain three ways to derive it. First, it immediately follows from the BCH formula since $[\hat{q},\hat{p}]$ is central.

A second way to derive (15.286) is to evaluate both operators on a wavefunction in the position representation. So, on the one hand:

$$\begin{aligned} ((U(\alpha)V(\beta))\cdot\psi)(q) &= (V(\beta)\cdot\psi)(q+\hbar\alpha) \\ &= e^{\mathrm{i}\beta(q+\hbar\alpha)}\psi(q+\hbar\alpha) \end{aligned} \tag{15.287}$$

On the other hand

$$\begin{aligned} ((V(\beta)U(\alpha))\cdot\psi)(q) &= e^{\mathrm{i}\beta q}(U(\alpha)\cdot\psi)(q+\hbar\alpha) \\ &= e^{\mathrm{i}\beta q}\psi(q+\hbar\alpha) \end{aligned} \tag{15.288}$$

Comparing (15.287) with (15.288) we arrive at (15.288). The reader should compare this with our discussion of quantum mechanics with a finite number of degrees of freedom, especially the derivation of (11.515).

Here is a third derivation of (15.286): Using (11.703) it follows that

$$U(\alpha)\hat{q}U(\alpha)^{-1} = e^{\mathrm{i}\alpha\mathrm{Ad}(\hat{p})}\hat{q} = \hat{q}+\hbar\alpha \tag{15.289}$$

$$V(\beta)\hat{p}V(\beta)^{-1} = e^{\mathrm{i}\beta\mathrm{Ad}(\hat{q})}\hat{p} = \hat{q}-\hbar\beta \tag{15.290}$$

Now using (11.707) we obtain (15.286).

Returning to the group generated by the operators $U(\alpha)$ and $V(\alpha)$ for $\alpha\in\mathbb{R}$, which we'll denote $\mathrm{Heis}(\mathbb{R}\times\mathbb{R})$, it fits in a central extension:

$$1 \to U(1) \to \mathrm{Heis}(\mathbb{R}\times\mathbb{R}) \to \mathbb{R}\times\mathbb{R} \to 1 \tag{15.291}$$

With one (nice) choice of cocycle we can write the group law as:

$$(z_1,(\alpha_1,\beta_1))\cdot(z_2,(\alpha_2,\beta_2)) = (z_1z_2e^{\frac{i}{2}\hbar(\alpha_1\beta_2-\alpha_2\beta_1)},(\alpha_1+\alpha_2,\beta_1+\beta_2)) \tag{15.292}$$

Notice that the cocycle is expressed in terms of the anti-symmetric form

$$\omega(v_1,v_2) := \alpha_1\beta_2-\alpha_2\beta_1 = \begin{pmatrix}\alpha_1 & \beta_1\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\alpha_2\\ \beta_2\end{pmatrix} = v_1^{tr}Jv_2 \tag{15.293}$$

where

$$v = \begin{pmatrix}\alpha\\ \beta\end{pmatrix} \tag{15.294}$$

and the matrix

$$J = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix} \tag{15.295}$$

was used at the very beginning of the course (see equation (3.36)) to define the symplectic group $Sp(2n, \kappa)$ as the group of matrices such that $A^{tr} J A = J$.

$\omega$ is called a *symplectic form* Note that

$$\omega(A\vec{v}_1, A\vec{v}_2) = \omega(\vec{v}_1, \vec{v}_2) \tag{15.296}$$

for $A \in Sp(2, \mathbb{R})$. We say that $\omega$ is invariant under symplectic transformations.

---

**Exercise**

Referring to equations (15.286) and (15.284) et. seq.

a.) Show that the choice of section

$$s(\alpha, \beta) = U(\alpha)V(\beta) \tag{15.297}$$

leads to the cocycle

$$f((\alpha_1, \beta_1), (\alpha_2, \beta_2) = e^{i\hbar(\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_1\beta_2)} \tag{15.298}$$

b.) Show that the choice of section

$$s(\alpha, \beta) = \exp[\mathrm{i}(\alpha\hat{p} + \beta\hat{q})] \tag{15.299}$$

leads to the cocycle

$$f((\alpha_1, \beta_1), (\alpha_2, \beta_2) = e^{\frac{i}{2}\hbar(\alpha_1\beta_2 - \alpha_2\beta_1)} \tag{15.300}$$

c.) Find an explicit coboundary that relates the cocycle (15.298) to (15.300). ♣NEED TO PROVIDE ANSWER HERE. ♣

---

---

**Exercise** *Generalization To* Heis($\mathbb{R}^n \oplus \mathbb{R}^n$)

Consider the group generated by exponentiating linear combinations of $\hat{q}^i$ and $\hat{p}_i$, $i = 1, \dots, n$ with the nonzero commutators being:

$$[\hat{q}^i, \hat{p}_j] = \mathrm{i}\hbar\delta^i{}_j \tag{15.301}$$

Then for

$$\vec{v} = \begin{pmatrix} \alpha_i \\ \beta^j \end{pmatrix} \tag{15.302}$$

we can choose section:

$$s(\vec{v}) = \exp\left[\mathrm{i}(\alpha^i\hat{p}_i + \beta_i\hat{q}^i)\right] \tag{15.303}$$

Show that the resulting group is

$$(z_1, \vec{v}_1) \cdot (z_2, \vec{v}_2) = (z_1 z_2 e^{\mathrm{i}\frac{\hbar}{2}\omega(\vec{v}_1, \vec{v}_2)}, \vec{v}_1 + \vec{v}_2) \tag{15.304}$$

with $\omega$ defined by equation (11.732).

---

### 15.5.2 Example: The Magnetic Translation Group For Two-Dimensional Electrons

In the presence of an electromagnetic field the group of translations acting on charged particles definitely becomes centrally extended. This shows up naturally when discussing a charged nonrelativistic particle confined to two spatial dimensions and moving in a constant magnetic field $B$. In one convenient gauge the Hamiltonian is

$$H = \frac{1}{2m}\left((p_1 + \frac{eBx_2}{2})^2 + (p_2 - \frac{eBx_1}{2})^2\right) = \frac{1}{2m}(\tilde{p}_1^2 + \tilde{p}_2^2) \tag{15.305}$$

where the gauge invariant momenta are $\tilde{p}_i := p_i - eA_i$ are

$$\begin{aligned} \tilde{p}_1 &= p_1 + \frac{eB}{2}x_2 \\ \tilde{p}_2 &= p_2 - \frac{eB}{2}x_1 \end{aligned} \tag{15.306}$$

Ordinary translations are generated by $p_1, p_2$ and do not commute with the Hamitonian: We have lost translation invariance. Nevertheless we can define the *magnetic translation operators*:

$$\pi_1 := p_1 - \frac{eBx_2}{2} \qquad \pi_2 := p_2 + \frac{eBx_1}{2} \tag{15.307}$$

Compare this carefully with the definitions of $\tilde{p}_i$. Note the relative signs! These operators satisfy $[\pi_i, \tilde{p}_j] = 0$. In particular they are translation-like operators that commute with the Hamiltonian: $[\pi_i, H] = 0$. Hence the name. While they are called "translation operators" note that they do not commute:

$$[\pi_1, \pi_2] = -i\hbar eB \tag{15.308}$$

The "magnetic translation group" is generated by the operators

$$\begin{aligned} U(a_1) &= \exp[ia_1\pi_1/\hbar] \\ V(a_2) &= \exp[ia_2\pi_2/\hbar] \end{aligned} \tag{15.309}$$

The operators $U(a_1), V(a_2)$ satisfy the relations:

$$U(a_1)V(a_2) = \exp[\mathrm{i}eBa_1a_2/\hbar]V(a_2)U(a_1) \tag{15.310}$$

If we are interested in quantized values of $a_1, a_2$ (as, for example, if the charged particle is moving in a lattice, or is confined to a torus) then we obtain the basic relations (15.501). Note that

$$\exp[\mathrm{i}eBa_1a_2/\hbar] = \exp[2\pi\mathrm{i}\Phi/\Phi_0] \tag{15.311}$$

where $\Phi = Ba_1a_2$ is the flux through an area element $a_1a_2$ and $\Phi_0 = h/e$ is known as the *magnetic flux quantum.* [249]

[249] One should be careful about a factor of two here since in superconductivity the condensing field has charge $2e$ and hence the official definition of the term "flux quantum" used, for example, by NIST is $\Phi_0 = h/2e$, half the value we use.

**Remark**: The group generated by the operators $U = U(a_1)$ and $V = V(a_2)$ is a Heisenberg group, but it is also interesting to consider the *algebra* of operators generated by $U, V$. This algebra admits a $C*$-algebra structure and is sometimes referred to as the *algebra of functions on the noncommutative torus* or the *irrational rotation algebra.* Abstractly it is the $C^*$ algebra generated by unitary operators $U, V$ satisfying $UV = e^{2\pi i\theta} VU$ for some $\theta$. The properties of the algebra are very different for $\theta$ rational and irrational. The algebra $\mathcal{A}_\theta$ figures prominantly in applications of noncommutative geometry to the QHE and in applications of noncommutative geometry to toroidal compactifications of string theory.

#### 15.5.3 The Commutator Function And The Definition Of A General Heisenberg Group

Let us now step back and think more generally about central extensions of $G$ by $A$ where $G$ is *also abelian.* From the exercise (15.158) we know that for $G$ abelian the commutator is

$$[(a_1, g_1), (a_2, g_2)] = \left(\frac{f(g_1, g_2)}{f(g_2, g_1)}, 1\right) \tag{15.312}$$

(We are writing $1/f(g_2, g_1)$ for $f(g_2, g_1)^{-1}$ and since $A$ is abelian the order doesn't matter, so we write a fraction as above.)

The function $\kappa : G \times G \to A$ defined by

$$\kappa(g_1, g_2) = \frac{f(g_1, g_2)}{f(g_2, g_1)} \tag{15.313}$$

is known as the *commutator function.*

Note that:

1. The commutator function is *gauge invariant*, in the sense that it does not change under the change of 2-cocycle $f$ by a coboundary. (Check that! This uses the property that $G$ is abelian). It is therefore a more intrinsic quantity associated with the central extension.

2. $\kappa(g, 1) = \kappa(1, g) = 1$. (This follows from exercise (15.60) above.)

3. The extension $\tilde{G}$ is abelian iff $\kappa(g_1, g_2) = 1$, that is, iff there exists a symmetric cocycle $f$. [250]

4. $\kappa$ is *skew*:
$$\kappa(g_1, g_2) = \kappa(g_2, g_1)^{-1} \tag{15.314}$$

5. $\kappa$ is *alternating*:
$$\kappa(g, g) = 1 \tag{15.315}$$

6. $\kappa$ is *bimultiplicative*:
$$\kappa(g_1g_2, g_3) = \kappa(g_1, g_3)\kappa(g_2, g_3) \tag{15.316}$$
$$\kappa(g_1, g_2g_3) = \kappa(g_1, g_2)\kappa(g_1, g_3) \tag{15.317}$$

[250] Note that in our example of $Q, D_4$ as extensions the cocycle we computed was not symmetric.

All of these properties except, perhaps, the last, are obvious. To prove the bimultiplicative properties (it suffices to prove just one) we rewrite (15.316) as

$$f(g_1g_2,g_3)f(g_3,g_2)f(g_3,g_1)=f(g_2,g_3)f(g_1,g_3)f(g_3,g_1g_2) \tag{15.318}$$

Now multiply the equation by $f(g_1,g_2)$ and use the fact that $A$ is abelian to write

$$(f(g_1,g_2)f(g_1g_2,g_3))f(g_3,g_2)f(g_3,g_1)=f(g_2,g_3)f(g_1,g_3)(f(g_1,g_2)f(g_3,g_1g_2)) \tag{15.319}$$

We apply the cocycle identity on both the LHS and the RHS (and also use the fact that $G$ is abelian) to get

$$f(g_2,g_3)f(g_1,g_2g_3)f(g_3,g_2)f(g_3,g_1)=f(g_2,g_3)f(g_1,g_3)f(g_3,g_1)f(g_3g_1,g_2) \tag{15.320}$$

Now canceling some factors and using that $A$ is abelian we have

$$f(g_1,g_2g_3)f(g_3,g_2)=f(g_1,g_3)f(g_3g_1,g_2) \tag{15.321}$$

Now use the fact that $G$ is abelian to write this as

$$f(g_1,g_3g_2)f(g_3,g_2)=f(g_1,g_3)f(g_1g_3,g_2) \tag{15.322}$$

which is the cocycle identity. This proves the bimultiplicative property (15.316). ♠

We now define the *Heisenberg extensions*. The function $\kappa$ is said to be *nondegenerate* if for all $g_1 \neq 1$ there is a $g_2$ with $\kappa(g_1,g_2)\neq 1$. When this is the case the center of $\tilde{G}$ is precisely $A$:

$$Z(\widetilde{G}) \cong A\ . \tag{15.323}$$

This follows immediately from equation (15.312). If $\kappa$ is degenerate the center will be larger. In the extreme case that $\kappa(g_1,g_2)=1$ for all $g_1,g_2$ we get the direct product $\widetilde{G}=A\times G$ and

$$Z(\widetilde{G})=\widetilde{G}\ . \tag{15.324}$$

In general, we will have an intermediate situation and $A$ will be a proper subgroup of $Z(\tilde{G})$.

One definition which is used in the literature is

**Definition**: A *Heisenberg extension* is a central extension of an *abelian* group $G$ by an *abelian* group $A$ where the commutator function $\kappa$ is nondegenerate.

---

**Exercise** *Alternating implies skew*

Show that a map $\kappa : G\times G\to A$ which satisfies the bimultiplicative identity (15.316) and the alternating identity (15.315) is also skew, that is, satisfies (15.314).

---

---

**Exercise** *Commutator function for* Heis($\mathbb{R} \oplus \mathbb{R}$)

a.) Show that both of the cocycles (15.298) and (15.300) defining groups isomorphic to Heis($\mathbb{R} \oplus \mathbb{R}$) have the same commutator function

$$\kappa((\alpha_1, \beta_1), (\alpha_2, \beta_2) = e^{i\hbar(\alpha_1\beta_2 - \alpha_2\beta_1)} = e^{i\hbar\omega(v_1,v_2)} \tag{15.325}$$

b.) Similarly show that for Heis($\mathbb{R}^n \oplus \mathbb{R}^n$) we have

$$\kappa(v_1, v_2) = \exp\left(i\hbar\omega(v_1, v_2)\right) \tag{15.326}$$

where

$$\omega(v_1, v_2) = v_1^{tr} J v_2 \tag{15.327}$$

---

### 15.5.4 Classification Of $U(1)$ Central Extensions Using The Commutator Function

For a large class of abelian groups $G$, there is a nice theorem regarding arbitrary central extensions by $U(1)$. We consider

1. Finitely generated Abelian groups. As we will prove in sections 16.2 and 16.3 below, these can be (noncanonically) written as products of cyclic groups $\mathbb{Z}_n$, for various $n$, and a lattice $\mathbb{Z}^d$ for some $d$ (possibly $d = 0$).
2. Vector spaces. [251]
3. Tori. These are isomorphic to $V/\mathbb{Z}^d$ where $V$ is a $d$-dimensional real vector space.
4. Direct products of the above three.

**Remark**: This class of groups can be characterized as the set of Abelian groups $A$ which are topological groups so there is an exact sequence:

$$0 \to \pi_1(A) \to \mathrm{Lie}(A) \to A \to \pi_0(A) \to 0 \tag{15.328}$$

where Lie($A$) is a vector space that projects to $A$ by an exponential map.

For this class of groups we have the following theorem:

**Theorem** Let $G$ be a topological Abelian group of the above class. The isomorphism classes of central extensions of $G$ by $U(1)$ are in one-one correspondence with continuous bimultiplicative maps

$$\kappa : G \times G \to U(1) \tag{15.329}$$

which are alternating (and hence skew).

[251] Topological, separable.

I do not know who originally proved this theorem, but one proof can be found in [252]

**Remarks**

1. In other words, given the commutator function $\kappa$ one can always find a corresponding cocycle $f$. This theorem is useful because $\kappa$ is invariant under change of $f$ by a coboundary, and moreover the bimultiplicative property is simpler to check than the cocycle identity. (In fact, one can show that it is always possible to find a cocycle $f$ which is bimultiplicative. This property automatically ensures the cocycle relation.)

2. It is important to realize that $\kappa$ only characterizes $\tilde{G}$ up to *noncanonical* isomorphism: to give a definite group one must choose a definite cocycle.

♣Explain this comment more. ♣

3. In this theorem we can replace $U(1)$ by any subgroup of $U(1)$, such as $\mathbb{Z}_n$ realized as the group of $n^{th}$ roots of unity.

#### 15.5.5 Pontryagin Duality And The Stone-von Neumann-Mackey Theorem

The theorem of section 15.5.4 is nicely illustrated by the Heisenberg of extension of the Abelian group $G = S \times \widehat{S}$ by $A = U(1)$. Here $S$ is any locally compact Abelian group and $\widehat{S}$ is the Pontryagin dual.

Now, there is a very natural alternating, skew, bilinear map on the product $S \times \widehat{S}$ defined by

$$\kappa\left((s_1,\chi_1),(s_2,\chi_2)\right) := \frac{\chi_2(s_1)}{\chi_1(s_2)} \tag{15.330}$$

and according to the above theorem this defines a general Heisenberg extension

$$1 \to U(1) \to \mathrm{Heis}(S\times\widehat{S}) \to S\times\widehat{S} \to 1 \tag{15.331}$$

at least up to isomorphism.

**Remarks**

1. Note that one natural cocycle giving the commutator function (15.330) is

$$f\left((s_1,\chi_1),(s_2,\chi_2)\right) := \frac{1}{\chi_1(s_2)} \tag{15.332}$$

2. There is a very natural representation of (15.331). We need a Haar measure on $S$ so that we can define $V = L^2(S)$, a Hilbert space of complex-valued functions on $S$.

[252] D. Freed, G. Moore, G. Segal, "The uncertainty of fluxes," Commun.Math.Phys. 271 (2007) 247-274, arXiv:hep-th/0605198, Proposition A.1.

For our key examples we have

$$
\begin{aligned}
\langle \psi_1, \psi_2 \rangle &= \int_{\mathbb{R}} \psi_1^*(x)\psi_2(x)dx \qquad S = \mathbb{R} \\
&= \sum_{n\in\mathbb{Z}} \psi_1^*(n)\psi_2(n) \qquad S = \mathbb{Z} \\
&= \frac{1}{2\pi}\int_0^{2\pi} \psi_1^*(e^{\mathrm{i}\theta})\psi_2(e^{\mathrm{i}\theta})d\theta \qquad S = U(1) \\
&= \frac{1}{n}\sum_{k=0}^{n-1} \psi_1^*(\bar{k})\psi_2(\bar{k}) \qquad S = \mathbb{Z}/n\mathbb{Z}
\end{aligned}
\tag{15.333}
$$

We represent $s_0 \in S$ as a translation operator:

$$
(T_{s_0} \cdot \Psi)(s) := \Psi(s + s_0) \tag{15.334}
$$

and we represent $\chi_0 \in \widehat{S}$ as a multiplication operator

$$
(M_{\chi_0} \cdot \Psi)(s) := \chi_0(s)\Psi(s) \tag{15.335}
$$

Then one checks that this does <u>not</u> define a representation of the direct product $S \times \widehat{S}$ but rather we have the operator equation:

$$
T_{s_0} M_{\chi_0} = \chi_0(s_0) M_{\chi_0} T_{s_0} \tag{15.336}
$$

If $\mathcal{O}$ is the group of operators generated by $T_s$, $M_\chi$ and $z \in U(1)$ acting on $V$ then the map

$$
(z; (s, \chi)) \to z T_s M_\chi \tag{15.337}
$$

is a homomorphism, and in fact an isomorphism of $\mathrm{Heis}(S \times \widehat{S})$ with $\mathcal{O}$ where we use the cocycle (15.332).

3. This construction is extremely important in quantum mechanics and also in the description of free quantum field theories. In these cases we take $S$ to be a vector space. For example, for the quantum mechanics of a particle in $\mathbb{R}^n$ we have $V = \mathbb{R}^n$ as an Abelian group. For the quantum field theory of a free real scalar field in $d+1$ dimensions we would take $V$ to be a suitable space of real-valued functions on a $d$-dimensional spatial slice. [253] Then we introduce the dual vector space $V^\vee \cong \widehat{S}$, and the canonical pairing $V \times V^\vee \to \mathbb{R}$ gives the character:

$$
\chi_k(v) = e^{\mathrm{i}k\cdot v} \tag{15.338}
$$

Then, $T_s$ in our general discussion is the operator $U(s)$ of the basic motivating example, and $M_{\chi_k}$ is the operator $V(k)$ of the basic motivating example and the general identity (15.336) becomes our starting point:

$$
U(s)V(k) = e^{\mathrm{i}k\cdot s} V(k) U(s) \tag{15.339}
$$

[253] Technical point: This will not be locally compact. So, if one wishes to be rigorous, further considerations are required.

4. *Stone-von Neumann-Mackey Theorem.* Up to isomorphism (equivalence) there is a unique irreducible unitary representation of $\text{Heis}(S \times \widehat{S})$ such that $A = U(1)$ acts by scalar multiplication. That is, if $\xi \in U(1)$ then we require $\rho(\xi) = \xi \mathbf{1}_V$. In addition we need some further technical hypotheses. [254] This is called the Stone-von Neumann theorem or sometimes the Stone-von Neumann-Mackey theorem. For a relatively short proof see. [255] The main idea is to consider the maximal Abelian subgroups of $\text{Heis}(S \times \widehat{S})$. One such subgroup is isomorphic to $U(1) \times S$, another is $U(1) \times \widehat{S}$. Let us consider $U(1) \times \widehat{S}$. Over the subgroup $\{1\} \times \widehat{S}$ we can split the sequence and consider the elements:

$$M_\chi := \rho(1, (0, \chi)) \tag{15.340}$$

where $1 \in U(1)$, $0 \in S$, and $\chi \in \widehat{S}$. These operators commute for different choices of $\chi$ and are simultaneously diagonalizable so we have a complete basis $\psi_\alpha$ of simultaneous eigenvectors for the representation with eigenvalues: [256]

$$M_\chi \psi_\alpha = \lambda_\alpha(\chi) \psi_\alpha \tag{15.341}$$

Clearly $\chi \mapsto \lambda_\alpha(\chi)$ must be a character on $\widehat{S}$, so the eigenvalues can be identified with characters on $\widehat{S}$ and therefore, by Pontryagin duality, can be identified with elements $s_\alpha \in S$. So there is some set $\{s_\alpha\}$ whose elements are drawn from $S$ with corresponding eigenbasis $\psi_\alpha$ with

$$M_\chi \psi_\alpha = \chi(s_\alpha) \psi_\alpha \tag{15.342}$$

Now, for any $s \in S$ define the operator

$$T_s := \rho(1, (s, 1)) \tag{15.343}$$

Choose some particular $\alpha_0$ and consider the vectors $T_s \psi_{\alpha_0}$ for $s \in S$. Using the Heisenberg relations and the fact that the central $U(1)$ group just acts by scalars we know that:

$$M_\chi \left(T_s \psi_{\alpha_0}\right) = \chi(s - s_{\alpha_0}) \left(T_s \psi_{\alpha_0}\right) \tag{15.344}$$

Therefore the span of $\{T_s \psi_{\alpha_0}\}_{s \in S}$, which is the span of $\{T_{s+s_{\alpha_0}} \psi_{\alpha_0}\}_{s \in S}$ is a copy of the representation constructed above. So if $(V, \rho)$ is irreducible, it must be equivalent to our representation constructed above. This demonstrates uniqueness of the irreducible representation.

5. *Simple Proof Of Irreducibility for $S = \mathbb{R}^n$.* For $v = (\alpha, \beta) \in \mathbb{R}^{2n}$ we introduced the section

$$s(v) := \exp[\mathrm{i}(\alpha \hat{q} + \beta \hat{p})] \tag{15.345}$$

Now, consider the standard representation of $\text{Heis}(\mathbb{R}^n \times \widehat{\mathbb{R}^n})$ on $\mathcal{H} = L^2(\mathbb{R}^n)$. For

♣This remark assumes at least a little bit of knowledge from the linear algebra chapter 2 and the idea of an irreducible representation, from chapter 4. ♣

♣The $\alpha, \beta$ here are reversed from the convention in previous section. ♣

[254] We need the representation to be continuous in the norm topology and we need $S$ to have a translation-invariant measure so that $L^2(A)$ makes sense. Then we replace $V$ above by the Hilbert space $L^2(S)$.

[255] A. Prasad, "An easy proof of the Stone-von Neumann-Mackey Theorem," arXiv:0912.0574.

[256] It is exactly at this point that we are using unitarity, and a careful discussion requires more functional analysis.

any two vectors $\psi_1, \psi_2 \in \mathcal{H}$ define the *Wigner function*:

$$W(\psi_1, \psi_2)(v) := \langle s(v)\psi_1, \psi_2 \rangle \tag{15.346}$$

This is a function on the phase space $v \in \mathbb{R}^{2n}$. Now we compute:

$$(s(v)\psi_1)(q) = e^{\mathrm{i}\frac{\alpha\beta}{2}} e^{\mathrm{i}\alpha q} \psi_1(q+\beta) \tag{15.347}$$

Shifting the integration variable by $\beta/2$ in the inner product we have

$$W(\psi_1, \psi_2)(v) = \int_{\mathbb{R}^n} e^{-\mathrm{i}\alpha q} \psi_1^*(q+\beta/2)\psi_2(q-\beta/2) dq \tag{15.348}$$

Now, an elementary computation shows that, on $L^2(\mathbb{R}^{2n})$ with the standard measure we have

$$\| W(\psi_1, \psi_2) \|^2 = \| \psi_1 \|^2 \| \psi_2 \|^2 \tag{15.349}$$

Here are some details

♣Need to clean up some $2\pi$'s here... ♣

$$\begin{aligned}
\| W(\psi_1, \psi_2) \|^2 &= \int_{\mathbb{R}^{2n}} |W(\psi_1, \psi_2)|^2 d^{2n} v \\
&= \int d\alpha d\beta dq_1 dq_2 e^{\mathrm{i}\alpha(q_1 - q_2)} \psi_1(q_1+\beta/2)\psi_2^*(q_1-\beta/2)\psi_1^*(q_2+\beta/2)\psi_2(q_2-\beta/2) \\
&= \int dq d\beta \psi_1^*(q+\beta/2)\psi_1(q+\beta/2)\psi_2^*(q-\beta/2)\psi_2(q-\beta/2) \\
&= \| \psi_1 \|^2 \| \psi_2 \|^2
\end{aligned} \tag{15.350}$$

Now with the key result (15.349) we can show that $\mathcal{H}$ is irreducible. [257] Suppose that $\mathcal{H}_0 \subset \mathcal{H}$ is preserved preserved by the Heisenberg group. Suppose there is a vector $\psi_\perp \notin \mathcal{H}_0$. WLOG we can take $\psi_\perp$ to be perpendicular to $\mathcal{H}_0$ (hence the notation). But then,

$$W(\psi, \psi_\perp)(v) = \langle s(v)\psi, \psi_\perp \rangle = 0 \tag{15.351}$$

for all $v \in \mathbb{R}^{2n}$ and all $\psi \in \mathcal{H}_0$, because $s(v)\psi \in \mathcal{H}_0$. But then by (15.349) we know that $\| \psi \|^2 \| \psi_\perp \|^2 = 0$. Therefore either $\psi = 0$ or $\psi_\perp = 0$. If $\mathcal{H}_0$ is not the zero vector space then we can always choose $\psi \neq 0$ and hence $\psi_\perp = 0$. But if $\mathcal{H}_0$ were proper then there would be a nonzero choice for $\psi_\perp$. Therefore, there is no nonzero and proper subspace $\mathcal{H}_0 \subset \mathcal{H}$ preserved by the group action of Heis. Therefore $\mathcal{H}$ is irreducible.

6. *Stone von-Neumann And Fourier* Now let us combine the Stone-von-Neumann theorem with Pontryagin duality. Because $\widehat{\widehat{S}} \cong S$ we can write

$$\mathrm{Heis}(S \times \widehat{S}) \cong \mathrm{Heis}(\widehat{S} \times \widehat{\widehat{S}}) \tag{15.352}$$

[257] See chapter 4 for a thorough discussion of reducible vs. irreducible representations. Briefly - if a representation $\mathcal{H}$ has a nonzero and proper subspace $\mathcal{H}_0$ preserved by the group action then it is said to be *reducible*. A representation which is not reducible is said to be *irreducible*.

so, we could equally well give a unitary representation of the group by taking $\widehat{V} := Fun(\widehat{S} \to \mathbb{C})$ with inner product

$$\langle \hat{\psi}_1, \hat{\psi}_2 \rangle := \int_{\widehat{S}} d\chi \hat{\psi}_1^*(\chi) \hat{\psi}_2(\chi) \tag{15.353}$$

Now we represent translation and multiplication operators by

$$(\hat{T}_{\chi_0} \hat{\psi})(\chi) := \hat{\psi}(\chi_0 \chi) \tag{15.354}$$

$$(\hat{M}_{s_0} \hat{\psi})(\chi) := \chi(s_0) \hat{\psi}(\chi) \tag{15.355}$$

and check the commutator function. So, by SvN there must be a unitary isomorphism

$$\mathcal{S} : V \to \widehat{V} \tag{15.356}$$

mapping

$$\begin{aligned} \mathcal{S} T_{s_0} \mathcal{S}^{-1} &= \hat{M}_{s_0} \\ \mathcal{S} M_{\chi_0} \mathcal{S}^{-1} &= \hat{T}_{\chi_0} \end{aligned} \tag{15.357}$$

Indeed there is: It is the Fourier transform:

$$\psi \mapsto \hat{\psi}(\chi) := \int_S \chi(s)^* \psi(s) ds \tag{15.358}$$

Moreover, $\mathcal{S}$ is an isometry by the Plancherel/Parseval theorem noted above.

7. Of course, there are many different Lagrangian subgroups of $G$. For example, if $G$ is a symplectic vector space there will be many different Lagrangian subspaces, all related by the linear action of the symplectic group. Each choice of Lagrangian subspace gives a representation of the Heisenberg group, but they all must be isomorphic, by the Stone-von Neumann-Mackey theorem. This leads to interesting isomorphisms between seemingly different representations and interesting (projective) actions of symplectic groups. These kinds of considerations are central to some simple examples of "duality symmetries" in quantum field theory. We will next investigate how the groups of symplectic automorphisms lift to automorphism groups of Heisenberg groups.

---

**Exercise** *Bimultiplicative cocycle*
a.) Check that (15.332) satisfies the cocycle relation.
b.) Show that, in fact, (15.332) is bimultiplicative.

---

---

**Exercise** *Irreducibility Of The Stone-von Neumann Representation Of* Heis$(S \times \widehat{S})$

Generalize the proof of irreducibility for $S = \mathbb{R}^n$ to other locally compact Abelian groups. [258]

---

### 15.5.6 Some More Examples Of Heisenberg Extensions

For this section the reader might wish to consult section [**** 2.1 *** ?] of chapter two for the definition of a ring. The reader won't lose much by taking $R = \mathbb{Z}$ or $R = \mathbb{Z}/N\mathbb{Z}$ with abelian group structure + and extra multiplication structure $\bar{n}_1\bar{n}_2 = \overline{n_1 n_2}$.

**Example 1**: Suppose $R$ is a commutative ring with identity. Then we can consider the group of $3 \times 3$ matrices over $R$ of the form

$$M(a,b,c) := \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \tag{15.359}$$

The multiplication law is easily worked out to be

$$M(a,b,c)M(a',b',c') = M(a+a', b+b', c+c'+ab') \tag{15.360}$$

Therefore, if we define Heis$(R \times R)$ to be the group of matrices $M(a,b,c)$ and take

$$\pi : M(a,b,c) \to (a,b) \tag{15.361}$$

and

$$\iota : c \mapsto M(0,0,c) \tag{15.362}$$

we have an extension

$$0 \to R \to \mathrm{Heis}(R \times R) \to R \oplus R \to 0 \tag{15.363}$$

with cocycle $f((a,b),(a',b')) = ab'$. Note that we are writing our Abelian group $R$ additively so the cocycle identity becomes

$$f(v_1, v_2) + f(v_1 + v_2, v_3) = f(v_1, v_2 + v_3) + f(v_2, v_3) \tag{15.364}$$

where $v = (a,b) \in \mathbb{R} \oplus \mathbb{R}$. In this additive notation the commutator function is

$$\kappa((a,b),(a',b')) = ab' - a'b \tag{15.365}$$

In the literature one will sometimes find the above class of groups defined as the "Heisenberg groups." It is a special case of what we have defined as general Heisenberg groups.

---

[258] *Answer* For the general case define the Wigner function as the function $W(\psi_1,\psi_2) : S \times \widehat{S} \to \mathbb{C}$ by $W(\psi_1,\psi_2)(s,\chi) := \langle T_s M_\chi \psi_1, \psi_2 \rangle$. Show that (15.349) continues to hold. You will need to use the orthogonality relation for characters.

As a special case of the above construction let us take $R = \mathbb{Z}/n\mathbb{Z}$. We will now show that we recover the group $\mathrm{Heis}(\mathbb{Z}_n \times \mathbb{Z}_n)$ discussed in section 11.17.2 in the context of a particle on a discrete approximation to a circle.

First consider $\mathbb{Z}_n$ written additively. So if $a \in \mathbb{Z}$, then $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$ is just $\bar{a} = a + n\mathbb{Z}$ is the coset. Then we define

$$U = \begin{pmatrix} \bar{1} & \bar{1} & 0 \\ 0 & \bar{1} & 0 \\ 0 & 0 & \bar{1} \end{pmatrix} \qquad V = \begin{pmatrix} \bar{1} & 0 & 0 \\ 0 & \bar{1} & \bar{1} \\ 0 & 0 & \bar{1} \end{pmatrix} \qquad q = \begin{pmatrix} \bar{1} & 0 & \bar{1} \\ 0 & \bar{1} & 0 \\ 0 & 0 & \bar{1} \end{pmatrix} \tag{15.366}$$

We easily check that for $a \in \mathbb{Z}$,

$$U^a = \begin{pmatrix} \bar{1} & \bar{a} & 0 \\ 0 & \bar{1} & 0 \\ 0 & 0 & \bar{1} \end{pmatrix} \qquad V^a = \begin{pmatrix} \bar{1} & 0 & 0 \\ 0 & \bar{1} & \bar{a} \\ 0 & 0 & \bar{1} \end{pmatrix} \qquad q^a = \begin{pmatrix} \bar{1} & 0 & \bar{a} \\ 0 & \bar{1} & 0 \\ 0 & 0 & \bar{1} \end{pmatrix} \tag{15.367}$$

so

$$U^n = V^n = q^n = 1 \tag{15.368}$$

Moreover:

$$UV = qVU \qquad qU = Uq \qquad qV = Vq \tag{15.369}$$

Thus we obtain the presentation:

$$\mathrm{Heis}(\mathbb{Z}_n \times \mathbb{Z}_n) = \langle U, V, q | U^n = V^n = q^n = 1, \quad UV = qVU, \quad Uq = qU, \quad Vq = qV \rangle \tag{15.370}$$

we saw before.

A simple and useful generalization of the previous construction is to take any bilinear map $c : R \times R \to \mathcal{Z}$ where $\mathcal{Z}$ is Abelian. Thus $c(a_1 + a_2, b) = c(a_1, b) + c(a_2, b)$ and $c(a, b_1 + b_2) = c(a, b_1) + c(a, b_2)$. Then we can define a central extension

$$0 \to \mathcal{Z} \to \tilde{G} \to R \oplus R \to 0 \tag{15.371}$$

by the law

$$(z_1, (a, b)) \cdot (z_2, (a', b')) = (z_1 + z_2 + c(a, b'), (a + a', b + b')) \tag{15.372}$$

The corresponding group cocycle is $f((a, b), (a', b')) = c(a, b')$. The cocycle relation is satisfied simply by virtue of $c$ being bilinear. It will be a Heisenberg extension if $\kappa : (R \times R) \times (R \times R) \to \mathcal{Z}$ given by $\kappa((a, b), (a', b')) = c(a, b') - c(a', b)$ is nondegenerate. In particular, if we take $\mathcal{Z} = R$ and $c(a, b') = ab'$ using the ring multiplication then we recover (15.359).

**Example 2**: *Clifford algebra representations and Extra-special groups.* Suppose we have a set of matrices $\gamma_i$, $1 \leq i \leq n$ such that

$$\{\gamma_i, \gamma_j\} = 2\delta_{ij} \tag{15.373}$$

Such a set of matrices can indeed be constructed, by taking suitable tensor products of Pauli matrices. They are called "gamma matrices." They form a matrix representation of what is called a *Clifford algebra* and we will study them in more detail and more abstractly in chapter ****. For the moment the reader should be content with the explicit representation: [259]

$$\begin{aligned} \gamma_1 &= \sigma^1 \\ \gamma_2 &= \sigma^2 \end{aligned} \tag{15.374}$$

for $n = 2$,

$$\begin{aligned} \gamma_1 &= \sigma^1 \\ \gamma_2 &= \sigma^2 \\ \gamma_3 &= \sigma^3 \end{aligned} \tag{15.375}$$

for $n = 3$,

$$\begin{aligned} \gamma_1 &= \sigma^1 \otimes \sigma^1 \\ \gamma_2 &= \sigma^1 \otimes \sigma^2 \\ \gamma_3 &= \sigma^1 \otimes \sigma^3 \\ \gamma_4 &= \sigma^2 \otimes 1 \end{aligned} \tag{15.376}$$

for $n = 4$,

$$\begin{aligned} \gamma_1 &= \sigma^1 \otimes \sigma^1 \\ \gamma_2 &= \sigma^1 \otimes \sigma^2 \\ \gamma_3 &= \sigma^1 \otimes \sigma^3 \\ \gamma_4 &= \sigma^2 \otimes 1 \\ \gamma_5 &= \sigma^3 \otimes 1 \end{aligned} \tag{15.377}$$

for $n = 5$,

$$\begin{aligned} \gamma_1 &= \sigma^1 \otimes \sigma^1 \otimes \sigma^1 \\ \gamma_2 &= \sigma^1 \otimes \sigma^1 \otimes \sigma^2 \\ \gamma_3 &= \sigma^1 \otimes \sigma^1 \otimes \sigma^3 \\ \gamma_4 &= \sigma^1 \otimes \sigma^2 \otimes 1 \\ \gamma_5 &= \sigma^1 \otimes \sigma^3 \otimes 1 \\ \gamma_6 &= \sigma^2 \otimes 1 \otimes 1 \end{aligned} \tag{15.378}$$

for $n = 6$, and so on. So for the Clifford algebra with $n$ generators we have constructed a representation by $2^{[n/2]} \times 2^{[n/2]}$ matrices.

Of course, the above choice of matrices is far from a unique choice of matrices satisfying the Clifford relations (15.373). If the $\gamma_i \in \mathrm{Mat}_d(\mathbb{C})$ then for any $S \in GL(d, \mathbb{C})$ we can

[259]See Chapter 2, section 5.3 for a detailed discussion of tensor product $\otimes$.

change $\gamma_i \to S\gamma_i S^{-1}$. These give *equivalent* representations of the Clifford algebra. We could also modify $\gamma_i \to \epsilon_i \gamma_i$ where $\epsilon_i \in \{\pm 1\}$ and still get a representation, although it might not be an equivalent one.

For example, note that for $n = 3$

$$\gamma_1\gamma_2\gamma_3 = \mathrm{i}\mathbf{1}_{2\times 2} \tag{15.379}$$

and for $n = 5$

$$\gamma_1 \cdots \gamma_5 = -\mathbf{1}_{4\times 4} \tag{15.380}$$

We cannot change the sign on the RHS by conjugating with $S$. So in this case we conclude that there are at least two inequivalent representations of the Clifford algebra.

The general story, proved in detail in Chapters 11-12 is that

1. If $n$ is even there is a unique irreducible representation of dimension $d = 2^{n/2}$.
2. If $n$ is odd there are precisely two irreducible representation of dimension $d = 2^{(n-1)/2}$ and they are distinguished by the sign of the "Clifford volume element" $\omega = \gamma_1 \cdots \gamma_n$.

In Chapter 11 these statements, and more are generalized to real Clifford algebras for a quadratic form of any signature,

For $w \in \mathbb{Z}_2^n$ (where we will think of $\mathbb{Z}_2 = \mathbb{Z}/2\mathbb{Z}$ as a ring in this example) we define

$$\gamma(w) := \gamma_1^{w_1} \cdots \gamma_n^{w_n} \tag{15.381}$$

Then

$$\gamma(w)\gamma(w') = \epsilon(w, w')\gamma(w + w') = \kappa(w, w')\gamma(w')\gamma(w) \tag{15.382}$$

where

$$\epsilon(w, w') = (-1)^{\sum_{i>j} w_i w'_j} \tag{15.383}$$

defines a cocycle with commutator function

$$\kappa(w, w') = (-1)^{\sum_{i\neq j} w_i w'_j} \tag{15.384}$$

When is the commutator function $\kappa$ nondegenerate? We need to consider two cases:

<u>Case 1a</u> $\sum_i w_i = 1 \bmod 2$ and there is an $i_0$ so that $w_{i_0} = 0$. Then $\gamma(w)$ anticommutes with $\gamma_{i_0}$.

<u>Case 1b</u> $\sum_i w_i = 1 \bmod 2$ and $w_i = 1$ for all $i$. In this case $n$ must be odd. Then in fact $\gamma_1 \cdots \gamma_n$ commutes with all the $\gamma_i$ and the commutator function is <u>degenerate</u>.

<u>Case 2</u>. $\sum_i w_i = 0 \bmod 2$ and some $w_{i_0} \neq 0$. Then $\gamma_{i_0}$ anticommutes with $\gamma(w)$.

We conclude that for the even degree Clifford algebras $\kappa$ is nondegenerate and the group generated by taking products of the matrices $\pm\gamma(w)$ defined by an irreducible representation in fact defines a Heisenberg extension:

$$1 \to \mathbb{Z}_2 \to \mathcal{E}_{2m} \to \mathbb{Z}_2^{2m} \to 1 \tag{15.385}$$

In finite group theory this group is an example of what is known as an "extra-special group" and is denoted $\mathcal{E}_{2m} = 2_+^{1+2m}$. [260]

In the case when $n$ is odd then in fact the Clifford volume form $\gamma_1 \cdots \gamma_n$ commutes with all the $\gamma_i$ and the cocycle is degenerate.

**Example 3**: *Heisenberg Construction Of Nontrivial $U(1)$ Bundle Over Symplectic Tori.* Consider a torus $T := \mathbb{R}^n/\Gamma$ where $\Gamma$ is an integral lattice. Suppose we have an integral-valued symplectic form on $\Gamma$. This is a blinear, anti-symmetric, nondegenerate map:

$$\Omega : \Gamma \times \Gamma \to \mathbb{Z} \tag{15.386}$$

It is shown in the Linear Algebra User's Manual that we can choose an ordered basis $\{\gamma_1, \dots, \gamma_n\}$ for $\Gamma$ so that the matrix $\Omega(\gamma_i, \gamma_j)$ is of the form

$$\begin{pmatrix} 0 & d_1 \\ -d_1 & 0 \end{pmatrix} \oplus \begin{pmatrix} 0 & d_2 \\ -d_2 & 0 \end{pmatrix} \cdots \tag{15.387}$$

where the $d_i$ are integers. We will assume that $\Omega$ is nondegenerate so therefore $n = 2m$ is even and all the integer $d_i$ are nonzero.

If $\Gamma$ is full rank we can extend $\Omega$ to a bilinear antisymmetric form on [261]

$$V = \Gamma \otimes_{\mathbb{Z}} \mathbb{R} \cong \mathbb{R}^{2m}$$

to get an antisymmetric bilinear form:

$$\Omega : V \times V \to \mathbb{R} \tag{15.388}$$

If all the $d_i$ are nonzero then this is a symplectic form. Using an invertible matrix $S$ we can bring $S^{tr}\Omega S$ to the standard form $J$. We can define a commutator function on $\mathbb{R}^n$:

$$\kappa(v_1, v_2) = e^{2\pi \mathrm{i}\Omega(v_1, v_2)} \tag{15.389}$$

and this defines the isomorphism class of a Heisenberg extension of $V \cong \mathbb{R}^{2m}$. To be concrete our extension is

$$1 \to U(1) \to \mathrm{Heis}(V, \Omega) \to V \to 0 \tag{15.390}$$

[260] For any prime $p$ an *extra-special group* is a group $G$ that fits in a central extension $1 \to \mathbb{Z}_p \to G \to \mathbb{Z}_p^n \to 1$ where the center is minimal, that is, is isomorphic to $\mathbb{Z}_p$. For example, for $p = 2$ we have seen that both $D_4$ and $Q$ are extra-special groups. Up to isomorphism there are two such groups, sometimes denoted $p_\pm^{1+n}$.

[261] The symbol $\otimes_{\mathbb{Z}}$ is explained in the Linear Algebra User's Manual. It means we are taking a tensor product of $\mathbb{Z}$-modules, i.e. Abelian groups.

where we make the explicit choice of cocycle $f(v_1,v_2) = e^{\mathrm{i}\pi\Omega(v_1,v_2)}$. So, $\mathrm{Heis}(V,\Omega)$ is the group of pairs $(z,v)$ with $z \in U(1)$ and $v \in V$ with multiplication:

$$(z_1,v_1)\cdot(z_2,v_2) := (z_1 z_2 e^{\mathrm{i}\pi\Omega(v_1,v_2)}, v_1+v_2) \tag{15.391}$$

We stress that this gives a Heisenberg group and in particular the sequence does not split.

Now consider the pullback of the sequence under the inclusion $\iota : \Gamma \to V$. We claim that the pulled-back sequence splits: Let us try to choose a section

$$s(\gamma) = (\epsilon_\gamma, \gamma) \tag{15.392}$$

then to split the sequence we will need

$$\begin{aligned}(\epsilon_{\gamma_1+\gamma_2}, \gamma_1+\gamma_2) &= s(\gamma_1+\gamma_2) \\ &= s(\gamma_1)s(\gamma_2) \\ &= (\epsilon_{\gamma_1}\epsilon_{\gamma_2}e^{\mathrm{i}\pi\Omega(\gamma_1,\gamma_2)}, \gamma_1+\gamma_2)\end{aligned} \tag{15.393}$$

In other words, to split the sequence over $\Gamma$ we need to find a function $\epsilon : \Gamma \to U(1)$ so that

$$\epsilon_{\gamma_1}\epsilon_{\gamma_2} = e^{-\mathrm{i}\pi\Omega(\gamma_1,\gamma_2)}\epsilon_{\gamma_1+\gamma_2} \tag{15.394}$$

It is indeed possible to find such functions. See the exercise.

Since the sequence splits over $\Gamma$ we consider the Abelian subgroup $s(\Gamma) \subset \tilde{V}$. Then define the quotient space:

$$P(T,\Omega,\epsilon) := \mathrm{Heis}(V,\Omega)/s(\Gamma) \tag{15.395}$$

Explicitly $P(T,\Omega,\epsilon)$ is the quotient $(U(1)\times V)/\Gamma$ with the equivalence relation

$$(z,v) \sim (z,v)\cdot(\epsilon_\gamma,\gamma) = (z\epsilon_\gamma e^{\mathrm{i}\pi\Omega(v,\gamma)}, v+\gamma) \tag{15.396}$$

for all $\gamma \in \Gamma$.

Note that there is a continuous map

$$\pi : P(T,\Omega,\epsilon) \to T \tag{15.397}$$

whose fiber is $U(1)$. This space, together with its projection map is an example of a principal $U(1)$ bundle over the torus $T$: Each fiber is a principal homogeneous space for the group $U(1)$, under the natural action of $U(1)$ on $P(T,\Omega,\epsilon)$. (Since the $U(1)$ is central we can consider it either as a left- or right- action.) Our construction of the bundle depended on a choice of splitting $\epsilon_\gamma$, but a change of splitting defines isomorphic bundles.

**Remarks**

1. The above construction comes up, either explicitly or implicitly in discussions of the quantum Hall efect, Chern-Simons theory, and the quantization of $p$-form gauge theories.

2. Note that while $s(\Gamma)$ is a subgroup of $\mathrm{Heis}(V,\Omega)$ it is <u>not</u> a normal subgroup, so that, while, $P(T,\Omega,\epsilon)$ is a bundle, it is not a group.

3. If we view $\mathrm{Heis}(V,\Omega)$ as a principal $U(1)$ bundle over $V$ then we can construct a very natural connection on this bundle. Parallel transport of the point $(z, v_0) \in P(T,\Omega,\epsilon)$ over a straightline path $\wp_{v_0,w} := \{v_0 + tw | 0 \leq t \leq 1\}$ in $V$ is defined by left-multiplication by $(1, w)$:

$$U(\wp_{v_0,w}) : (z, v_0) \to (1, w) \cdot (z, v_0) = (z e^{\mathrm{i}\pi\Omega(w,v_0)}, w + v_0) \tag{15.398}$$

Now, if one considers parallel transport around a small square loop starting at $v_0$ by composing paths

$$\wp_{v,w_1,w_2} := \wp_{v_0,w_1} \star \wp_{v_0+w_1,w_2} * \wp_{v_0+w_1+w_2,-w_1} * \wp_{v_0+w_2,-w_2} \tag{15.399}$$

one obtains the holonomy

$$U(\wp_{v_0,w_1,w_2}) : (z, v_0) \to (z e^{2\pi\mathrm{i}\Omega(w_1,w_2)}, v_0) \tag{15.400}$$

showing that the curvature of this connection is $\Omega$ regarded as a 2-form on $V$. The connection and 2-form descend to the principal $U(1)$ bundle $P(T,\Omega,\epsilon)$. The cohomology class of $[\Omega]$, which is the first Chern class of $P(T,\Omega,\epsilon)$ is characterized by the integers $\vec{d} = (d_1, d_2, \dots)$. A nonzero value of $\vec{d}$ obstructs topological triviality.

---

**Exercise**

We illustrated how $Q$ and $D_4$ are the only two non-Abelian groups that sit in an extension of $\mathbb{Z}_2 \times \mathbb{Z}_2$ by $\mathbb{Z}_2$. Which one is the Heisenberg extension?

---

---

**Exercise** *Finite Heisenberg group in multiplicative notation*

It is interesting to look at the Heisenberg extension

$$1 \to \mathbb{Z}_n \to \mathrm{Heis}(\mathbb{Z}_n \times \mathbb{Z}_n) \to \mathbb{Z}_n \times \mathbb{Z}_n \to 1 \tag{15.401}$$

where we think of $\mathbb{Z}_n$ as the *multiplicative* group of $n^{th}$ roots of unity. Let $\omega = \exp[2\pi\mathrm{i}/n]$. We distinguish the three $\mathbb{Z}_n$ factors by writing generators as $\omega_1, \omega_2, \omega_3$.

a.) Show that one natural choice of cocycle is:

$$f\left((\omega_1^s, \omega_2^t), (\omega_1^{s'}, \omega_2^{t'})\right) := \omega_3^{st'} \tag{15.402}$$

b.) Compute the commutator function

$$\kappa\left((\omega_1^s, \omega_2^t), (\omega_1^{s'}, \omega_2^{t'})\right) := \omega_3^{st'-ts'} \tag{15.403}$$

c.) Connect to our general theory of extensions by defining $U := (1, (\omega_1, 1))$, $V := (1, (1, \omega_2))$ and computing

$$\begin{aligned} UV &= (f((\omega_1, 1), (1, \omega_2)), (\omega_1, \omega_2)) \\ &= (\omega_3, (\omega_1, \omega_2)) \\ VU &= (f((1, \omega_2), (\omega_1, 1)), (\omega_1, \omega_2)) \\ &= (1, (\omega_1, \omega_2)) \end{aligned} \tag{15.404}$$

or in other words, since the center is generated by $q = (\omega_3, (1, 1))$ we can write:

$$UV = qVU \tag{15.405}$$

---

---

**Exercise** *Degenerate Heisenberg extensions*

Suppose $n = km$ is composite and suppose we use the function $c_k(a, b') = kab'$ in defining an extension of $\mathbb{Z}_n \times \mathbb{Z}_n$.

a.) Show that the commutator function is now degenerate.

b.) Show that the center of the central extension is larger than $\mathbb{Z}_n$. Compute it. [262]

While these are not - strictly speaking - Heisenberg extensions people will often refer to them as Heisenberg extensions. We might call them "degenerate Heisenberg extensions."

---

---

**Exercise** *Constructing The Splitting* (15.394)

Give an explicit construction of a function $\epsilon : \Gamma \to U(1)$ satisfying (15.394). [263]

---

---

**Exercise** *Different Splittings Give Isomorphic Bundles*

a.) Describe the relation between two splittings of the pullback of the sequence (15.390) to $\Gamma$.

b.) Let $\epsilon_1$, $\epsilon_2$ denote two splittings of the pullback of the sequence (15.390) to $\Gamma$. Show that the bundles $P(T, \Omega, \epsilon_1)$ with $P(T, \Omega, \epsilon_2)$.

---

[262] *Answer*: The center is generated by $q, U^m, V^m$ and is $\mathbb{Z}_n \times \mathbb{Z}_k \times \mathbb{Z}_k$.

[263] *Answer*: Choose an ordered basis $\gamma_1, \dots, \gamma_n$ for $\Gamma$ and define

$$\epsilon_\gamma := e^{-i\pi \sum_{i<j} n_i n_j \Omega_{ij}} \tag{15.406}$$

where $\gamma = \sum_i n_i \gamma_i$ and $\Omega_{ij} = \Omega(\gamma_i, \gamma_j)$. One can check this satisfies the desired identity. Note that it is crucial that $\Omega_{ij}$ and $n_i$ are integral.

**Exercise** *Two Dimensions*
a.) Suppose $T = \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ with $\text{Im}\tau > 0$. So $\Gamma = \mathbb{Z} + \tau\mathbb{Z}$. Choose

$$\Omega(1,\tau) = -\Omega(\tau,1) = k \in \mathbb{Z}. \tag{15.407}$$

b.) Show that

$$\Omega(z_1,z_2) = k\frac{\text{Im}(\bar{z}_1 z_2)}{\text{Im}\tau} \tag{15.408}$$

### 15.5.7 Lagrangian Subgroups And Induced Representations

Let us compare the general Heisenberg extension

$$1 \to \mathcal{Z} \to \tilde{G} \to G \to 0 \tag{15.409}$$

with (15.331) and (15.371) with $\mathcal{Z}$ any subgroup of $U(1)$. The difference from the general case is that in these examples $G$ is explicitly presented as a product of subgroups $G = L \times L'$ where $L$ and $L'$ are *maximal Lagrangian subgroups*. A subgroup $L \subset G$ is said to be a *Lagrangian subgroup* if $\kappa(g_1,g_2) = 1$ for all pairs $(g_1,g_2) \in L$ and similarly for $L'$.

When discussing Heisenberg groups of the form $\text{Heis}(S \times \widehat{S})$ the group $S \times \widehat{S}$ has two canonical Lagrangian subgroups, namely $S$ and $\widehat{S}$. With the choice of $\kappa$ we made above these are maximal Lagrangian subgroups.

The case of $G = S \times \widehat{S}$ should be contrasted with other examples where $G$ is $\mathbb{R}^{2n}$ or $\mathbb{Z}_2^{2n}$. These groups certainly can be presented as products of maximal Lagrangian subgroups, but there is no canonical decomposition. Consider, for example the Heisenberg extension $\text{Heis}(\mathbb{Z}_2^{2n})$ constructed using the gamma-matrices. Note that $\mathbb{Z}_2 = \mathbb{F}_2$ is a field, and we can consider $\mathbb{F}_2^{2n}$ to be a vector space over $\kappa = \mathbb{F}_2$. We could take $L$ to be any half-dimensional Lagrangian subspace.

In the general situation, with no canonical choice of $L$ one would often like to construct an explicit unitary representation of the Heisenberg group. One way to do this is the following:

Choose a maximal Lagrangian subgroup $L \subset G$. The inverse image $\tilde{L} \subset \tilde{G}$ is a maximal commutative subgroup of $\tilde{G}$.

We now choose a character of $\tilde{L}$ such that $\rho(z,x) = z\rho(x)$. Note that such a character must satisfy

$$\rho(x)\rho(x') = f(x,x')\rho(x+x') \qquad \forall x,x' \in L \tag{15.410}$$

♣This is bad notation, because $\rho(z,x)$ is later used for the full SvN representation. The letter here should be changed and it propagates through the discussion. ♣

Note that $f$ need not be trivial on $L$, but it does define an Abelian extension $\tilde{L}$ which must therefore be isomorphic to a product $L \times U(1)$, albeit noncanonically. Different choices of $\rho$ are different choices of splitting. Indeed note that (15.410) says that, when restricted to $L \times L$ the cocycle is trivialized by $\rho$.

The carrier space of our representation will be the space $\mathcal{F}$ of functions $\psi : \tilde{G} \to \mathbb{C}$ such that

$$\psi\left((z,x)(z',x')\right) = \rho(z',x')^{-1}\psi\left(z,x\right) \qquad \forall (z',x') \in \tilde{L} \tag{15.411}$$

Setting $(z',x') = (z^{-1},0)$ we note that equation (15.411) implies $\psi(z,x) = z^{-1}\psi(1,x)$, so defining $\Psi(x) := \psi(1,x)$ we can simplify the description of $\mathcal{F}$ by identifying it with the space of functions $\Psi : G \to \mathbb{C}$ such that:

$$\Psi\left(x + x'\right) = \frac{f(x,x')}{\rho(x')}\Psi(x) \qquad \forall x' \in L. \tag{15.412}$$

If our group is continuous or noncompact we should state an $L^2$ condition. We take $\rho$ to be a unitary character so that $|\Psi(x)|^2$ descends to a function on $G/L$ and we demand:

$$\int_{G/L} |\Psi(x)|^2 dx < +\infty. \tag{15.413}$$

The group action is simply left-action of $\tilde{G}$ on the functions $\psi(z,x)$. When written in terms of $\Psi(x)$ the representation of $(x,z) \in \tilde{G}$ is:

$$\begin{aligned}(\rho(z,x)\cdot\Psi)(y) &= (\rho(z,x)\cdot\psi)(1,y)\\ &= \psi\left((z,x)^{-1}\cdot(1,y)\right)\\ &= \psi\left((z^{-1}f(x,-x)^{-1}f(-x,y), y-x)\right)\\ &= zf(x,y-x)\Psi(y-x)\end{aligned} \tag{15.414}$$

where in the last line we assumed that $f$ is a normalized cocycle so that $f(0,y) = 1$.

Let us see how we recover the standard Stone-von Neumann representation of Heis$(S \times \widehat{S})$ from this viewpoint. Let us choose $L = \widehat{S}$. Then the equivariance condition (15.412) becomes

$$\Psi(s,\chi\chi') = \frac{1}{\rho(\chi')}\Psi(s,\chi) \tag{15.415}$$

Now set $\chi' = 1/\chi$ and conclude that

$$\Psi(s,\chi) = \frac{1}{\rho(\chi)}\Psi(s,1) \tag{15.416}$$

So the dependence on $\chi \in \widehat{S}$ is completely fixed by equivariance. Defining $\tilde{\psi}(s) := \Psi(s,1)$ we obtain a vector $\tilde{\psi} \in L^2(S)$, and thus if we take $L = \widehat{S}$ then our space of equivariant functions is naturally identified with $L^2(S)$.

We can now work out the representations of

$$\begin{aligned}T_s &= \rho(1,(s,1))\\ M_\chi &= \rho(1,(0,\chi))\end{aligned} \tag{15.417}$$

on $L^2(S)$. Working through the above definitions it should not be surprising that one recovers:

$$\begin{aligned}(T_s\tilde{\psi})(s_0) &= \tilde{\psi}(s_0 - s)\\ (M_\chi\tilde{\psi})(s_0) &= \frac{\rho(\chi)}{\chi(s_0)}\tilde{\psi}(s_0)\end{aligned} \tag{15.418}$$

**Example**. Consider $\mathbb{F}_2^{2m}$ with $\kappa(w,w') = (-1)^{\sum_{i\neq j} w_i w'_j}$. We can give a Lagrangian decomposition

$$\mathbb{F}_2^{2m} \cong L \oplus N \tag{15.419}$$

in many different ways. For special values of $m$ there are special Lagrangian subspaces provided by classical error correcting codes. A Heisenberg representation can be given by taking $V$ to be the space of functions $\psi : L \to \mathbb{C}$. This has complex dimension $2^m$. $N$ can be identified with the group of characters on $L$ since we can set

$$\chi_n(\ell) = \kappa(n,\ell) \tag{15.420}$$

The usual translation and multiplication operators $T_\ell$ and $M_n$ generate an algebra isomorphic to $Mat_d(\mathbb{C})$. $V$ is also a representation of the Clifford algebra (and hence the extra-special group). So the Clifford representation matrices $\gamma_i$ can be expressed in terms of these, and vice versa.

**Remarks**

1. The representation is, geometrically, just the space of $L^2$-sections of the associated line bundle $\tilde{G} \times_{\tilde{L}} \mathbb{C}$ defined by $\rho$. The representation is independent of the choice of $\rho$, and any two choices are related by an automorphism of $L$ given by the restriction of an inner automorphism of $\tilde{G}$.
2. This is an example of an *induced representation*, $Ind_{\tilde{L}}^{\tilde{G}}(\mathbb{C})$ which we will study more systematically in chapter 4.

---

**Exercise**

Construct explicit Lagrangian subspaces of $\mathbb{F}_2^{2m}$ for small values of $m$ and write out the matrices of the Heisenberg representation. [264]

---

### 15.5.8 Automorphisms Of Heisenberg Extensions

In several of our examples above, such as $\mathrm{Heis}(V,\Omega)$ and the extraspecial group we have noted that to define a representation one must make a choice of a Lagrangian subgroup, where $\kappa$ restricts to 1. On the other hand, we also noted that in some situations there is no natural choice of such a Lagrangian group. There are many such Lagrangian subgroups and they are related by "symplectic automorphisms" of $G$.

We say that *an automorphism* $\alpha \in \mathrm{Aut}(G)$ *is symplectic* if it preserves the commutator function:

$$\begin{aligned} \alpha^*\kappa(g_1,g_2) &:= \kappa(\alpha(g_1),\alpha(g_2)) \\ &= \kappa(g_1,g_2) \end{aligned} \tag{15.421}$$

[264] *Answer.* Start with $m = 1$. Any vector is isotropic so let $L$ be spanned by $\ell = (1,0)$ and $N$ spanned by $n = (0,1)$. Choose a basis for $V$ by taking the delta function supported on basis vector $e_i$. Then $T_\ell = \sigma^1$ and $M_n = \sigma_3$ relative to this basis.*********** CONTINUE ***********.

In the first line we defined the general notion of pullback $\kappa \to \alpha^*\kappa$ and the second line is the invariance condition. In physics, such symplectic transformations are relevant to canonical transformations. We can ask whether such automorphisms of "phase space" actually lift to automorphisms of the full Heisenberg group, and then whether and how this lifted group acts on the representations of the Heisenberg group. This would be the "quantum mechanical implementation of symplectic transformations." In this section we will investigate those questions from the group-theoretical viewpoint.

Quite generally, (there is no need for $G$ to be Abelian in this paragraph), if $\pi : \widetilde{G} \to G$ is a homomorphism and $\alpha \in \mathrm{Aut}(G)$ is an automorphism of $G$ we say that $\alpha$ *lifts to an automorphism of* $\widetilde{G}$ if there is an automorphism $\widetilde{\alpha} \in \mathrm{Aut}(\widetilde{G})$ that completes the diagram:

$$\begin{array}{ccc} \widetilde{G} & \xrightarrow{\pi} & G \\ \downarrow{\scriptstyle \widetilde{\alpha}} & & \downarrow{\scriptstyle \alpha} \\ \widetilde{G} & \xrightarrow{\pi} & G \end{array} \tag{15.422}$$

Or, in equations, for every $\alpha \in \mathrm{Aut}(G)$ we seek a corresponding $\widetilde{\alpha} \in \mathrm{Aut}(\widetilde{G})$ such that

$$\pi(\widetilde{\alpha}(\widetilde{g})) = \alpha(\pi(\widetilde{g})) \tag{15.423}$$

for all $\widetilde{g} \in \widetilde{G}$.

<u>Some General Theory</u> [265]

Consider a group extension

$$1 \to A \to \widetilde{G} \to G \to 1 \tag{15.424}$$

where $A$ and $G$ are Abelian and we write the group operations on both $A$ and $G$ additively so $\widetilde{G}$ is the group of pairs $(a, g)$ with group multiplication

$$(a_1, g_1)(a_2, g_2) = (a_1 + a_2 + f(g_1, g_2), g_1 + g_2) \tag{15.425}$$

and $f(g_1, g_2)$ is a cocycle satisfying the additive version of the cocycle identity:

$$f(g_1 + g_2, g_3) + f(g_1, g_2) = f(g_1, g_2 + g_3) + f(g_2, g_3) \tag{15.426}$$

Now suppose $\alpha \in \mathrm{Aut}(G)$. If $\alpha$ preserves the commutator function then we can hope to lift it to an automorphism $T_\alpha$ of $\widetilde{G}$. As explained above, "lifting" means that

$$\pi(T_\alpha(a, g)) = \alpha(\pi(a, g)) = \alpha(g)\ . \tag{15.427}$$

Therefore, $T_\alpha$ must be of the form:

$$T_\alpha(a, g) = (\xi_\alpha(a, g), \alpha(g)) \tag{15.428}$$

[265] What follows is an elaboration of the ideas from Appendix A of arXiv:1707.08888. I also got some useful help from Graeme Segal.

where $\xi_\alpha$ is some function $\xi_\alpha : A \times G \to A$. We can write constraints on this function from the requirement that $T_\alpha$ must be an automorphism with the group law defined by the cocycle $f$. In particular $T_\alpha$ must be a group homomorphism:

$$T_\alpha((a_1, g_1), (a_2, g_2)) = T_\alpha(a_1, g_1) \cdot T_\alpha(a_2, g_2) \tag{15.429}$$

which is true iff

$$\xi_\alpha(a_1 + a_2 + f(g_1, g_2), g_1 + g_2) = \xi_\alpha(a_1, g_1) + \xi_\alpha(a_2, g_2) + f(\alpha(g_1), \alpha(g_2)) \tag{15.430}$$

Now, specialize this equation by putting $g_1 = 0$ and assuming (WLOG) that we have a normalized cocycle, so that $f(g, 0) = f(0, g) = 0$. Then equation (15.430) simplifies to

$$\xi_\alpha(a_1 + a_2, g) = \xi_\alpha(a_1, 0) + \xi_\alpha(a_2, g) \tag{15.431}$$

Putting $a_2 = 0$ in (15.431) we now learn that

$$\xi_\alpha(a, g) = \xi_\alpha(a, 0) + \xi_\alpha(0, g) \tag{15.432}$$

On the other hand, putting $g = 0$ in (15.431) we now learn that $a \mapsto \xi_\alpha(a, 0)$ is just an automorphism of $A$. Composing lifts of $\alpha$ with such automorphisms is an inherent ambiguity in lifting $\alpha$. Thus, it is useful to make the simplifying assumption that $\xi_\alpha(a, 0) = a$, since we can always arrange this by composition with an automorphism of $A$. Therefore we can write equation (15.432) in the general form

$$T_\alpha(a, g) = (a + \tau_\alpha(g), \alpha(g)) \tag{15.433}$$

If we restrict to automorphisms of the type (15.433) then one easily checks that $T_\alpha$ is indeed a group homomorphism iff

$$(\alpha^* f - f)(g_1, g_2) = \tau_\alpha(g_1 + g_2) - \tau_\alpha(g_1) - \tau_\alpha(g_2) \tag{15.434}$$

where

$$\alpha^* f(g_1, g_2) := f(\alpha(g_1), \alpha(g_2)) \tag{15.435}$$

is known as the "pulled-back cocycle." The conceptual meaning of this equation is the following: $\alpha \in \mathrm{Aut}(G)$ is a "symmetry of $G$." We are asking how badly the cocycle $f$ breaks that symmetry. We say that "$f$ is invariant under pullback" if $\alpha^* f = f$ and in that case we can take $\tau_\alpha = 0$ and we can easily lift the group $\mathrm{Aut}(G)$ to a group of automorphisms of $\widetilde{G}$. The more general condition (15.434) says that the amount by which $f$ is not symmetric under $\alpha$, that is $\alpha^* f - f$, must be a trivializable cocycle. Put this way, it is clear that the condition is unchanged under shifting $f$ by a coboundary. Indeed, if we change $f$ by a coboundary so

$$\tilde{f}(g_1, g_2) = f(g_1, g_2) + b(g_1 + g_2) - b(g_1) - b(g_2) \tag{15.436}$$

then we can solve (15.434) by taking

$$\tilde{\tau}_\alpha(g) = \tau_\alpha(g) + (\alpha^* b - b)(g) \tag{15.437}$$

so the existence of a solution to (15.434) is gauge invariant. Put more simply, *the cohomology class* $[f] \in H^2(G,A)$ *must be invariant under the action of* $\alpha^*$ *on* $H^2(G,A)$.

♣The above paragraph should be said more succinctly. ♣

Thus, in general we cannot lift all automorphisms of $G$, only those for which (15.434) holds. The set of such automorphisms forms a subgroup of $\mathrm{Aut}(G)$ that we will denote as $\mathrm{Aut}_0(G)$. Note that, in our additive notation we have

$$\kappa(g_1,g_2) = f(g_1,g_2) - f(g_2,g_1) \tag{15.438}$$

which, as mentioned above, is a generalization of a symplectic form. Any automorphism satisfying (15.434) automatically satisfies $\alpha^*(\kappa) = \kappa$ and is therefore a "symplectic automorphism." However, $\mathrm{Aut}_0(G)$ is in general a subgroup of the group of symplectic automorphisms of $G$. We will give an example with $G = SL(2,\mathbb{Z}_n)$ below.

We now restrict attention to $\alpha \in \mathrm{Aut}_0(G)$. If, in addition

$$\tau_{\alpha_1\circ\alpha_2}(g) = \tau_{\alpha_1}(\alpha_2(g)) + \tau_{\alpha_2}(g) \tag{15.439}$$

that is,

$$\tau_{\alpha_1\circ\alpha_2} = \alpha_2^*\tau_{\alpha_1} + \tau_{\alpha_2} \tag{15.440}$$

then in fact $T_{\alpha_1}\circ T_{\alpha_2} = T_{\alpha_1\circ\alpha_2}$ generate a subgroup of $\mathrm{Aut}(\widetilde{G})$ isomorphic to $\mathrm{Aut}(G)$.

♣ Relate this to equation (15.521)) the condition for a twisted homomorphism $\tau : \mathrm{Aut}(G) \to A$. ♣

In general, even if we can find a solution to (15.434) the criterion (15.440) will not hold. Nevertheless, the automorphisms $T_\alpha$ will generate a subgroup of $\mathrm{Aut}(\widetilde{G})$. To see what subgroup it is we introduce, for $\ell \in \mathrm{Hom}(G,A)$, the automorphism

$$P_\ell(a,g) = (a + \ell(g), g) \tag{15.441}$$

Then we note that

$$T_{\alpha_1}\circ T_{\alpha_2}(a,g) = T_{\alpha_1\circ\alpha_2}\circ P_{\ell_{\alpha_1,\alpha_2}} \tag{15.442}$$

where

$$\ell_{\alpha_1,\alpha_2}(g) := \tau_{\alpha_1}(\alpha_2(g)) + \tau_{\alpha_2}(g) - \tau_{\alpha_1\circ\alpha_2}(g) \tag{15.443}$$

A little computation shows that $\ell_{\alpha_1,\alpha_2} \in \mathrm{Hom}(G,A)$ is indeed a homomorphism from $G$ to $A$. A little more computation reveals

$$\begin{aligned} P_{\ell_1}\circ P_{\ell_2} &= P_{\ell_1+\ell_2} \\ T_\alpha \circ P_\ell &= P_{\alpha^*(\ell)}\circ T_\alpha \end{aligned} \tag{15.444}$$

Therefore, we can write any word in $P$'s and $T$'s in the form $P_{\ell'}\circ T_{\alpha'}$ for some $(\ell',\alpha')$. Altogether, equations (15.442), (15.443), and (15.444) mean that $T_\alpha$ generate a subgroup of $\mathrm{Aut}(\widetilde{G})$. Including all transformations $P_\ell$ defines a subgroup $\widetilde{\mathrm{Aut}(G)}$ of $\mathrm{Aut}(\widetilde{G})$ which fits in an exact sequence:

$$1 \to \mathrm{Hom}(G,A) \to \widetilde{\mathrm{Aut}(G)} \to \mathrm{Aut}_0(G) \to 1 \tag{15.445}$$

Finally, restoring $\mathrm{Hom}(G,A) \to \mathrm{Hom}(G,A) \rtimes \mathrm{Aut}(A)$ gives the group of (possible) lifts of automorphisms of $G$ to automorphisms of $\tilde{G}$.

Example: Heis($\mathbb{R}\oplus\mathbb{R}$)

Let us consider the basic example from quantum mechanics based on a phase space $\mathbb{R}\oplus\mathbb{R}$ with symplectic form defined by $J$. In this case there is a very nice way of thinking of the matrix group $Sp(2,\mathbb{R})$. We note that if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{15.446}$$

then

$$\begin{aligned} A^{tr} J A &= \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \\ &= (ad-bc)J \end{aligned} \tag{15.447}$$

Therefore, $A \in Sp(2,\mathbb{R})$ iff $ad - bc = 1$. But this is precisely the condition that defines $SL(2,\mathbb{R})$. Therefore

$$SL(2,\mathbb{R}) = Sp(2,\mathbb{R}) \tag{15.448}$$

are identical as matrix groups. The same argument applies if we replace $\mathbb{R}$ by any ring $R$. This kind of isomorphism is definitely not true if we consider higher rank groups $SL(n,\mathbb{R})$ and $Sp(2n,\mathbb{R})$.

Now, as we have seen, the section $s(\alpha,\beta) = \exp[\mathrm{i}(\alpha\hat{p} + \beta\hat{q})]$ leads to the choice of section, written additively

$$f\left((\alpha_1,\beta_1),(\alpha_2,\beta_2)\right) = \frac{\hbar}{2}(\alpha_1\beta_2 - \alpha_2\beta_1) \tag{15.449}$$

This is symplectic invariant, so we can take $\tau_\alpha(g) = 0$ in the equation (15.434) and we conclude that the symplectic group acts as a group of automorphisms on the Heisenberg group Heis($\mathbb{R}^{2n}$).

It is interesting however, that the symplectic group only acts projectively on the Stone-von-Neumann representation of the Heisenberg group. We now explain this point.

As we will discuss in detail in the chapter on Lie groups, $SL(2,\mathbb{R})$ and $Sp(2,\mathbb{R})$ are examples of Lie groups. It is useful to look at group elements infinitesimally close to the identity matrix. These can be written as

$$A = 1 + \epsilon m + \mathcal{O}(\epsilon^2) \tag{15.450}$$

We learn from the defining equation that

$$\mathrm{Tr}(m) = 0 \tag{15.451}$$

is required to satisfy the defining conditions to order $\epsilon$. (Exercise: Prove this using both the definition of $Sp(2,\mathbb{R})$ and of $SL(2,\mathbb{R})$.)

The infinitesimal group elements are thus characterized by the vector space $\mathfrak{sp}(2,\mathbb{R}) = \mathfrak{sl}(2,\mathbb{R})$, which is the vector space of $2\times 2$ real traceless matrices. These form a Lie algebra with the standard matrix commutator. Generic (but not all) group elements are obtained by exponentiating such matrices. In particular, in a neighborhood of the identity all group elements are obtained by exponentiating elements of the Lie algebra.

Recall the standard basis of $\mathfrak{sl}(2,\mathbb{R})$:

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad h = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad f = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \tag{15.452}$$

(Check the signs carefully!) Compute:

$$[h,e] = -2e \qquad [e,f] = h \qquad [h,f] = +2f \tag{15.453}$$

From this one can in principle multiply exponentiated matrices using the BCH formula.

We now consider the quantum implementation of these operators on $L^2(\mathbb{R})$ with $\rho(e) = \hat{e}$ etc. with:

$$\hat{e} := \frac{\mathrm{i}}{2\hbar}\hat{p}^2 \qquad \hat{h} := \frac{\mathrm{i}}{2\hbar}(\hat{q}\hat{p} + \hat{p}\hat{q}) \qquad \hat{f} := \frac{\mathrm{i}}{2\hbar}\hat{q}^2 \tag{15.454}$$

Now, using the useful identities: [266]

$$\begin{aligned} [AB,CD] &= A[B,C]D + [A,C]BD + CA[B,D] + C[A,D]B \\ &= AC[B,D] + A[B,C]D + C[A,D]B + [A,C]DB \end{aligned} \tag{15.456}$$

we can check that this is indeed a representation:

$$[\hat{h},\hat{e}] = -2\hat{e} \qquad [\hat{e},\hat{f}] = \hat{h} \qquad [\hat{h},\hat{f}] = +2\hat{f} \tag{15.457}$$

However, we have to be careful about exponentiating these operators.

Let us consider the one-parameter subgroup of $SL(2,\mathbb{R})$:

$$\exp[\theta(e+f)] = \cos\theta + \sin\theta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{15.458}$$

This is in fact a maximal compact subgroup of $SL(2,\mathbb{R})$. (See chapter **** on 2x2 matrix groups.) Note carefully that it has period $\theta \sim \theta + 2\pi$.

The quantum implementation of $e+f$ is just the standard harmonic oscillator Hamiltonian!

$$\hat{e} + \hat{f} = \frac{\mathrm{i}}{2}(\hat{p}^2 + \hat{q}^2) = \mathrm{i}(\bar{a}a + \frac{1}{2}) \tag{15.459}$$

where

$$\begin{aligned} a &= \frac{1}{\sqrt{2}}(q + \mathrm{i}p) \\ \bar{a} &= \frac{1}{\sqrt{2}}(q - \mathrm{i}p) \end{aligned} \tag{15.460}$$

[266] These identities are very useful, but a bit hard to remember. If you are on a desert island you can easily reconstruct them from the special cases:

$$\begin{aligned} [AB,C] &= A[B,C] + [A,C]B \\ [A,CD] &= C[A,D] + [A,C]D \end{aligned} \tag{15.455}$$

both of which are easy to remember.

Now, in the Stone-von-Neumann representation $\frac{\mathrm{i}}{2}(\hat{p}^2 + \hat{q}^2)$ has the spectrum $\mathrm{i}(n + \frac{1}{2})$, $n = 0, 1, 2, \dots$. Therefore, the one-parameter subgroup $\exp[\theta(\hat{e}+\hat{f})]$ has period $\theta \sim \theta + 4\pi$. We see that the group generated by $\hat{e}, \hat{f}, \hat{h}$ is at least a double cover of $Sp(2, \mathbb{R})$. In fact, it turns out to be exactly a double cover, and it is known as the *metaplectic group*.

One very interesting aspect of the metaplectic group is that this is a Lie group with no finite-dimensional faithful representation. We now explain that fact, [267] and a few other important things in the following remarks:

**Remarks**

1. *Application to the metaplectic group.* The Lie algebra of the metaplectic group is $\mathfrak{sl}(2, \mathbb{R})$. Any finite dimensional representation of the metaplectic group would give a finite-dimensional representation of the Lie algebra $\mathfrak{sl}(2, \mathbb{R})$ and we discussed in detail what these are above. Note that $e + f$ corresponds to $\mathrm{i}J^1$. So: In any finite-dimensional complex representation of $\mathfrak{sl}(2, \mathbb{R})$, the operator $\rho(e) + \rho(f)$ is diagonalizable and all the eigenvalues are of the form $\mathrm{i}\ell$, where $\ell$ is an integer. So $\exp[\theta\rho(e + f)]$ has period $\theta \sim \theta + 2\pi$. Therefore no finite dimensional representation of $Mpl$ can be faithful. In particular, $Mpl$ is an example of a Lie group which is not a matrix group: It cannot be embedded as a subgroup of $GL(N, \mathbb{C})$ for any $N$.

2. It is very interesting to consider the action of the one-parameter family $\exp[\theta(\hat{e} + \hat{f})]$ in the standard "position space" representation $L^2(\mathbb{R})$ with $\hat{p} = -\mathrm{i}\frac{d}{dq}$. Let us compute the integral kernel:
$$\langle x|\exp[\theta(\hat{e} + \hat{f})]|y\rangle \tag{15.461}$$
since
$$\left(e^{\theta(\hat{e}+\hat{f})}\psi\right)(x) = \int_{-\infty}^{+\infty} \langle x|\exp[\theta(\hat{e} + \hat{f})]|y\rangle\psi(y)dy \tag{15.462}$$
Clearly, to evaluate (15.461) we should insert a complete set of eigenstates of the harmonic oscillator Hamiltonian. So we introduce the *Hermite functions*: [268]
$$\psi_n(x) := (2^n n!\sqrt{\pi})^{-1/2} e^{-\frac{x^2}{2}} H_n(x) \qquad n \in \mathbb{Z}_+ \tag{15.463}$$
where $H_n(x)$ is the Hermite polynomial
$$H_n(x) = e^{x^2}\left(-\frac{d}{dx}\right)^n e^{-x^2} \tag{15.464}$$
The Hermite functions $\psi_n(x)$ satisfy
$$\left(-\frac{d^2}{dx^2} + x^2\right)\psi_n = (2n + 1)\psi_n \tag{15.465}$$

[267] Our demonstration of this surprising fact follows the discussion in G. Segal in *Lectures On Lie Groups and Lie Algebras.*

[268] What follows here are completely standard facts. We used Wikipedia.

and we have normalized them so that so that

$$\int_{-\infty}^{+\infty} \psi_n(x)\psi_m(x)dx = \delta_{n,m} \tag{15.466}$$

Now we have

$$\sum_{n=0}^{\infty} u^n \psi_n(x)\psi_m(x) = \frac{1}{\sqrt{\pi(1-u^2)}}\exp\left[-\frac{1-u}{1+u}\left(\frac{x+y}{2}\right)^2 - \frac{1+u}{1-u}\left(\frac{x-y}{2}\right)^2\right] \tag{15.467}$$

To prove this write

$$H_n(x) = e^{x^2}\left(-\frac{d}{dx}\right)^n e^{-x^2} = e^{x^2}\left(-\frac{d}{dx}\right)^n\left(\frac{1}{\sqrt{4\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{4}s^2+\mathrm{i}sx}ds\right) \tag{15.468}$$

Apply this to both $\psi_n(x)$ and $\psi_n(y)$ apply the derivatives, exchange integration and sum and get

$$\begin{aligned}\sum_{n=0}^{\infty} u^n \psi_n(x)\psi_m(x) &= \frac{e^{\frac{1}{2}(x^2+y^2)}}{4\pi^{3/2}}\int dsdte^{-\frac{1}{4}(s^2+t^2)-\frac{1}{2}stu+\mathrm{i}sx+\mathrm{i}ty}\\ &= \frac{1}{\sqrt{\pi(1-u^2)}}\exp\left[-\frac{1-u}{1+u}\left(\frac{x+y}{2}\right)^2 - \frac{1+u}{1-u}\left(\frac{x-y}{2}\right)^2\right]\end{aligned} \tag{15.469}$$

Now we have

$$\langle x|\exp[\theta(\hat{e}+\hat{f})]|y\rangle = e^{\mathrm{i}\theta/2}\sum_{n=0}^{\infty} e^{\mathrm{i}n\theta}\psi_n(x)\psi_n(y) \tag{15.470}$$

so we apply the above identity with $u = e^{\mathrm{i}\theta}$. We should be careful about convergence: The Gaussian integral in (15.469) has quadratic form

$$A = \frac{1}{4}\begin{pmatrix} 1 & u \\ u & 1 \end{pmatrix} \tag{15.471}$$

which has eigenvalues $\frac{1}{4}(1 \pm u)$. The zero-modes at $u = \pm 1$ indicate a divergent Gaussian. In fact we have

$$\begin{aligned}\sum_{n=0}^{\infty}\psi_n(x)\psi_n(y) &= \delta(x-y)\\ \sum_{n=0}^{\infty}(-1)^n\psi_n(x)\psi_n(y) &= \delta(x+y)\end{aligned} \tag{15.472}$$

The second line follows easily from the first since the parity of $\psi_n(x)$ as a function of $x$ is the parity of $n$.

The quadratic form $A$ has a positive definite real part for $|u| \le 1$ except for $u = \pm 1$. The values for $|u| > 1$ have to be defined by analytic continuation and there is a branch point at $u = \pm 1$. Note that at $\theta = \pi/2$ we have

$$\langle x|\exp[\frac{\pi}{2}(\hat{e}+\hat{f})]|y\rangle = \frac{e^{\mathrm{i}\pi/4}}{\sqrt{2\pi}}e^{\mathrm{i}xy} \tag{15.473}$$

and we recognize the kernel for the Fourier transform:

$$e^{\frac{\pi}{2}(\hat{e}+\hat{f})}\psi = e^{\mathrm{i}\pi/4}\mathcal{F}(\psi) \tag{15.474}$$

where $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is the Fourier transform. This is the quantum implementation of the canonical transformation exchanging position and momenta. Note that $\theta = \pi$ is the square of the Fourier transform (up to a scalar multiplication by i) but this is not a scalar operator on the space of functions. Indeed at $\theta = \pi$ we have

$$\left(e^{\pi(\hat{e}+\hat{f})}\psi\right)(x) = \mathrm{i}\psi(-x) \tag{15.475}$$

The Fourier transform is of order four not order two. Also note that at $\theta = 2\pi$ the operator is just multiplication by $-1$.

3. We note that there are beautiful general formulae for expectation value operators defined by exponentiating general quadratic forms in the $\hat{p}_i$ and $\hat{q}^i$, or, equivalently in $a$'s and $a^\dagger$'s. This is useful when working with coherent states and squeezed states. But it is best presented in the Bargmann or geometric quantization formalism. ♣say more? ♣

---

**Exercise**

a.) Check that, for $2 \times 2$ matrices the condition $\mathrm{Tr}(m) = 0$ is identical to the condition $(mJ)^{tr} = mJ$.

b.) Show that for any $n \times n$ matrix, the infinitesimal version of the condition $\det A = 1$ is that $A = 1 + \epsilon m + \mathcal{O}(\epsilon^2)$ with $\mathrm{Tr}(m) = 0$.

c.) Show that for any $n \times n$ matrix, the infinitesimal version of the condition $A^{tr} J A = J$ is that $A = 1 + \epsilon m + \mathcal{O}(\epsilon^2)$ with $(mJ)^{tr} = mJ$.

d.) Show that the conditions $\mathrm{Tr}(m) = 0$ and $(mJ)^{tr} = mJ$ are inequivalent different for $n > 2$.

---

♣Following exercise belongs in the Linear Algebra chapter. ♣

---

**Exercise** *Linear independence of eigenvectors with distinct eigenvalues*

Suppose a set of nonzero vectors $v_1, \dots, v_n$ are eigenvectors of some operator $A$ with distinct eigenvalues. Show that they are linearly independent. [269]

---

[269] *Answer*: Suppose $\sum_s c_s v_s = 0$ for some coefficients $c_s$. Applying powers of $A$ we determine that $\sum c_s \lambda_s^k v_s = 0$. If all the $c_s$ are nonzero then we learn that the matrix the matrix $VC$ must have determinant zero where $V_{ij} = \lambda_i^j$ and $C$ is the diagonal matrix with entries $c_1, \dots, c_n$. If some of the $c_s$ are nonzero then we have a minor of the matrix $V$ times the diagonal matrix of the nonzero $c_s$. In any case, none of the minors of $V_{ij}$ have zero determinant, provided the $\lambda_i$ are distinct. Therefore, $VC$ (or the appropriate minor) has nonzero determinant and hence no kernel. So $\{v_s\}$ is a linearly independent set.

**Exercise** $\mathfrak{su}(2)$ *vs.* $\mathfrak{sl}(2,\mathbb{R})$

A basis for the real Lie algebra of $2\times 2$ traceless anti-Hermitian matrices is

$$T^i = -\frac{\mathrm{i}}{2}\sigma^i \qquad\qquad i = 1,2,3 \tag{15.476}$$

with Lie algebra

$$[T^i, T^j] = \epsilon^{ijk} T^k \tag{15.477}$$

Can one make real linear combinations of $T^i$ to produce the generators $e, h, f$ of $\mathfrak{sl}(2,\mathbb{R})$ above?

---

**Exercise** *Representation matrices*

a.) Show that if $v_0$ is the vector used above with $\rho(e)v_0 = 0$ and $\rho(h)v_0 = -Nv_0$ then [270]

$$\rho(e)\rho(f)^\ell v_0 = -\ell(N+1-\ell)\rho(f)^{\ell-1} v_0 \tag{15.478}$$

b.) Suppose that we choose a highest weight vector $v_h$ so that $\rho(f)w = 0$ and $\rho(h)w = +Nw$. Write the representation matrices in the ordered basis

$$w, \rho(e)w, \dots, \rho(e)^N w \tag{15.479}$$

c.) Put a unitary structure on the vector space $V$ so that $\rho(T^j)$ are anti-Hermitian matrices and relate the above bases to the standard basis $|j,m\rangle$, $m = -N/2, -N/2+1, \dots, N/2-1, N/2$ appearing in quantum-mechanics textbooks. That is, find a rescaling of the vectors $\rho(f)^k v_0$ so that in the new basis, after defining $\rho(T^i)$ using (11.819) one obtains anti-Hermitian matrices for $\rho(T^i)$.

---

**Exercise** *Quantum Implementation Of Symplectic Transformations*

Define a vector of operators

$$\hat{V}_\alpha := \begin{pmatrix} \hat{p} \\ \hat{q} \end{pmatrix} \tag{15.480}$$

with $\alpha = 1,2$.

Show that

$$\begin{aligned} [\hat{e}, \hat{V}_\alpha] &= e_{\beta\alpha}\hat{V}_\beta \\ [\hat{h}, \hat{V}_\alpha] &= h_{\beta\alpha}\hat{V}_\beta \\ [\hat{f}, \hat{V}_\alpha] &= f_{\beta\alpha}\hat{V}_\beta \end{aligned} \tag{15.481}$$

[270] *Hint*: Use $[\rho(e), \rho(f)^\ell] = \sum_{i=0}^{\ell-1} \rho(e)^i [\rho(e), \rho(f)] \rho(e)^{\ell-1-i}$.

Conclude that if $A \in SL(2,\mathbb{R})$ has the form

$$A = \exp[\alpha e + \beta h + \gamma f] \tag{15.482}$$

then

$$\hat{A} := \exp[\alpha \hat{e} + \beta \hat{h} + \gamma \hat{f}] \tag{15.483}$$

satisfies

$$[\hat{A}, \hat{V}_\alpha] = A_{\beta\alpha} \hat{V}_\beta \tag{15.484}$$

---

---

**Exercise** *Explicit Bogoliubov Transformation Of The Vacuum*

Use the previous exercise to prove the following: Let $\zeta = re^{\mathrm{i}\phi}$ and define the unitary operator:

$$S(\zeta) := \exp[\frac{1}{2}\zeta^* a^2 - \frac{1}{2}\zeta (a^\dagger)^2] \tag{15.485}$$

a.) Show that

$$S(\zeta)^\dagger a S(\alpha) = (\cosh r) a - e^{-\mathrm{i}\phi}(\sinh r) a^\dagger \tag{15.486}$$

b.) If we have a Bogoliubov transformation in $SU(1,1)$ with

$$b = (\cosh r) a - e^{-\mathrm{i}\phi}(\sinh r) a^\dagger \tag{15.487}$$

Conclude that

$$|0\rangle_b = \mathcal{N} S(\zeta)^\dagger |0\rangle_a \tag{15.488}$$

for some $\mathcal{N}$, where $|0\rangle_b$ is a normalized vacuum define by $b|0\rangle_b = 0$.

c.) Show that $|\mathcal{N}|^2 = 1$. (This is easy). Since the condition $b|0\rangle_b = 0$ only defines the normalized vacuum up to phase we cannot fix the phase $\mathcal{N}$ given just the above information.

d.) Using the BCH formula show that

$$S(\zeta) = (\cosh r)^{-1/2} e^{-\frac{1}{2}\Gamma (a^\dagger)^2} e^{-\log(\cosh r) a^\dagger a} e^{\frac{1}{2}\Gamma a^2} \tag{15.489}$$

where $\Gamma = e^{\mathrm{i}\phi}\tanh(r)$. (This is hard.) Conclude that

$$|0\rangle_b = \mathcal{N}(\cosh r)^{-1/2} e^{-\frac{1}{2}\Gamma (a^\dagger)^2} |0\rangle_a \tag{15.490}$$

(this is easy).

---

Example: $SL(2,\mathbb{Z})$ action on $Heis(\mathbb{Z}_n \times \mathbb{Z}_n)$.

Consider again the example of the quantum mechanics of a particle on a discrete approximation to a ring.

1. Because the position is periodic, the momentum is quantized.
2. Because the position is quantizied, the momentum is periodic.

So, the momentum is both periodic and discrete, just like the position. Recall the position operator $Q$ and the momentum operator $P$ both had a spectrum which is given by $n^{th}$ roots of unity.

So there is a symmetry between momentum and position. This is part of a kind of symplectic symmetry in this discrete system related to $SL(2,\mathbb{Z}_n)$. All such matrices arise from reduction modulo $n$ of matrices in $SL(2,\mathbb{Z})$. Recall from Section [**** 8.3 ****] that $SL(2,\mathbb{Z})$ is generated by $S$ and $T$ with relations

$$(ST)^3 = S^2 = -1 \tag{15.491}$$

Therefore, $S$ and $T$ (reduced modulo $n$) will generate $SL(2,\mathbb{Z}_n)$, although there will be further relations, such as $T^n = 1$. The $SL(2,\mathbb{Z}_n)$ symmetry plays and important role in string theory and Chern-Simons theory and illustrates nicely some ideas of duality.

We now take the cocycle for $\text{Heis}(\mathbb{Z}_n \times \mathbb{Z}_n)$ to be

$$f((a_1,b_1),(a_2,b_2)) = a_1 b_2 \in \mathbb{Z}_n \tag{15.492}$$

so the commutator function is

$$\kappa((a_1,b_1),(a_2,b_2)) = a_1 b_2 - a_2 b_1 = v_1^{tr} J v_2 \tag{15.493}$$

which we can recognize as a symplectic form on $\mathbb{Z}_n \oplus \mathbb{Z}_n$.

Now an important subtlety arises. In general we cannot find an equivalent cocycle so that

$$\tilde{f}((a_1,b_1),(a_2,b_2)) = \frac{1}{2}(a_1 b_2 - a_2 b_1) \tag{15.494}$$

because, in general, we are not allowed to divide by 2 in $\mathbb{Z}_n$. After all, $x = \frac{1}{2}(a_1 b_2 - a_2 b_1)$ should be the solution to $2x = (a_1 b_2 - a_2 b_1)$, but, if $n$ is even, then if $x$ is a solution $x + n/2$ is a different solution, so the expression is ambiguous. If $n$ is <u>odd</u> then 2 is invertible and we can divide by 2. Therefore, it is not obvious if $Sp(2,\mathbb{Z}_n)$ will lift to automorphisms of the finite Heisenberg group.

Consider the transformation:

$$S : (a,b) \rightarrow (b,-a) \tag{15.495}$$

This is a symplectic transformation: $S^*\kappa = \kappa$, and it satisfies $S^2 = -1$. Now we compute

$$(S^* f - f)((a_1,b_1),(a_2,b_2)) = -(a_1 b_2 + b_1 a_2) \tag{15.496}$$

Although the cocycle is not invariant under $S$ nevertheless the difference $S^* f - f$ can indeed be trivialized by

$$\tau_S(a,b) = -ab \tag{15.497}$$

Now consider the transformation

$$T : (a, b) \to (a + b, b) \tag{15.498}$$

We compute

$$(T^* f - f)((a_1, b_1), (a_2, b_2)) = b_1 b_2 \tag{15.499}$$

This can be trivialized by

$$\tau_T(a, b) = \frac{1}{2} b^2 \tag{15.500}$$

*PROVIDED* we are able to divide by 2!! This is possible if $n$ is odd, but not when $n$ is even. Indeed, when $n = 2$ the cocycle $(T^* f - f)$ is not even a trivializable cocycle! (Why not? Apply the triviality test described in [ **** Remark 5, section 11.3 ****] above.)

One way to determine the lifted group, and how the group lifts when $n$ is even is the following. Suppose we have any two operators $U, V$ that satisfy

$$UV = e^{2\pi i\theta} VU \tag{15.501}$$

for some $\theta$ (which, at this point, need not even be rational). If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}) \tag{15.502}$$

then

$$\widetilde{U} := U^a V^c \qquad \widetilde{V} = U^b V^d \tag{15.503}$$

♣Need to add phases here to get a proper group action. These don't compose properly yet. ♣

also satisfy

$$\widetilde{U}\widetilde{V} = e^{2\pi i\theta} \widetilde{V}\widetilde{U} \tag{15.504}$$

Now suppose, in addition that $\theta = 1/n$ and $U^n = V^n = 1$. Then we can compute

$$\begin{aligned} \widetilde{U}^n &= e^{-i\pi\theta n(n-1)ac} = e^{-i\pi(n-1)ac} \\ \widetilde{V}^n &= e^{-i\pi\theta n(n-1)bd} = e^{-i\pi(n-1)bc} \end{aligned} \tag{15.505}$$

Now, when $n$ is odd, the conditions (15.505) place no restriction on $A$. In that case, the group $SL(2, \mathbb{Z})$ acts on $\mathrm{Heis}(\mathbb{Z}_n \oplus \mathbb{Z}_n)$ as a group of automorphisms, but the normal subgroup

$$\Gamma(n) := \{A \in SL(2, \mathbb{Z}) | A = \mathbf{1} \mathrm{mod} n\} \tag{15.506}$$

acts trivially so that

$$SL(2, \mathbb{Z})/\Gamma(n) \cong SL(2, \mathbb{Z}_n) \tag{15.507}$$

indeed acts as a group of automorphisms.

However, when $n$ is even, we must consider the subgroup of $SL(2, \mathbb{Z})$ with the extra conditions $ac = bd = 0 \mathrm{mod} 2$. Only this subgroup acts as a group of automorphisms. Again a finite quotient group acts effectively.

Lifting Symplectomorphisms In Geometric Quantization

************************************************

************************************************

TO BE WRITTEN OUT IN DETAIL: A good example of the general problem of lifting automorphisms is to consider a line bundle with connection $(L, \nabla)$ over a symplectic manifold such that the curvature of the connection is the symplectic form $\Omega$. Then, in geometric quantization we can attempt to lift the symplectomorphisms which preserve the line bundle with connection. We will see all the above phenomena, and only the "Hamiltonian automorphisms" will lift. For some discussion see section 6 of:

https://arxiv.org/pdf/hep-th/0605200.pdf

IN PARTICULAR EXPLAIN THE RELATION OF $\mathrm{Aut}_0(G)$ to the kernel of a homomorphism $\mathrm{Aut}(G) \to \mathrm{Hom}(G, H^1(G, A))$.

****************************************

****************************************

#### 15.5.9 Coherent State Representations Of Heisenberg Groups: The Bargmann Representation

EXPLAIN BARGMANN REPRESENTATION.

NICE FORMULAE FOR COHERENT STATES AND VEV'S OF EXPONENTIATED QUADRATIC EXPRESSIONS IN OSCILLATORS.

#### 15.5.10 Some Remarks On Chern-Simons Theory

In Chern-Simons theory (and similar topological field theories) it is quite typical for the Wilson line operators to generate finite Heisenberg groups. For example for $U(1)$ Chern-Simons of level $k$ on a torus we have an action

$$S = \frac{k}{4\pi} \int_{\mathbb{R}} dt \int_{T^2} A_1 \partial_t A_2 + \cdots \tag{15.508}$$

so upon quantization $A_2 \sim \frac{4\pi}{k} \frac{\delta}{\delta A_1}$. The consequence is that Wilson lines along the $a$- and $b$-cycles generate a finite Heisenberg group with $q = e^{2\pi \mathrm{i}/k}$. The Hilbert space of states is a finite-dimensional irreducible representation of this group.

************************

EXPLAIN MORE. THETA FUNCTIONS AND METAPLECTIC REPRESENTATION.

***************

### 15.6 Non-Central Extensions Of A General Group $G$ By An Abelian Group $A$: Twisted Cohomology

Let us now generalize central extensions to extensions of the form:

$$1 \to A \overset{\iota}{\to} \tilde{G} \overset{\pi}{\to} G \to 1 \tag{15.509}$$

Here $G$ can be any group, not necessarily Abelian. We continue to assume that $N = A$ is Abelian, but now we no longer assume $\iota(A)$ is central in $\tilde{G}$. So we allow for the possibility of non-central extensions by an Abelian group.

Much of our original story goes through, but now the map

$$\omega : G \to \mathrm{Aut}(A) \tag{15.510}$$

of our general discussion (defined in equations (15.12) and (15.14)) is canonically defined and is actually a group homomorphism. As we stressed below (15.14), in general it is not a group homomorphism. There are two ways to understand that:

1. $\tilde{G}$ acts on $A$ by conjugation on the isomorphic image of $A$ in $\tilde{G}$ which, because the sequence is exact, is still a normal subgroup. In equations, we can define

   $$\iota(\tilde{\omega}_{\tilde{g}}(a)) := \tilde{g}\iota(a)\tilde{g}^{-1} \tag{15.511}$$

   But now $\tilde{\omega}_{\tilde{g}}$ only depends on the equivalence class $[\tilde{g}] \in \tilde{G}/\iota(A)$ beccause

   $$(\tilde{g}\iota(a_0))\,\iota(a)\,(\tilde{g}\iota(a_0))^{-1} = \tilde{g}\iota(a)\tilde{g}^{-1} \tag{15.512}$$

   so $\tilde{\omega}_{\tilde{g}\iota(a_0)} = \tilde{\omega}_{\tilde{g}}$ and since $\tilde{G}/\iota(A) \cong G$ we can use this to define $\omega_g$. However, from this definition it is clear that $g \mapsto \omega_g$ is a group homomorphism.

2. Or you can just choose a section and define $\omega_g$ exactly as in (15.14). To stress the dependence on $s$ we write

   $$\iota(\omega_{g,s}(a)) = s(g)\iota(a)s(g)^{-1} \tag{15.513}$$

   However, now if we change section so that [271] $\hat{s}(g) = \iota(t(g))s(g)$ is another section then we compute

   $$\begin{aligned}\iota\left(\omega_{g,\hat{s}}(a)\right) &:= \{\iota(t(g))s(g)\} \cdot \iota(a) \cdot \{\iota(t(g))s(g)\}^{-1} \\ &= \iota(t(g)) \cdot \iota(\omega_{g,s}(a)) \cdot \iota(t(g))^{-1} \\ &= \iota\left\{t(g) \cdot \omega_{g,s}(a) \cdot (t(g))^{-1}\right\} \\ &= \iota(\omega_{g,s}(a))\end{aligned} \tag{15.514}$$

   and since $\iota$ is injective $\omega_{g,s}$ is independent of section and we can just denote it as $\omega_g$.

   Note carefully that only in the very last line did we use the assumption that $A$ is Abelian. We will come back to this when we discuss general extensions in section 15.7.

   Moreover, given a choice of section we can define $f_s(g_1, g_2)$ just as we did in equation (15.58). This definition works for all group extensions:

   $$s(g_1)s(g_2) = \iota(f_s(g_1, g_2))s(g_1g_2) \tag{15.515}$$

[271] Note that here the order of the two factors on the RHS matters, since $\iota(A)$ is not necessarily central in $\widetilde{G}$

We can now compute, just as in (15.15):

$$\begin{aligned}\iota\left(\omega_{g_1}\circ\omega_{g_2}(a)\right) &= s(g_1)\iota(\omega_{g_2}(a))s(g_1)^{-1}\\ &= s(g_1)s(g_2)\iota(a)(s(g_1)s(g_2))^{-1}\\ &= \iota(f_s(g_1,g_2))\cdot\iota(\omega_{g_1g_2}(a))\cdot\iota(f_s(g_1,g_2))^{-1}\\ &= \iota\left\{f_s(g_1,g_2))\omega_{g_1g_2}(a)f_s(g_1,g_2))^{-1}\right\}\\ &= \iota\left(\omega_{g_1g_2}(a)\right)\end{aligned} \tag{15.516}$$

and again notice that only in the very last line did we use the hypothesis that $A$ is Abelian. Again, since $\iota$ is injective, we conclude that $\omega_{g_1}\circ\omega_{g_2}=\omega_{g_1g_2}$ so that the map $\omega$ is a group homomorphism.

Now, computing $s(g_1)s(g_2)s(g_3)$ in two ways, just as before, we derive the *twisted cocycle relation*:

$$\omega_{g_1}(f_s(g_2,g_3))f_s(g_1,g_2g_3)=f_s(g_1,g_2)f_s(g_1g_2,g_3) \tag{15.517}$$

Conversely, given a homomorphism $\omega: G\to \mathrm{Aut}(A)$ and a twisted cocycle for $\omega$ we can define a group law on the set $A\times G$:

$$(a_1,g_1)\cdot(a_2,g_2)=(a_1\omega_{g_1}(a_2)f(g_1,g_2),g_1g_2) \tag{15.518}$$

The reader should check that this really does define a valid group law on the set $A\times G$.

**Remark**: Note that (15.518) simultaneously generalizes the twisted product of a semidirect product (14.2) and the twisted product of a central extension (15.70).

Now suppose that we change section from $s$ to $\hat{s}(g):=\iota(t(g))s(g)$ using some arbitrary function $t: G\to A$. Then one can compute that the new cocycle is related to the old one by

$$f_{\hat{s}}(g_1,g_2)=t(g_1)\omega_{g_1}(t(g_2))f_s(g_1,g_2)t(g_1g_2)^{-1} \tag{15.519}$$

Note that since $A$ is Abelian the order of the factors on the RHS do not matter, but in the analogous formula for general extensions, equation (15.599) below, the order definitely does matter.

We say two different twisted cocycles are related by a twisted coboundary if they are related as in (15.519) for some function $t: G\to A$. One can check that if $f$ is a twisted cocycle and we define $f'$ as in (15.519) then $f'$ is also a twisted cocycle. We again have an equivalence relation and we define the *twisted cohomology group* $H^{2+\omega}(G,A)$ to be the abelian group of equivalence classes. It is again an Abelian group, as in the untwisted case, as one shows by a similar argument.

The analog of the main theorem of section 15.3 above is:

**Theorem**: Let $\omega: G\to \mathrm{Aut}(A)$ be a fixed group homomorphism. Denote the set of isomorphism classes of extensions of the form

$$1\to A\to\tilde{G}\to G\to 1 \tag{15.520}$$

which induce $\omega$ by $\mathrm{Ext}^{\omega}(G,A)$. Then the set $\mathrm{Ext}^{\omega}(G,A)$ is in 1-1 correspondence with the twisted cohomology group $H^{2+\omega}(G,A)$.

The proof is very similar to the untwisted case and we will skip it. Now the trivial element of the Abelian group $H^{2+\omega}(G,A)$ corresponds to the semidirect product determined by $\omega$.

Now we can observe an interesting phenomenon which happens often in cohomology theory: Suppose that a twisted cocycle $f$ is trivializable so that $[f]=0$. Then our group extension is equivalent to a semidirect product. Nevertheless, the sequence (15.509) can be split in many different ways: There are many distinct trivializations and the different trivializations have meaning. Equivalently, there are many different coboundary transformations that preserve the trivial cocycle. A glance at (15.519) reveals that this will happen when

$$t(g_1g_2)=t(g_1)\omega_{g_1}(t(g_2)) \tag{15.521}$$

This is known as a *twisted homomorphism*. Of course, in the case that $\omega: G\to \mathrm{Aut}(A)$ takes every $g\in G$ to the identity automorphism of $\mathrm{Aut}(A)$ (that is, the identity element of $\mathrm{Aut}(A)$, the condition specializes to the definition of a homomorphism.

For the later discussion of group cohomology is useful:

A 1-cochain $t\in C^1(G,A)$ is simply a map $t: G\to A$.

A twisted homomorphism is also known as a *twisted one-cocycle*. That is, a 1-cocycle $t\in Z^{1+\omega}(G,A)$ with twisting $\omega$ is a 1-cochain that satisfies (15.521).

To define group cohomology $H^{1+\omega}(G,A)$ we need an appropriate notion of equivalence of one-cocycles. This is motivated by noting that if $s: G\to\tilde{G}$ is a section that is also a homomorphism (that is, a splitting) then for any $a\in A$ we can produce a new splitting

$$\tilde{s}(g)=\iota(a)s(g)\iota(a)^{-1} \tag{15.522}$$

This corresponds to the change of section $\tilde{s}(g)=\iota(t(g))s(g)$ where the function $t(g)$ is:

$$t(g)=t_a(g):=a\omega_g(a)^{-1} \quad . \tag{15.523}$$

To check this you write

$$\begin{aligned}
\tilde{s}(g)&=\iota(a)s(g)\iota(a)^{-1}\\
&=\iota(a)\cdot\left(s(g)\iota(a)^{-1}s(g)^{-1}\right)\cdot s(g)\\
&=\iota(a)\cdot\left(\iota(\omega_g(a^{-1}))\right)\cdot s(g)\\
&=\left(\iota(a\omega_g(a^{-1}))\right)\cdot s(g)
\end{aligned} \tag{15.524}$$

One easily checks that if $t$ is a one-cocycle, then $t\cdot t_a$ is also a one-cocycle. So, in defining the cohomology group $H^{1+\omega}(G,A)$ we use the equivalence relation $t\sim t'$ if there exists an $a$ so that $t=t't_a$.

**Theorem:** When the sequence (15.509) splits, that is, when the cohomology class of the twisted cocycle is trivial $[f]=0$, then the inequivalent splittings are in one-one correspondence with the inequivalent trivializations of a trivializable cocycle, and these are in one-one correspondence with the cohomology group $H^{1+\omega}(G,A)$.

**Example 1**: Consider the sequence associated with the Euclidean group

$$0 \to \mathbb{R}^d \stackrel{\iota}{\to} \mathrm{Euc}(d) \stackrel{\pi}{\to} O(d) \to 1 \tag{15.525}$$

Recall that if $v \in \mathbb{R}^d$ then $\iota(v) = T_v$ is the translation operator on affine space $\mathbb{A}^d$. We have $T_v(p) = p + v$. As we saw in (14.74) and (14.75) and the discussion preceding that exercise, for any $p \in \mathbb{A}^d$ we have a section $R \mapsto s_p(R) \in \mathrm{Euc}(d)$ where $s_p(R)$ is the transformation that takes

$$s_p(R) : p + v \mapsto p + Rv \tag{15.526}$$

In other words, we define rotation-reflections by choosing $p$ as the origin. Then from

$$T_{\omega_R(v_0)} = s_p(R) T_{v_0} s_p(R)^{-1} \tag{15.527}$$

we compute that

$$\omega_R(v_0) = R v_0 \tag{15.528}$$

thus $\omega_R \in \mathrm{Aut}(\mathbb{R}^d)$, and indeed $R \mapsto \omega_R$ is a group homomorphism. If we have two difference sections $s_{p'}$ and $s_p$ then

$$s_{p'}(R) = T_{t(R)} s_p(R) \tag{15.529}$$

where

$$t(R) = (1 - R)(p' - p) = (1 - R)w \tag{15.530}$$

where we have put $p' = p + w$, $w \in \mathbb{R}^d$.

Note that One easily checks that for fixed $w \in \mathbb{R}^d$

$$R \mapsto t(R) := (1 - R)w \tag{15.531}$$

is indeed a twisted homomorphism $O(d) \to \mathbb{R}^d$. (It is not a homomorphism.) However, one also checks that it is of the form $t_w(R) = w - \omega_R(w)$, so it is a trivial one-cocycle: All the splittings are equivalent in the sense defined above.

**Example 2**: Now restrict the sequence (15.525) to

$$0 \to \mathbb{Z}^d \to G \to \{1, \sigma\} \to 1 \tag{15.532}$$

where $\{1, \sigma\} \subset O(d)$ is a subgroup isomorphic to $\mathbb{Z}_2$ with $\sigma = -\mathbf{1}_{d \times d}$. Then $\omega : \mathbb{Z}_2 \to \mathrm{Aut}(\mathbb{Z}^d)$ is inherited from $\omega_R$ in (15.525) and $\omega_\sigma(\vec{n}) = -\vec{n}$. Now the sequence splits and the most general possible splitting is $s_{\vec{n}_0}$ where $s_{\vec{n}_0}(1) = \{0|\mathbf{1}\}$ and

$$s_{\vec{n}_0}(\sigma) = \{\vec{n}_0|\sigma\} \tag{15.533}$$

for some $\vec{n}_0 \in \mathbb{Z}^d$. Indeed one checks $s_{\vec{n}_0}(\sigma)^2 = 1$. Now for $\vec{a} \in \mathbb{R}^d$ we have

$$t_{\vec{a}}(\sigma) = \vec{a} - \omega_\sigma(\vec{a}) = 2\vec{a} \in 2\mathbb{Z}^d \tag{15.534}$$

So not all splittings are equivalent! The equivalent ones have $\vec{n}_0 - \vec{n}_0' \in 2\mathbb{Z}^d$. Therefore

$$H^{1+\omega}(\mathbb{Z}_2, \mathbb{Z}^d) \cong \mathbb{Z}^d / 2\mathbb{Z}^d \cong (\mathbb{Z}_2)^d \tag{15.535}$$

**Remarks**

1. Different trivializations of something trivializable can have physical meaning. In the discussion on crystallographic groups below the different trivializations are related to a choice of origin for rotation-reflection symmetries of the crystal.

2. An analogy to bundle theory might help some readers: Let $G$ be a compact Lie group. Then the isomorphism classes of principal $G$-bundles over $S^3$ are in 1-1 correspondence with $\pi_2(G)$ and a theorem states that $\pi_2(G) = 0$ for all compact Lie groups. Therefore, every principal $G$-bundle over $S^3$ is trivializable. Distinct trivializations differ by maps $t : S^3 \to G$ and the set of inequivalent trivializations is classified by $\pi_3(G)$, which is, in general nontrivial. This can have physical meaning. For example, in Yang-Mills theory in $3+1$ dimensions on $S^3 \times \mathbb{R}$ the principal $G$-bundle on space $S^3$ is trivializable. But if there is an instanton between two time slices then the trivialization jumps by an element of $\pi_3(G)$.

---

**Exercise** *Due diligence*

Derive equation (15.517) and show that if we change $f$ by a coboundary using (15.519) then indeed we produce another twisted cocycle.

---

---

**Exercise**

Suppose that a twisted cocycle $f(g_1, g_2)$ can be trivialized by two different functions $t_1, t_2 : G \to A$. Show that $t_{12}(g) := t_1(g)/t_2(g)$ is a trivialization that preserves the trivial cocycle. That is, show that $t_{12}$ is a twisted 1-cocycle.

---

### 15.6.1 Crystallographic Groups

A *crystal* is a subset of affine space $C \subset \mathbb{A}^d$ that is invariant under translations by a lattice $L \subset \mathbb{R}^d$ (actually, that's an embedded lattice). As an example, see Figure 43. Then restricting the exact sequence of the Euclidean group (equation (15.525) above ) to the subgroup $G(C) \subset \mathrm{Euc}(d)$ of those transformations that preserve $C$ we have an exact sequence

$$1 \to L(C) \to G(C) \to P(C) \to 1 \tag{15.536}$$

where $P(C) \cong G(C)/L(C)$ is a subgroup of $O(d)$ known as the *point group of the crystal.*

**Remark**: In solid state physics when the sequence (15.536) does not split the crystallographic group $G(C)$ is said to be *nonsymmorphic.*

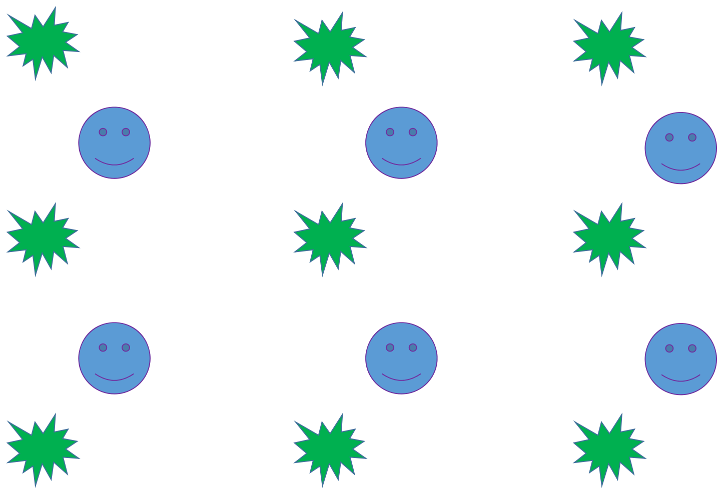


**Figure 43:** A portion of a crystal in the two-dimensional plane.

**Example 1**: Take $C = \mathbb{Z} \amalg (\mathbb{Z} + \delta) \subset \mathbb{R}$ where $0 < \delta < 1$. Then of course $L(C) = \mathbb{Z}$ acts by translations, preserving the crystal. But note that it is also true that

$$\begin{aligned} \{\delta|\sigma\} &: n \mapsto \delta - n = -n + \delta \\ &: n + \delta \mapsto \delta - (n + \delta) = -n \end{aligned} \tag{15.537}$$

where $\sigma \in O(1)$ is the reflection around 0, $\sigma : x \to -x$ in $\mathbb{R}$. The transformation $\{\delta|\sigma\}$ maps $\mathbb{Z}$ to $\mathbb{Z} + \delta$ and $\mathbb{Z} + \delta$ to $\mathbb{Z}$ so that the whole crystal is preserved. Since $O(1) = \mathbb{Z}_2$, this is all we can do. We thus find that $G(C)$ fits in a sequence

$$0 \to L(C) \cong \mathbb{Z} \to G(C) \to O(1) \cong \mathbb{Z}_2 \to 1 \tag{15.538}$$

But we can split this sequence by choosing a section $s(\sigma) = \{\delta| - 1\}$. Note that

$$\{\delta|\sigma\} \cdot \{\delta|\sigma\} = \{0|1\} \tag{15.539}$$

so $s : O(1) \to G(C)$ is a homomorphism. Another way of thinking about this is that $s(\sigma)$ is just reflection, not around the origin, but around the point $\frac{1}{2}\delta$. So, by a shift of origin for defining our rotation-inversion group $O(1)$ we just have reflections and integer translations. In any case we can recognize $G(C)$ as the infinite dihedral group.

**Example 2**: More generally, consider a lattice $L \subset \mathbb{R}^d$ and a generic vector $\vec{\delta} \in \mathbb{R}^d$. Consider the crystal

$$C = L \amalg (L + \vec{\delta}) \tag{15.540}$$

If $L$ and $\vec{\delta}$ are generic then the point group is just $\mathbb{Z}_2$ generated by $-1 \in O(d)$. Denoting the action of $-1$ on $\mathbb{R}^d$ by $\sigma$ we can lift this to the involution

$$\{\vec{\delta}|\sigma\} \in G(C) \tag{15.541}$$

which exchanges $L$ with $(L+\vec{\delta})$. This group is symmorphic (because $\{\vec{\delta}|\sigma\}$ is an involution). In fact, this operation is just inversion about the new origin $\frac{1}{2}\vec{\delta}$:

$$\{\vec{\delta}|\sigma\} : \frac{1}{2}\vec{\delta} + \vec{y} \to \frac{1}{2}\vec{\delta} - \vec{y} \tag{15.542}$$

♣NEED TO HAVE A FIGURE HERE. THIS WOULD HELP. ♣

**Example 3**: For another very similar example consider

$$C = L \amalg (L + \vec{\delta}) \subset \mathbb{R}^2 \tag{15.543}$$

where

$$L = a_1\mathbb{Z} \oplus a_2\mathbb{Z} \subset \mathbb{R}^2 \tag{15.544}$$

As we have just discussed, for generic $a_1, a_2$ and $\vec{\delta}$ the symmetry group will be isomorphic to the semidirect product $\mathbb{Z}^2 \rtimes \mathbb{Z}_2$.

However, now let $0 < \delta < \frac{1}{2}$ and specialize $\vec{\delta}$ to $\vec{\delta} = (\delta a_1, \frac{1}{2}a_2)$. Then the crystal has more symmetry and in particular the point group is enhanced from $\mathbb{Z}_2$ to $\mathbb{Z}_2 \times \mathbb{Z}_2$:

$$1 \to \mathbb{Z}^2 \to G(C) \to \mathbb{Z}_2 \times \mathbb{Z}_2 \to 1 \tag{15.545}$$

To see this let $\sigma_1, \sigma_2$ be generators of $\mathbb{Z}_2 \times \mathbb{Z}_2$ acting by reflection around the $x_2$ and $x_1$ axes, respectively. Then the operations:

$$\hat{\sigma}_1 : (x_1, x_2) \mapsto \vec{\delta} + (-x_1, x_2) \tag{15.546}$$

$$\hat{\sigma}_2 : (x_1, x_2) \mapsto (x_1, -x_2) \tag{15.547}$$

are symmetries of the crystal $G(C)$. In Seitz notation (or rather, its improvement - see equations (14.31) and (14.32) above) we have:

$$\hat{\sigma}_1 = \{\vec{\delta}| \begin{pmatrix} -1 & \\ & 1 \end{pmatrix}\} \tag{15.548}$$

$$\hat{\sigma}_2 = \{0| \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}\} \tag{15.549}$$

Now, we can define a section $s(\sigma_1) = \hat{\sigma}_1$ and $s(\sigma_2) = \hat{\sigma}_2$. Note that the square of the lift

$$\hat{\sigma}_1^2 = \{(0, a_2)|1\} \tag{15.550}$$

is a nontrivial translation. Thus $\sigma_i \to \hat{\sigma}_i$ is *not* a splitting. Moreover, $\hat{\sigma}_1$ does not have finite order. Therefore, it cannot be in a discrete group of rotations about any point!

Just because we chose a section that wasn't a splitting doesn't mean that a splitting doesn't actually exist. Here is how we can prove that in fact no splitting exists: The most general section is of the form

$$s(\sigma_1) = \{\vec{\delta} + \vec{v}|\sigma_1\} \tag{15.551}$$

where $\vec{v} = (n_1 a_1, n_2 a_2) \in L$ where $n_1, n_2 \in \mathbb{Z}$. Now consider the square:

$$s(\sigma_1)^2 = \{(0, a_2(1 + 2n_2)|1\}. \tag{15.552}$$

Since $n_2 \in \mathbb{Z}$ there is no lifting that makes this an involution. Therefore, there is no section. Therefore the sequence (15.545) does not split.

**Example 4**: It is interesting to see what happens to the previous example when $a_1 = a_2 = a$ and we take $\delta = a(\frac{1}{2}, \frac{1}{2})$. Then, clearly

$$C = L \amalg (L + \vec{\delta}) \subset \mathbb{R}^2 \tag{15.553}$$

has a point group symmetry $D_4$. So this becomes a symmorphic crystal. In fact, this is just a square lattice in disguise! We can take basis vectors $\delta$ and $R(\pi/2)\delta$.

***************************

NEED TO RELATE THE ABOVE FACTS MORE DIRECTLY TO THE PREVIOUS DISCUSSION OF GROUP COHOMOLOGY. SHOULD DO MORE ON CASE WHERE THE SEQUENCE SPLITS BUT THERE ARE INEQUIVALENT SPLITTINGS: PROBABLY A good example is Zincblend structure with tetrahedral symmetry. For example GaAs has this structure. There are two tetrahedra around the Ga and As but they are rotated.

*****************************

---

**Exercise**

Why does the argument of example 3 fail in the special case of example 4? [272]

---

---

**Exercise** *Honeycomb*

Consider a honeycomb crystal in the plane. Discuss the crystal group, the point group, and decide if it is symmorphic or not.

---

♣Need to provide answer in a footnote ♣

[272] *Answer*: The wrong step is in equation (15.551). When $\delta$ takes the special form $\frac{1}{2}(a, a)$ this is not the most general lifting. One has now translation symmetry by multiples of $\vec{\delta}$, so there is an obvious lifting of $\sigma_1$.

### 15.6.2 Time Reversal

A good example of a physical situation in which it is useful to know about how twisted cocycles define non-central extensions is when there are anti-unitary symmetries in a quantum mechanical system. A typical example where this happens is when there is a time-orientation-reversing symmetry. In this case there is a homomorphism

$$\tau : G \to \{\pm 1\} \cong \mathbb{Z}_2 \tag{15.554}$$

telling us whether the symmetry $g \in G$ preserves or reverses the orientation of time.

In quantum mechanics it is often (but not always! - see below) the case that time-reversal is implemented as an anti-unitary operator (see Chapter 2 below for a precise definition of this term) and therefore when looking at the way the symmetry is implemented quantum mechanically we should consider the nontrivial automorphism of $U(1)$ defined by complex conjugation.

Recall that

$$\mathrm{Aut}(U(1)) \cong \mathrm{Out}(U(1)) \cong \mathbb{Z}_2 \tag{15.555}$$

and the nontrivial element of $\mathrm{Aut}(U(1)$ is the automorphism $z \to z^* = z^{-1}$.

So, when working with a symmetry group $G$ that includes time-orientation-reversing symmetries we will need to consider the group homomorphism

$$\omega : G \to \mathrm{Aut}(U(1)) \tag{15.556}$$

where:

$$\omega(g)(z) = \begin{cases} z & \tau(g) = +1 \\ z^{-1} & \tau(g) = -1 \end{cases} \tag{15.557}$$

**Example 1**. The simplest example is where we have a symmetry group $G = \mathbb{Z}_2$ interpreted as time reversal. It will be convenient to denote $M_2 = \{1, \bar{T}\}$, with $\bar{T}^2 = 1$. Of course, $M_2 \cong \mathbb{Z}_2$. In quantum mechanics the representation of $\bar{T}$ will be an operator $\rho(T) := \tilde{T}$ on the Hilbert space and we will get a possibly twisted central extension of $M_2$. Let $\omega : M_2 \to \mathrm{Aut}(U(1))$. There are two possibilities: $\omega(\bar{T}) = 1$ (so the operation is unitary) and $\omega(\bar{T})$ is the complex conjugation automorphism (so the operation is anti-unitary). Assuming the anti-unitary case is the relevant one, so that $\omega$ is the nontrivial homomorphis $M_2 \to \mathrm{Aut}(U(1))$ (both are isomorphic to $\mathbb{Z}_2$ so $\omega$ is just the identity homomorphism of $\mathbb{Z}_2$) we need the group cohomology:

$$H^{2+\omega}(\mathbb{Z}_2, U(1)) = \mathbb{Z}_2 \tag{15.558}$$

To prove this we look at the twisted cocycle identity. Exactly the same arguments as in Remark 5 of section 15.3 show that we can choose a gauge with $f(g,1) = f(1,g) = 1$ for all $g$. This leaves only $f(\bar{T}, \bar{T})$ to be determined. Now, the twisted cocycle relation for the case $g_1 = g_2 = g_3 = \bar{T}$ says that

$$f(\bar{T}, \bar{T})^* = \omega_{\bar{T}}(f(\bar{T}, \bar{T})) = f(\bar{T}, \bar{T}) \tag{15.559}$$

and since $f(\bar{T},\bar{T}) \in U(1)$ this means $f(\bar{T},\bar{T}) \in \{\pm 1\}$. We need to check that we can't gauge $f(\bar{T},\bar{T})$ to one using a twisted coboundary relation. That relation says that we can gauge $f$ to

$$\tilde{f}(\bar{T},\bar{T}) = t(\bar{T})\omega_{\bar{T}}(t(\bar{T}))f(\bar{T},\bar{T})/t(1) \tag{15.560}$$

Now $t(1) = 1$ since we want to preserve the gauge $f(g,1) = f(1,g) = 1$ and $t(\bar{T})\omega_{\bar{T}}(t(\bar{T})) = |t(\bar{T})|^2 = 1$ so $f(\bar{T},\bar{T})$ is gauge invariant.

(Note that $\bar{T}$ is an involution so our old criterion from Remark 5 of section 15.3 would ask us to find a square root of $f(\bar{T},\bar{T}) = -1$. Indeed such a square root exists, it is $\pm \mathrm{i}$, but our old criterion no longer applies because we are in the twisted case.)

So, choosing $\omega$ to be the nontrivial homomorphism $M_2 \to \mathrm{Aut}(U(1))$ there are two extensions:

$$1 \longrightarrow U(1) \longrightarrow M_2^{\pm} \xrightarrow{\tilde{\pi}} M_2 \longrightarrow 1 \tag{15.561}$$

Let us write these out more explicitly:

Choose a lift $\tilde{T}$ of $\bar{T}$. Then $\pi(\tilde{T}^2) = 1$, so $\tilde{T}^2 = z \in U(1)$. But, then

$$\tilde{T}z = \tilde{T}\tilde{T}^2 = \tilde{T}^2\tilde{T} = z\tilde{T} \tag{15.562}$$

On the other hand, since we take $\omega(\bar{T})$ to be the nontrivial automorphism of $U(1)$ then

$$\tilde{T}z = z^{-1}\tilde{T} \tag{15.563}$$

Therefore $z^2 = 1$, so $z = \pm 1$, and therefore $\tilde{T}^2 = \pm 1$. Thus the two groups are

$$M_2^{\pm} = \{z\tilde{T} | z\tilde{T} = \tilde{T}z^{-1} \qquad \& \qquad \tilde{T}^2 = \pm 1\} \tag{15.564}$$

These possibilities are really distinct: If $\tilde{T}'$ is another lift of $\bar{T}$ then $\tilde{T}' = \mu\tilde{T}$ for some $\mu \in U(1)$ and so

$$(\tilde{T}')^2 = (\mu\tilde{T})^2 = \mu\bar{\mu}\tilde{T}^2 = \tilde{T}^2 \tag{15.565}$$

So the sign of the square of the lift of the time-reversing symmetry is an invariant.

The extension corresponding to the identity element of $H^{2+\omega}(\mathbb{Z}_2, U(1))$ is the semidirect product. This is just $O(2)$, using $SO(2) \cong U(1)$:

$$O(2) = SO(2) \rtimes \mathbb{Z}_2 \tag{15.566}$$

But the nontrivial extension is a new group for us. It double-covers $O(2)$ and is known as $\mathrm{Pin}^-(2)$. Indeed we can define homomorphisms

$$\pi^{\pm} : M_2^{\pm} \to O(2) \tag{15.567}$$

where $\pi^{\pm}(\tilde{T}) = P \in O(2)$ and $\pi^{\pm}(z = e^{\mathrm{i}\alpha}) = R(2\alpha)$. Note that $-1 = e^{\mathrm{i}\pi} \mapsto R(e^{2\mathrm{i}\pi}) = +\mathbf{1}$. In $\mathrm{Pin}^+(2)$ the double cover of a reflection, $\tilde{T} = (\pi^+)^{-1}(P)$ , squares to one. In $\mathrm{Pin}^-(2)$ the double cover of a reflection, $\tilde{T} = (\pi^-)^{-1}(P)$ squares to $-1$.

**Remark**: In QM textbooks it is shown that if we write Schrödinger equation for an electron in a potential with spin-orbit coupling then there is a time-reversal symmetry:

$$(\tilde{T} \cdot \Psi)(\vec{x}, t) = \mathrm{i}\sigma^2(\Psi(\vec{x}, -t)^* \tag{15.568}$$

where here $\Psi$ is a 2-component spinor function of $(\vec{x}, t)$. [273] Note that this implies:

$$\begin{aligned}
(\tilde{T}^2 \cdot \Psi)(\vec{x}, t) &= \mathrm{i}\sigma^2 \cdot \Big((\tilde{T} \cdot \Psi)(\vec{x}, -t)\Big)^* \\
&= \mathrm{i}\sigma^2 \cdot \big(\mathrm{i}\sigma^2 \cdot (\Psi(\vec{x}, t))^*\big)^* \\
&= \mathrm{i}\sigma^2 \mathrm{i}\sigma^2 \Psi(\vec{x}, t) \\
&= -\Psi(\vec{x}, t)
\end{aligned} \tag{15.569}$$

So, in this example, $\tilde{T}^2 = -1$. More generally, in analogous settings for spin $j$ particles $\tilde{T}^2 = (-1)^{2j}$. See section 15.6.3 below for an explanation. The fact that $\tilde{T}^2 = (-1)^{2j}$ in the spin $j$ representation has a very important consequence known as *Kramer's theorem*: In these situations the energy eigenspaces must have even degeneracy. For if $\Psi$ is an energy eigenstate $H\Psi = E\Psi$ and we have a time-reversal invariant system then $\tilde{T} \cdot \Psi$ is also an energy eigenstate. We can prove that it is linearly independent of $\Psi$ as follows: Suppose to the contrary that

$$\tilde{T} \cdot \Psi = z\Psi \tag{15.570}$$

for some complex number $z$. Then act with $\tilde{T}$ again and use the fact that it is anti-unitary and squares to $-1$:

$$-\Psi = z^* \tilde{T} \cdot \Psi \tag{15.571}$$

but this implies that $z = -1/z^*$ which implies $|z|^2 = -1$, which is impossible. Therefore, (15.570) is impossible. Therefore $\Psi$ and $\tilde{T} \cdot \Psi$ are independent energy eigenstates. A slight generalization of the argument shows that the dimension of the energy eigenspace must be even. A more conceptual way of understanding this is that the energy eigenspace must be a quaternionic vector space because we have an anti-linear operator on it that squares to $-1$. See the discussion of real, complex, and quaternionic vector spaces in Chapter 2 below.

**Example 2**: In general a system can have time-orientation reversing symmetries but the simple transformation $t \to -t$ is not a symmetry. Rather, it must be accompanied by other transformations so that the symmetry group is <u>not</u> of the simple form $G = G_0 \times \mathbb{Z}_2$ where $G_0$ is a group of time-orientation-preserving symmetries. (Such a structure is often assumed in the literature.) As a simple example consider a crystal

$$C = \big(\mathbb{Z}^2 + (\delta_1, \delta_2)\big) \amalg \big(\mathbb{Z}^2 + (-\delta_2, \delta_1)\big) \amalg \big(\mathbb{Z}^2 + (-\delta_1, -\delta_2)\big) \amalg \big(\mathbb{Z}^2 + (\delta_2, -\delta_1)\big) \tag{15.572}$$

where $\vec{\delta}$ is generic so, as we saw above we have a symmorphic crystal with $P(C) \cong D_4$. The action of $D_4$ is just given by rotation around the origin $\{0|R(\frac{\pi}{2})\}$ which we will denote by $R$ and reflection, say, in the $y$-axis, which we will denote by $P$. So $R^4 = 1$, $P^2 = 1$, and $PRP = R^{-1}$. We have

$$G(C) = \mathbb{Z}^2 \rtimes D_4 \tag{15.573}$$

[273] This is most elegantly derived from the time-reversal transformation on the Dirac equation.

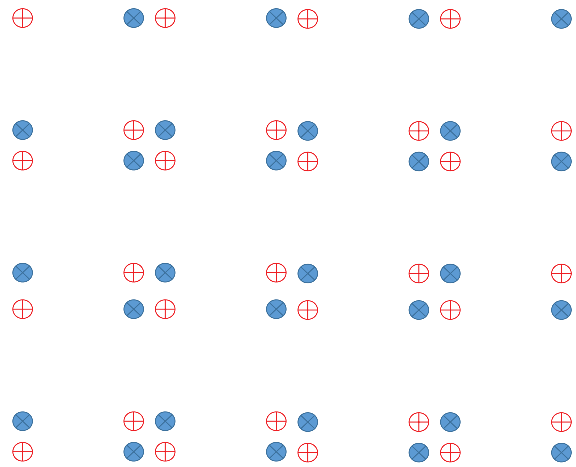


**Figure 44:** In this figure the blue crosses represent an atom with a local magnetic moment pointing up while the red crosses represent an atom with a local magnetic moment pointing down. The magnetic point group is isomorphic to $D_4$ but the homomorphism $\tau$ to $\mathbb{Z}_2$ has a kernel $\mathbb{Z}_2 \times \mathbb{Z}_2$ (generated by $\pi$ rotation around a lattice point together with a reflection in a diagonal). Since $D_4$ is nonabelian the sequence $1 \to \widehat{P}_0 \to \widehat{P} \xrightarrow{\tau} \mathbb{Z}_2 \to 1$ plainly does not split.

But now suppose there is a dipole moment, or spin $S$. We model this with a set of two elements $\mathcal{S} = \{S, -S\}$ for dipole moment up and down and now our crystal with spin is a subset of $\mathbb{R}^2 \times \mathcal{S}$. This subset is of the form

$$\widehat{C} = \widehat{C}_+ \amalg \widehat{C}_- \tag{15.574}$$

with

$$\widehat{C}_+ = \left(\mathbb{Z}^2 + (\delta_1, \delta_2)\right) \times \{S\} \amalg \left(\mathbb{Z}^2 + (-\delta_1, -\delta_2)\right) \times \{S\} \tag{15.575}$$

but a spin $-S$ on points of the complementary sub-crystal

$$\widehat{C}_- = \left(\mathbb{Z}^2 + (-\delta_2, \delta_1)\right) \times \{-S\} \amalg \left(\mathbb{Z}^2 + (\delta_2, -\delta_1)\right) \times \{-S\} \tag{15.576}$$

Now let $\mathbb{Z}_2 = \{1, \sigma\}$ act on $\mathbb{R}^2 \times \mathcal{S}$ by acting trivially on the first factor and $\sigma : S \to -S$ on the second factor. Now reversal of time orientation exchanges $S$ with $-S$. So the symmetries of the crystal with dipole is a subgroup $\widehat{G(C)} \subset \mathrm{Euc}(2) \times \mathbb{Z}_2$ known as the *magnetic crystallographic group*. The subgroup of translations by the lattice is still a normal subgroup and the quotient by the lattice of translations is the *magnetic point group*. In the present example:

$$0 \to \mathbb{Z}^2 \to \widehat{G(C)} \to \widehat{P(C)} \to 1 \tag{15.577}$$

The elements in $\widehat{P(C)}$ are

$$\{(1,1),(R,\sigma),(R^2,1),(R^3,\sigma),(P,\sigma),(PR,1),(PR^2,\sigma),(PR^3,1)\} \tag{15.578}$$

This magnetic point group is isomorphic to $D_4$ but the time reversal homomorphism takes $\tau(R,\sigma) = -1$ and $\tau(P,\sigma) = -1$ so that we have

$$1 \to \mathbb{Z}_2 \times \mathbb{Z}_2 \to \widehat{P(C)} \xrightarrow{\tau} \mathbb{Z}_2 \to 1 \tag{15.579}$$

The induced automorphism on $\mathbb{Z}_2 \times \mathbb{Z}_2$ is trivial so clearly this sequence does not split, since $\widehat{P(C)} \cong D_4$ is nonabelian.

**Remarks**:

1. With the possible exception of exotic situations in which quantum gravity is important, physics takes place in space and time. Except in unusual situations associated with nontrivial gravitational fields we can assume our spacetime is time-orientable. Then, any physical symmetry group $G$ must be equipped with a homomorphism

   $$\tau : G \to \mathbb{Z}_2 \tag{15.580}$$

   telling us whether the symmetry operations preserve or reverse the orientation of time. That is $\tau(g) = +1$ are symmetries which preserve the orientation of time while $\tau(g) = -1$ are symmetries which reverse it.

   Now, suppose that $G$ is a symmetry of a quantum system. Then Wigner's theorem gives $G$ another grading $\phi : G \to \mathbb{Z}_2$, telling us whether the operator $\rho(g)$ implementing the symmetry transformation $g$ on the Hilbert space is unitary or anti-unitary. Thus, on very general grounds, a symmetry of a quantum system should be *bigraded* by a pair of homomorphisms $(\phi, \tau)$, or what is the same, a homomorphism to $\mathbb{Z}_2 \times \mathbb{Z}_2$.

   It is natural to ask whether $\phi$ and $\tau$ are related. A natural way to try to relate them is to study the dynamical evolution.

   In quantum mechanics, time evolution is described by unitary evolution of states. That is, there should be a family of unitary operators $U(t_1, t_2)$, strongly continuous in both variables and satisfying composition laws $U(t_1, t_3) = U(t_1, t_2)U(t_2, t_3)$ so that the density matrix $\varrho$ evolves according to:

   $$\varrho(t_1) = U(t_1, t_2)\varrho(t_2)U(t_2, t_1) \tag{15.581}$$

   Let us - for simplicity - make the assumption that our physical system has time-translation invariance so that $U(t_1, t_2) = U(t_1 - t_2)$ is a strongly continuous group of unitary transformations. [274]

   By Stone's theorem, $U(t)$ has a self-adjoint generator $H$, the Hamiltonian, so that

♣There is an obvious generalization of this statement for $U(t_1, t_2)$. Is it proved rigorously somewhere? ♣

[274] In the more general case we would need an analog of Stone's theorem to assert that there is a family of self-adjoint operators with $U(t_1, t_2) = \mathrm{Pexp}[-\frac{i}{\hbar}\int_{t_1}^{t_2} H(t')dt']$. Then, the argument we give below would lead to $\rho(g)H(t)\rho(g)^{-1} = \phi(g)\tau(g)H(t)$ for all $t$.

we may write

$$U(t) = \exp\left(-\frac{it}{\hbar}H\right) \tag{15.582}$$

Now, suppose we have a group [275] of operators on the Hilbert space: $\rho : G \to \mathrm{Aut}_{\mathbb{R}}(\mathcal{H})$. We say this group action is a *symmetry of the dynamics* if for all $g \in G$:

$$\rho(g)U(t)\rho(g)^{-1} = U(\tau(g)t) \tag{15.583}$$

where $\tau : G \to \mathbb{Z}_2$ is the indicator of time-orientation-reversal.

Now, substituting (15.582) and paying proper attention to $\phi$ we learn that the condition for a symmetry of the dynamics (15.583) is equivalent to

$$\phi(g)\rho(g)H\rho(g)^{-1} = \tau(g)H \tag{15.584}$$

in other words,

$$\rho(g)H\rho(g)^{-1} = \phi(g)\tau(g)H \tag{15.585}$$

Thus, the answer to our question is that $\phi$ and $\tau$ are *unrelated* in general. We should therefore define a third homomorphism $\chi : G \to \mathbb{Z}_2$

$$\chi(g) := \phi(g)\tau(g) \in \{\pm 1\} \tag{15.586}$$

Note that

$$\phi \cdot \tau \cdot \chi = 1 \tag{15.587}$$

2. It is very unusual for physical systems to have nontrivial homomorphisms $\chi$. That is, it is very unusual to have physical systems with time-orientation-reversing symmetries which are $\mathbb{C}$-linear or time-orientation-preserving symmetries which act $\mathbb{C}$-anti-linearly. But it is not impossible. To see why it is unusual note that:

$$\rho(g)H\rho(g)^{-1} = \chi(g)H \tag{15.588}$$

implies that if any group element has $\chi(g) = -1$ then the spectrum of $H$ must be symmetric around zero. In particular, if the spectrum is bounded below but not above this condition must fail. In many problems, e.g. in the standard Schrödinger problem with potentials which are bounded below, or in relativistic QFT with $H$ bounded below we must have $\chi(g) = 1$ for all $g$ and hence $\phi(g) = \tau(g)$, which is what one reads in virtually every physics textbook: "A symmetry is anti-unitary iff it reverses the orientation of time." Not true, in general.

3. However, there *are* physical examples where $\chi(g)$ can be non-trivial, that is, there can be symmetries which are both anti-unitary and time-orientation preserving. An example are the so-called "particle-hole" symmetries in free fermion systems.

[275] As explained in the 2012 article of Freed and Moore, this group might be an extension of the original group of quantum symmetries $\bar{\rho} : \bar{G} \to \mathrm{Aut}_{\mathrm{qtm}}(\mathbb{P}\mathcal{H})$.

### 15.6.3 $T^2 = (-1)^{2j}$ and the Clebsch-Gordon Decomposition

Above we checked that $T^2 = -1$ on spin 1/2 particles whose wavefunction obeys the usual Schrödinger equation with spin-orbit coupling. The generalization to spin $j$ particles is $T^2 = (-1)^{2j}$.

One simple way to see this is to note that the spin $j$ representation is obtained by decomposing the tensor product of $(2j)$ copies of the spin 1/2 representation. In general we have the very important Clebsch-Gordon decomposition:

$$V(j_1) \otimes V(j_2) \cong V(|j_1 - j_2|) \oplus V(|j_1 - j_2| + 1) \oplus \cdots\cdots \oplus V(j_1 + j_2) \tag{15.589}$$

which we discussed in section 11.20.

Note that every representation on the RHS of (15.589) has the same parity of $(-1)^{2j}$. Also note the triangular structure of the Clebsch-Gordon decomposition of $V(\frac{1}{2})^{\otimes n}$ allowing for an inductive proof. Finally $\tilde{T}^2$ on $V(\frac{1}{2})^{\otimes n}$ is just $(-1)^n$, so it is $(-1)^n$ on the highest summand $V(n/2)$.

## 15.7 General Extensions

Let us briefly return to the general extension (15.1). Thus, we are now not assuming that $N$ or $Q$ is abelian. We might ask what happens if we try to continue following the reasoning of section (15.1) in this general case, but now keeping in mind the nice classification of central extensions using group cohomology.

What we showed is that for *any* group extension a choice of a section $s : Q \to G$ automatically gives us two maps:

1. $\omega_s : Q \to \mathrm{Aut}(N)$
2. $f_s : Q \times Q \to N$

These two maps are <u>defined</u> by

$$\iota(\omega_{s,q}(n)) := s(q)\iota(n)s(q)^{-1} \tag{15.590}$$

and

$$s(q_1)s(q_2) := \iota(f_s(q_1, q_2))s(q_1 \cdot q_2) \tag{15.591}$$

respectively.

Now (15.590) defines an element of $\mathrm{Aut}(N)$ for fixed $s$ and $q$, but the map $q \mapsto \omega_{s,q}$ need not be a homomorphism, as we have repeatedly stressed. Rather, using (15.590) and (15.591) we can derive a twisted version of the homomorphism rule:

$$\omega_{s,q_1} \circ \omega_{s,q_2} = I(f_s(q_1, q_2)) \circ \omega_{s,q_1 q_2} \tag{15.592}$$

Recall that for $a \in N$, $I(a) \subset \mathrm{Aut}(N)$ denotes the inner automorphism given by conjugation by $a$. The proof of (15.592) follows exactly the same steps as (15.516), except for the very last line.

Moreover, using (15.591) to relate $s(q_1)s(q_2)s(q_3)$ to $s(q_1q_2q_3)$ in two ways gives a twisted cocycle relation:

$$\omega_{s,q_1}(f_s(q_2,q_3))f_s(q_1,q_2q_3) = f_s(q_1,q_2)f_s(q_1q_2,q_3) \tag{15.593}$$

Note this is the same as (15.517), but unlike that equation now order of the terms is very important since we no longer assume that $N$ is abelian.

To summarize: Given a general extension (15.1) there exist maps $(\omega_s, f_s)$, associated with any section $s$ and defined by (15.590) and (15.591). The maps $(\omega_s, f_s)$ automatically satisfy the identities (15.592) and (15.593).

We now consider, more generally, functions satisfying identities (15.592) and (15.593). That is, we assume we are given two maps (not necessarily derived from some section):

1. A map $f : Q \times Q \to N$
2. A map $\omega : Q \to \mathrm{Aut}(N)$

And we suppose the data $(\omega, f)$ satisfy the two conditions

$$\omega_{q_1} \circ \omega_{q_2} = I(f(q_1,q_2)) \circ \omega_{q_1q_2} \tag{15.594}$$

$$\omega_{q_1}(f(q_2,q_3))f(q_1,q_2q_3) = f(q_1,q_2)f(q_1q_2,q_3) \tag{15.595}$$

then we can construct an extension (15.1) with the multiplication law:

$$(n_1,q_1)\cdot_{f,\omega}(n_2,q_2) := (n_1\omega_{q_1}(n_2)f(q_1,q_2), q_1q_2) \tag{15.596}$$

This is very similar to (15.518) but we stress that since $N$ might be nonabelian, the order of the factors in the first entry on the RHS matters!

With a few lines of algebra, using the identities (15.594) and (15.595) one can check the associativity law and the other group axioms. We have already seen this simultaneous generalization of the semidirect product (14.2) and the twisted product of a central extension (15.70) in our discussion of the case where $N = A$ is abelian. (See equation (15.518) above.) The new thing we have now learned is that this is the most general way of putting a group structure on a product $N \times Q$ so that the result fits in an extension of $Q$ by $N$.

Now, suppose again that we are given a group extension. As we showed, a choice of section $s$ gives us a pair of functions $(\omega_s, f_s)$ satisfying (15.594) and (15.595). Any other section $\tilde{s}$ is related to $s$ by a function $t : Q \to N$. Indeed that function $t$ is defined by:

$$\tilde{s}(q) = \iota(t(q))s(q) \tag{15.597}$$

and one easily computes that we now have

$$\omega_{\tilde{s},q} = I(t(q)) \circ \omega_{s,q} \tag{15.598}$$

$$f_{\tilde{s}}(q_1,q_2) = t(q_1)\omega_{s,q_1}(t(q_2))f_s(q_1,q_2)t(q_1q_2)^{-1} \tag{15.599}$$

The proof of (15.598) follows exactly the same steps as (15.514). To prove (15.599) we patiently combine the definition (15.597) with the definition (15.591).

♣We also skipped this proof for $N = A$ abelian. Probably should show the steps. ♣

These formulae for how $(\omega_s, f_s)$ change as we change the section now motivate the following:

Suppose we are given a pair $(\omega, f)$ satisfying (15.594) and (15.595) and an arbitrary function $t : Q \to N$. We can now define a new pair $(\omega', f')$ by the equations:

$$\omega'_q = I(t(q)) \circ \omega_q \tag{15.600}$$

$$f'(q_1, q_2) = t(q_1)\omega_{q_1}(t(q_2))f(q_1, q_2)t(q_1 q_2)^{-1} \tag{15.601}$$

Now, with some algebra (DO IT!) one can check that indeed $(\omega', f')$ really do satisfy (15.594) and (15.595) as well. Equations (15.600) and (15.601) generalize the coboundary relation (15.65) of central extension theory. Note that the equations relating $\omega$ and $f$ back to $\omega'$ and $f'$ are of the same form with $t(q) \to t(q)^{-1}$.

The relations (15.600) and (15.601) define an equivalence relation on the set of pairs $(\omega, f)$ satisfying (15.594) and (15.595). Moreover, if $(\omega, f)$ and $(\omega', f')$ are related by (15.600) and (15.601) then we can define a group structure on the set $N \times Q$ in two ways using the equation (15.596) for each pair. Nevertheless, there is a morphism between these two extensions in the sense of (15.4) above where we define

$$\varphi(n, q) := (nt(q)^{-1}, q) \tag{15.602}$$

So, to check this you need to check

$$\varphi((n_1, q_1) \cdot_{f,\omega} (n_2, q_2)) = \varphi(n_1, q_1) \cdot_{f',\omega'} \varphi(n_2, q_2) \tag{15.603}$$

Then note that $\varphi^{-1}(n, q) = (nt(q), q)$ is an inverse morphism of extensions, and hence we have an isomorphism of extensions.

Now we would like to state all this a little more conceptually. The first point to note is that a map $q \mapsto \omega_q \in \mathrm{Aut}(N)$ that satisfies (15.594) in fact canonically defines a homomorphism $\bar{\omega} : Q \to \mathrm{Out}(N)$ of $Q$ into the group of outer automorphisms of $N$. This homomorphism is defined more conceptually as the unique map that makes the diagram:

$$\begin{array}{ccccccccc} 1 & \longrightarrow & N & \xrightarrow{\iota} & G & \xrightarrow{\pi} & Q & \longrightarrow & 1 \\ & & \downarrow{\scriptstyle I} & & \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \overline{\omega}} & & \\ 1 & \longrightarrow & \mathrm{Inn}(N) & \longrightarrow & \mathrm{Aut}(N) & \longrightarrow & \mathrm{Out}(N) & \longrightarrow & 1 \end{array} \tag{15.604}$$

Here $I : N \to \mathrm{Inn}(N)$ is the map that takes $n$ to the inner automorphism $I(n) : n' \mapsto nn'n^{-1}$ and $\psi$ is the map from $G \to \mathrm{Aut}(N)$ defined by

$$\iota(\psi(g)(n)) = g\iota(n)g^{-1} \tag{15.605}$$

Now, we can ask the converse question: *Given an arbitrary homomorphism* $\bar{\omega} : Q \to \mathrm{Out}(N)$ *is there an extension of* $Q$ *by* $N$ *that induces it as in* (15.604)*?*

The most obvious thing to try when trying to answer this question is to use $\bar{\omega} : Q \to \mathrm{Out}(N)$ and the pullback construction (15.34) of the canonical exact sequence given by the lower line of (15.604). But this will only give an extension of $Q$ by $\mathrm{Inn}(N)$. Note that

$\mathrm{Inn}(N) \cong N/Z(N)$, and so the center of $N$ might cause some trouble. That is in fact what happens: The answer to the above question is, in general, "NO," and the obstruction has to do with the third cohomology group $H^{3+\bar{\omega}}(Q, Z(N))$ where $Z(N)$ is the center of $N$. See section 15.8.5 below.

But for now, let us suppose we have a choice of $\bar{\omega}$ such that extensions inducing it do exist. What can we say about the set $\mathrm{Ext}^{\bar{\omega}}(Q, N)$ of equivalence classes of such extensions?

To answer this we choose a lifting of the homomorphism, that is, a map $q \mapsto \omega_q \in \mathrm{Aut}(N)$. Now, if we have two extensions both inducing $\bar{\omega}$ and we choose two liftings $\omega_q^{(1)}$ and $\omega_q^{(2)}$ then they will be related by

$$\omega_q^{(1)} = I(t(q)) \circ \omega_q^{(2)} \tag{15.606}$$

for some function $t : Q \to N$. Note, please, that while this equation is formally very similar to (15.598) it is conceptually different. Nothing has been said about the relation of the two extensions, other than that they induce the same $\bar{\omega}$.

Now we try to relate the corresponding functions $f^{(1)}(q_1, q_2)$ and $f^{(2)}(q_1, q_2)$. To do that we compute

$$\begin{aligned}
\omega_{q_1}^{(1)} \circ \omega_{q_2}^{(1)}(n) &= t(q_1)\omega_{q_1}^{(2)}\left(t(q_2)\omega_{q_2}^{(2)}(n)t(q_2)^{-1}\right)t(q_1)^{-1} \\
&= t(q_1)\omega_{q_1}^{(2)}(t(q_2))\left(\omega_{q_1}^{(2)} \circ \omega_{q_2}^{(2)}(n)\right)\omega_{q_2}^{(2)}(t(q_2)^{-1})t(q_1)^{-1} \\
&= t(q_1)\omega_{q_1}^{(2)}(t(q_2))\left(f^{(2)}(q_1, q_2)\omega_{q_1 q_2}^{(2)}(n)f^{(2)}(q_1, q_2)^{-1}\right)\omega_{q_2}^{(2)}(t(q_2)^{-1})t(q_1)^{-1} \\
&= \left\{t(q_1)\omega_{q_1}^{(2)}(t(q_2))f^{(2)}(q_1, q_2)t(q_1 q_2)^{-1}\right\} \cdot \omega_{q_1 q_2}^{(1)}(n)\left\{t(q_1)\omega_{q_1}^{(2)}(t(q_2))f^{(2)}(q_1, q_2)t(q_1 q_2)^{-1}\right\}^{-1} \\
&= \hat{f}^{(2)}(q_1, q_2) \cdot \omega_{q_1 q_2}^{(1)}(n)\hat{f}^{(2)}(q_1, q_2)^{-1}
\end{aligned} \tag{15.607}$$

where we define

$$\hat{f}^{(2)}(q_1, q_2) := t(q_1)\omega_{q_1}^{(2)}(t(q_2))f^{(2)}(q_1, q_2)t(q_1 q_2)^{-1} \tag{15.608}$$

On the other hand, we know that

$$\omega_{q_1}^{(1)} \circ \omega_{q_2}^{(1)}(n) = f^{(1)}(q_1, q_2)\omega_{q_1 q_2}^{(1)}(n)f^{(1)}(q_1, q_2)^{-1} \tag{15.609}$$

Can we conclude that $\hat{f}^{(2)}(q_1, q_2) = f^{(1)}(q_1, q_2)$ ? Certainly not! Provided $\omega_{q_1 q_2}^{(1)}(n)$ is sufficiently generic all we can conclude is that

$$\hat{f}^{(2)}(q_1, q_2) = f^{(1)}(q_1, q_2)\zeta(q_1, q_2) \tag{15.610}$$

for some function $\zeta : Q \times Q \to Z(N)$. These two functions are not necessarily related by a coboundary and the extensions are not necessarily equivalent!

What is true is that if $\hat{f}^{(2)}$ and $f^{(1)}$ satisfy the twisted cocycle relation then $\zeta(q_1, q_2)$ in (15.610) also satisfies the twisted cocycle relation. (This requires a lot of patient algebra....) It follows that

$$\zeta \in Z^{2+\bar{\omega}}(Q, Z(N)) \tag{15.611}$$

Moreover, going the other way, given one extension and corresponding $(\omega^{(1)}, f^{(1)})$, and a $\zeta \in Z^{2+\bar{\omega}}(Q, Z(N))$ we can change $f$ as in (15.610). If $[z] \in H^{2+\bar{\omega}}(Q, Z(N))$ is nontrivial we will in general get a new, nonequivalent extension.

All this is summarized by the theorem:

**Theorem**: Let $\mathrm{Ext}^{\bar{\omega}}(Q, N)$ be the set of inequivalent extensions of $Q$ by $N$ inducing $\bar{\omega}$. Then either this set it is empty or it is a torsor [276] for $H^{2+\bar{\omega}}(Q, Z(N))$.

************************

NEED SOME EXAMPLES HERE. AND NEED SOME MORE INTERESTING EXERCISES.

*************************

---

**Exercise** *Checking the group laws*

Show that (15.596) really defines a group structure.

a.) Check the associativity relation.

b.) What is the identity element? [277]

c.) Check that every element has an inverse.

---

---

**Exercise**

a.) Check that (15.602) really does define a homomorphism of the group laws (15.596) defined by $(\omega, f)$ and $(\omega', f')$ if $(\omega', f')$ is related to $(\omega, f)$ by (15.600) and (15.601).

b.) Check that the diagram (15.4) really does commute if we use (15.602).

---

## 15.8 Group cohomology in other degrees

Motivations:

a.) The word "cohomology" suggests some underlying chain complexes, so we will show that there is such a formulation.

b.) There has been some discussion of higher degree group cohomology in physics in

1. The theory of anomalies (Faddeev-Shatashvili; Segal; Carey et. al.; Mathai et. al.; ... )

2. Classification of rational conformal field theories (Moore-Seiberg; Dijkgraaf-Vafa-Verlinde-Verlinde; Dijkgraaf-Witten; Kapustin-Saulina)

[276] A *torsor* $X$ for a group $G$ is a set $X$ with a $G$-action on it so that given any pair $x, x' \in X$ there is a unique $g \in G$ that maps $x$ to $x'$. In this chapter we have discussed an important example of a torsor quite extensively: Affine space $\mathbb{A}^d$ is a torsor for $\mathbb{R}^d$ with the natural action of $\mathbb{R}^d$ on $\mathbb{A}^d$ by translation.

[277] *Answer*: $(f(1,1)^{-1}, 1_Q)$.

3. Chern-Simons theory and topological field theory (Dijkgraaf-Witten,...)
4. Condensed matter/topological phases of matter (Kitaev; Wen et. al.; Kapustin et. al.; Freed-Hopkins;....)
5. Three-dimensional supersymmetric gauge theory.

Here we will be brief and just give the basic definitions:

### 15.8.1 Definition

Suppose we are given any group $G$ and an Abelian group $A$ (written additively in this sub-section) and a homomorphism

$$\omega : G \to \mathrm{Aut}(A) \tag{15.612}$$

**Definition**: An $n$-cochain is a function $\phi : G^{\times n} \to A$. The space of $n$-cochains is denoted $C^n(G, A)$. It is also useful to speak of 0-cochains. We interpret a 0-cochain $\phi_0$ to be some element $\phi_0 = a \in A$.

Note that $C^n(G, A)$, for $n \geq 0$, is an abelian group using the abelian group structure of $A$ on the values of $\phi$, that is: $(\phi_1 + \phi_2)(\vec{g}) := \phi_1(\vec{g}) + \phi_2(\vec{g})$.

Define a group homomorphism: $d : C^n(G, A) \to C^{n+1}(G, A)$

$$\begin{aligned}(d\phi)(g_1, \dots, g_{n+1}) &:= \omega_{g_1}\left(\phi(g_2, \dots, g_{n+1})\right)\\ -\phi(g_1 g_2, g_3, \dots, g_{n+1}) &+ \phi(g_1, g_2 g_3, \dots, g_{n+1}) \pm \cdots + (-1)^n \phi(g_1, \dots, g_{n-1}, g_n g_{n+1})\\ &+ (-1)^{n+1}\phi(g_1, \dots, g_n)\end{aligned} \tag{15.613}$$

Then we have, for $n = 0$:

$$(d\phi_0)(g) = \omega_g(a) - a \tag{15.614}$$

For $n = 1$, $n = 2$ and $n = 3$ the formula written out looks like:

$$(d\phi_1)(g_1, g_2) = \omega_{g_1}\left(\phi_1(g_2)\right) - \phi_1(g_1 g_2) + \phi_1(g_1) \tag{15.615}$$

$$(d\phi_2)(g_1, g_2, g_3) = \omega_{g_1}\left(\phi_2(g_2, g_3)\right) - \phi_2(g_1 g_2, g_3) + \phi_2(g_1, g_2 g_3) - \phi_2(g_1, g_2) \tag{15.616}$$

$$(d\phi_3)(g_1, g_2, g_3, g_4) = \omega_{g_1}\left(\phi_3(g_2, g_3, g_4)\right) - \phi_3(g_1 g_2, g_3, g_4) + \phi_3(g_1, g_2 g_3, g_4) - \phi_3(g_1, g_2, g_3 g_4) + \phi_3(g_1, g_2, g_3) \tag{15.617}$$

Next, one can check that for any $\phi$, we have the absolutely essential equation:

$$\boxed{d(d\phi) = 0} \tag{15.618}$$

We will give a simple proof of (15.618) below but let us just look at how it works for the lowest degrees: If $\phi_0 = a \in A$ is a 0-cochain then

$$\begin{aligned}(d^2\phi_0)(g_1, g_2) &= \omega_{g_1}(d\phi_0(g_2)) - d\phi_0(g_1 \cdot g_2) + d\phi_0(g_1)\\ &= \omega_{g_1}\left(\omega_{g_2}(a) - a\right) - \left(\omega_{g_1 g_2}(a) - a\right) + \left(\omega_{g_1}(a) - a\right)\\ &= \omega_{g_1}(\omega_{g_2}(a)) - \omega_{g_1 g_2}(a)\\ &= 0\end{aligned} \tag{15.619}$$

if $\phi_1$ is any 1-cochain then we compute:

$$\begin{aligned}(d^2\phi_1)(g_1,g_2,g_3) &= \omega_{g_1}(d\phi_1(g_2,g_3)) - (d\phi_1)(g_1g_2,g_3) + (d\phi_1)(g_1,g_2g_3) - (d\phi_1)(g_1,g_2)\\ &= \omega_{g_1}\left(\omega_{g_2}(\phi_1(g_3)) - \phi_1(g_2g_3) + \phi_1(g_2)\right)\\ &- \left(\omega_{g_1g_2}(\phi_1(g_3)) - \phi_1(g_1g_2g_3) + \phi_1(g_1g_2)\right)\\ &+ \left(\omega_{g_1}(\phi_1(g_2g_3) - \phi_1(g_1g_2g_3) + \phi_1(g_1)\right)\\ &- \left(\omega_{g_1}(\phi_1(g_2)) - \phi_1(g_1g_2) + \phi_1(g_1)\right)\\ &= 0\end{aligned} \tag{15.620}$$

where you can check that all terms cancel in pairs, once you use $\omega_{g_1} \circ \omega_{g_2} = \omega_{g_1g_2}$.

The set of ($\omega$-twisted) $n$-cocycles is defined to be the subgroup $Z^{n+\omega}(G,A) \subset C^n(G,A)$ of cochains that satisfy $d\phi_n = 0$.

Thanks to (15.618) we can define a subgroup $B^{n+\omega}(G,A) \subset Z^{n+\omega}(G,A)$, called the subgroup of coboundaries:

$$B^{n+\omega}(G,A) := \{\phi_n | \exists \phi_{n-1} \qquad s.t. \qquad d\phi_{n-1} = \phi_n\} \tag{15.621}$$

then, since $d^2 = 0$ we have $B^{n+\omega}(G,A) \subset Z^{n+\omega}(G,A)$.

Then the group cohomology is defined to be the quotient

$$H^{n+\omega}(G,A) = Z^{n+\omega}(G,A)/B^{n+\omega}(G,A) \tag{15.622}$$

**Example**: Let us take $G = \mathbb{Z}_2 = \{1,\sigma\}$ and $A = \mathbb{Z}$. Recall that

$$\mathrm{Aut}(A) = \mathrm{Aut}(\mathbb{Z}) = \{\mathrm{Id}_{\mathbb{Z}}, \mathcal{P}\} \cong \mathbb{Z}_2 \tag{15.623}$$

where $\mathcal{P}$ is the automorphism that takes $\mathcal{P} : n \to -n$. Now $\mathrm{Hom}(G, \mathrm{Aut}(\mathbb{Z})) \cong \mathbb{Z}_2$. Of course, $\omega_1 = \mathrm{Id}_{\mathbb{Z}}$ always and now we have two possibilities for $\omega_\sigma$. Either $\omega_\sigma = \mathrm{Id}_{\mathbb{Z}}$ in which case we denote $\omega = T$ ("$T$" for trivial) or $\omega_\sigma = \mathcal{P}$ which we will denote $\mathcal{I}$. Let us compute $H^{1+\omega}(\mathbb{Z}_2,\mathbb{Z})$ for these two possibilities. First look at the subgroup of coboundaries. If $\phi_0 = n_0 \in \mathbb{Z}$ is some integer then

$$\begin{aligned}(d\phi_0)(1) &= 0\\ (d\phi_0)(\sigma) &= \omega_\sigma(n_0) - n_0 = \begin{cases} 0 & \omega = T \\ -2n_0 & \omega = \mathcal{I}\end{cases}\end{aligned} \tag{15.624}$$

Now consider the differential of a one-cochain:

$$\begin{aligned}(d\phi_1)(1,1) &= \omega_1(\phi_1(1)) - \phi_1(1) + \phi_1(1) = \phi_1(1)\\ (d\phi_1)(1,\sigma) &= \omega_1(\phi_1(\sigma)) - \phi_1(\sigma) + \phi_1(1) = \phi_1(1)\\ (d\phi_1)(\sigma,1) &= \omega_\sigma(\phi_1(1)) - \phi_1(\sigma) + \phi_1(\sigma) = \omega_\sigma(\phi_1(1))\\ (d\phi_1)(\sigma,\sigma) &= \omega_\sigma(\phi_1(\sigma)) - \phi_1(1) + \phi_1(\sigma)\end{aligned} \tag{15.625}$$

Now the cocycle condition implies $\phi_1(1) = 0$, making the first three lines of (15.625) vanish. Using this the fourth line becomes:

$$(d\phi_1)(\sigma, \sigma) = \omega_\sigma(\phi_1(\sigma)) + \phi_1(\sigma) = \begin{cases} 2\phi_1(\sigma) & \omega = T \\ 0 & \omega = \mathcal{I} \end{cases} \tag{15.626}$$

Now, when is $\phi_1$ a cocycle? When $\omega = T$ is trivial then we must take $\phi_1(\sigma) = 0$ and hence $\phi_1 = 0$ moreover, there are no coboundaries. We find $H^{1+T}(\mathbb{Z}_2, \mathbb{Z}) = 0$ in this case, reproducing the simple fact that there are no nontrivial group homomorphisms from $\mathbb{Z}_2$ to $\mathbb{Z}$.

On the other hand, when $\omega = \mathcal{I}$ we can take $\phi_1(\sigma) = a$ to be any integer $a \in \mathbb{Z}$. The group of twisted cocycles is isomorphic to $\mathbb{Z}$. However, now there are nontrivial coboundaries, as we see from (15.624). We can shift $a$ by any even integer $a \to a - 2n_0$. So

$$H^{1+\mathcal{I}}(\mathbb{Z}_2, \mathbb{Z}) \cong \mathbb{Z}_2 \tag{15.627}$$

In addition to the interpretation in terms of splittings, this has a nice interpretation in topology in terms of the unorientability of even-dimensional real projective spaces.

**Remarks**:

1. Previously we were denoting the cohomology groups by $H^{n+\omega}(G, A)$. In the equations above the $\omega$ is still present, (see the first term in the definition of $d\phi$) but we leave the $\omega$ implicit in the notation. Nevertheless, we are talking about the same groups as before, but now generalizing to arbitrary degree $n$.

2. Remembering that we are now writing our abelian group $A$ additively, we see that the equation $(d\phi_2) = 0$ is just the twisted 2-cocycle conditions, and $\phi_2' = \phi_2 + d\phi_1$ are two different twisted cocycles related by a coboundary. See equations (15.517) and (15.519) above. Roughly speaking, you should "take the logarithm" of these equations.

3. *Homological Algebra*: What we are discussing here is a special case of a topic known as homological algebra. Quite generally, a *chain complex* is a sequence of Abelian groups $\{C_n\}_{n \in \mathbb{Z}}$ equipped with group homomorphisms

$$\partial_n : C_n \to C_{n-1} \tag{15.628}$$

such that $\partial_n \circ \partial_{n+1} = 0$ for all $n \in \mathbb{Z}$. A *cochain complex* is similarly a sequence of Abelian groups $\{C^n\}_{n \in \mathbb{Z}}$ with group homomorphisms $d_n : C^n \to C^{n+1}$ so that $d_{n+1} \circ d_n = 0$ for all $n \in \mathbb{Z}$. Note that these are NOT exact sequences. Indeed the failure to be an exact sequence is measured by the *homology groups* of the chain complex

$$H_n(C_*, \partial_*) := \ker(\partial_n)/\mathrm{im}(\partial_{n+1}) \tag{15.629}$$

and the *cohomology groups* of the cochain complex:

$$H^n(C^*, d_*) := \ker(d_n)/\mathrm{im}(d_{n-1}) \tag{15.630}$$

4. *Homogeneous cocycles*: A nice way to prove that $d^2 = 0$ is the following. We define *homogeneous* $n$-cochains to be maps $\varphi : G^{n+1} \to A$ which satisfy

$$\varphi(hg_0, hg_1, \dots, hg_n) = \omega_h\left(\varphi(g_0, g_1, \dots, g_n)\right) \tag{15.631}$$

Let $\mathcal{C}^n(G, A)$ denote the abelian group of such homogeneous group cochains. (Warning! Elements of $\mathcal{C}^n(G, A)$ have $(n+1)$ arguments!) Define

$$\delta : \mathcal{C}^n(G, A) \to \mathcal{C}^{n+1}(G, A) \tag{15.632}$$

by

$$\delta\varphi(g_0, \dots, g_{n+1}) := \sum_{i=0}^{n+1} (-1)^i \varphi(g_0, \dots, \widehat{g_i}, \dots, g_{n+1}) \tag{15.633}$$

where $\widehat{g_i}$ means the argument is omitted. Clearly, if $\varphi$ is homogeneous then $\delta\varphi$ is also homogeneous. It is then very straightforward to prove that $\delta^2 = 0$. Indeed, if $\varphi \in \mathcal{C}^{n-1}(G, A)$ we compute:

$$\begin{aligned}
\delta^2\varphi(g_0, \dots, g_{n+1}) &= \sum_{i=0}^{n+1} (-1)^i \Bigg\{ \sum_{j=0}^{i-1} (-1)^j \varphi(g_0, \dots, \widehat{g_j}, \dots, \widehat{g_i}, \dots, g_{n+1}) \\
&\qquad - \sum_{j=i+1}^{n+1} (-1)^j \varphi(g_0, \dots, \widehat{g_i}, \dots, \widehat{g_j}, \dots, g_{n+1}) \Bigg\} \\
&= \sum_{0 \le j < i \le n+1} (-1)^{i+j} \varphi(g_0, \dots, \widehat{g_j}, \dots, \widehat{g_i}, \dots, g_{n+1}) \\
&\qquad - \sum_{0 \le i < j \le n+1} (-1)^{i+j} \varphi(g_0, \dots, \widehat{g_i}, \dots, \widehat{g_j}, \dots, g_{n+1}) \\
&= 0
\end{aligned} \tag{15.634}$$

Now, we can define an isomorphism $\psi : \mathcal{C}^n(G, A) \to C^n(G, A)$ by defining

$$\phi_n(g_1, \dots, g_n) := \varphi_n(1, g_1, g_1 g_2, \dots, g_1 \cdots g_n) \tag{15.635}$$

That is, when $\phi_n$ and $\varphi_n$ are related this way we say $\phi_n = \psi(\varphi_n)$. Now one can check that the simple formula (15.633) becomes the more complicated formula (15.613). Put more formally: there is a unique $d$ so that $d\psi = \psi\delta$, or even more formally, there is a unique group homomorphism $d$ such that we have a commutative diagram:

$$\begin{array}{ccc}
\mathcal{C}^n(G, A) & \xrightarrow{\delta} & \mathcal{C}^{n+1}(G, A) \\
\downarrow \psi & & \downarrow \psi \\
C^n(G, A) & \xrightarrow{d} & C^{n+1}(G, A)
\end{array} \tag{15.636}$$

For example, if

$$\phi_1(g) = \psi(\varphi_1)(g) = \varphi_1(1, g) \tag{15.637}$$

then we can check that

$$
\begin{aligned}
(d\phi_1)(g_1,g_2) &= d(\psi(\varphi_1))(g_1,g_2) \\
&= \psi(\delta\varphi_1)(g_1,g_2) \\
&= \delta\varphi_1(1,g_1,g_1g_2) \\
&= \varphi_1(g_1,g_1g_2) - \varphi_1(1,g_1g_2) + \varphi_1(1,g_1) \\
&= \omega_{g_1}(\varphi_1(1,g_2)) - \varphi_1(1,g_1g_2) + \varphi_1(1,g_1) \\
&= \omega_{g_1}(\phi_1(g_2)) - \phi_1(g_1g_2) + \phi_1(g_1)
\end{aligned} \tag{15.638}
$$

in accord with the previous definition!

5. Where do all these crazy formulae come from? The answer is in topology. We will indicate it briefly in our discussion of categories and groupoids below.

6. The reader will probably find these formulae a bit opaque. It is therefore good to stop and think about what the cohomology is measuring, at least in low degrees.

---

**Exercise**

Derive the formula for the differential on an inhomogeneous cochain $d\phi_2$ starting with the definition on the analogous homogeneous cochain $\varphi_3$

---

---

**Exercise**

If $(C_n, \partial_n)$ is a chain complex show that one can define a cochain complex with groups:

$$C^n := \mathrm{Hom}(C_n, \mathbb{Z}) \tag{15.639}$$

---

### 15.8.2 Interpreting the meaning of $H^{0+\omega}$

A zero-cocycle is an element $a \in A$ so that for all $g$

$$\omega_g(a) = a \tag{15.640}$$

There are no coboundaries to worry about, so $H^0(G,A)$ is just the set of fixed points of the $G$ action on $A$.

### 15.8.3 Interpreting the meaning of $H^{1+\omega}$

We have interpreted $H^{1+\omega}(G,A)$ above as the set of nontrivial splittings of the semidirect product defined by $\omega$:

$$0 \to A \to A \rtimes G \to G \to 1 \tag{15.641}$$

### 15.8.4 Interpreting the meaning of $H^{2+\omega}$

Again, we have interpreted $H^{2+\omega}(G, A)$ as $\mathrm{Ext}^{\omega}(G, A)$, the set of equivalence classes of extensions

$$0 \to A \to \tilde{G} \to G \to 1 \tag{15.642}$$

inducing a fixed $\omega : G \to \mathrm{Aut}(A)$. The trivial element of the cohomology group corresponds to the semi-direct product and the set of inequivalent trivializations is the group $H^{1+\omega}(G, A)$ of splittings of the semi-direct product.

More generally, $\mathrm{Ext}^{\bar{\omega}}(Q, N)$ is a torsor for $H^{2+\bar{\omega}}(Q, Z(N))$.

### 15.8.5 Interpreting the meaning of $H^3$

To see one interpretation of $H^3$ in terms of extension theory let us return to the analysis of general extensions in §15.7.

Recall that, as we have discussed using (15.604), a general extension (15.1) has a canonically associated homomorphism

$$\bar{\omega} : Q \to \mathrm{Out}(N) \tag{15.643}$$

where $\mathrm{Out}(N)$ is the group of outer automorphisms of $N$.

The natural question arises: *Given a homomorphism* $\bar{\omega}$ *as in* (15.643) *is there a corresponding extension of* $Q$ *by* $N$ *inducing* $\bar{\omega}$ *as in equation* (15.604) *?*

To answer this question we could proceed by *choosing* for each $q \in Q$ an automorphism $\xi_q \in \mathrm{Aut}(N)$ such that $[\xi_q] = \bar{\omega}_q$ in $\mathrm{Out}(N)$. To do this, choose a section $s$ of $\pi : \mathrm{Aut}(N) \to \mathrm{Out}(N)$ and let $\xi_q := s(\bar{\omega}_q)$. If we cannot split the sequence

$$1 \to \mathrm{Inn}(N) \to \mathrm{Aut}(N) \to \mathrm{Out}(N) \to 1 \tag{15.644}$$

then $q \mapsto \xi_q$ will not be a group homomorphism. But we do know that for all $q_1, q_2 \in Q$

$$\xi_{q_1} \circ \xi_{q_2} \circ \xi^{-1}_{q_1 q_2} \in \mathrm{Inn}(N) \tag{15.645}$$

Therefore, for every $q_1, q_2$ we may *choose* an element $f(q_1, q_2) \in N$ so that

$$\xi_{q_1} \circ \xi_{q_2} \circ \xi^{-1}_{q_1 q_2} = I(f(q_1, q_2)) \tag{15.646}$$

i.e.

$$\xi_{q_1} \circ \xi_{q_2} = I(f(q_1, q_2)) \circ \xi_{q_1 q_2} \tag{15.647}$$

Of course, the choice of $f(q_1, q_2)$ is ambiguous by an element of $Z(N)$!

Equation (15.647) is of course just (15.594) written in slightly different notation. Therefore, as we saw in §15.7, if $f(q_1, q_2)$ were to satisfy the the "twisted cocycle condition" (15.595) then we could use (15.596) to define an extension inducing $\bar{\omega}$.

Therefore, let us check if some choice of $f(q_1, q_2)$ actually does satisfy the twisted cocycle condition (15.595). Looking at the RHS of (15.595) we compute:

$$\begin{aligned} I(f(q_1, q_2) f(q_1 q_2, q_3)) &= I(f(q_1, q_2)) I(f(q_1 q_2, q_3)) \\ &= \left(\xi_{q_1} \circ \xi_{q_2} \circ \xi^{-1}_{q_1 q_2}\right) \circ \left(\xi_{q_1 q_2} \circ \xi_{q_3} \circ \xi^{-1}_{q_1 q_2 q_3}\right) \\ &= \xi_{q_1} \circ \xi_{q_2} \circ \xi_{q_3} \circ \xi^{-1}_{q_1 q_2 q_3} \end{aligned} \tag{15.648}$$

On the other hand, looking at the LHS of (15.595) we compute:

$$
\begin{aligned}
I(\xi_{q_1}(f(q_2,q_3))f(q_1,q_2q_3)) &= I(\xi_{q_1}(f(q_2,q_3)))I(f(q_1,q_2q_3)) \\
&= \xi_{q_1}\circ I((f(q_2,q_3)))\circ\xi_{q_1}^{-1}\circ I(f(q_1,q_2q_3)) \\
&= \xi_{q_1}\circ\left(\xi_{q_2}\circ\xi_{q_3}\circ\xi_{q_2q_3}^{-1}\right)\circ\xi_{q_1}^{-1}\circ\left(\xi_{q_1}\circ\xi_{q_2q_3}\circ\xi_{q_1q_2q_3}^{-1}\right) \\
&= \xi_{q_1}\circ\xi_{q_2}\circ\xi_{q_3}\circ\xi_{q_1q_2q_3}^{-1}
\end{aligned}
\tag{15.649}
$$

Therefore, comparing (15.648) and (15.649) we conclude that

$$
I(\xi_{q_1}(f(q_2,q_3))f(q_1,q_2q_3)) = I(f(q_1,q_2)f(q_1q_2,q_3)) \tag{15.650}
$$

We cannot conclude that $f$ satisfies the twisted cocycle equation from this identity because inner transformations are trivial for elements in the center $Z(N)$. Rather, what we can conclude is that for every $q_1,q_2,q_3$ there is an element $z(q_1,q_2,q_3)\in Z(N)$ such that

$$
f(q_1,q_2)f(q_1q_2,q_3) = z(q_1,q_2,q_3)\xi_{q_1}(f(q_2,q_3))f(q_1,q_2q_3) \tag{15.651}
$$

Now, one can check (with a lot of algebra) that

1. $z$ is a cocycle in $Z^{3+\bar{\omega}}(Q,Z(N))$. (We are using $\mathrm{Aut}(Z(N))\cong\mathrm{Out}(Z(N))$.)
2. Changes in choices of $\xi_q$ and $f(q_1,q_2)$ lead to changes in $z$ by a coboundary.

and therefore we conclude:

**Theorem 15.8.5.1** : Given $\bar{\omega}: Q\to\mathrm{Out}(N)$ there exists an extension of $Q$ by $N$ iff the cohomology class $[z]\in H^3(Q,Z(N))$ vanishes.

Moreover, as we have seen, if $[z]=0$ then the trivializations of $z$ are in 1-1 correspondence with elements $H^2(Q,Z(N))$ and are hence in 1-1 correspondence with isomorphism classes of extensions of $Q$ by $N$. This is the analogue, one step up in degree, of our interpretation of $H^1(G,A)$.

**Examples**: As an example [278] where a degree three cohomology class obstructs the existence of an extension inducing a homomorphism $\bar{\omega}: Q\to\mathrm{Out}(N)$ we can take $N$ to be the generalized quaternion group of order 16. It is generated by $x$ and $y$ satisfying:

$$
x^4=y^2 \qquad x^8=1 \qquad yxy^{-1}=x^{-1} \tag{15.652}
$$

Using these relations every word in $x^{\pm1}$ and $y^{\pm1}$ can be reduced to either $x^m$, or $yx^m$, with $m=0,\dots,7$, and these words are all different. One can show the outer automorphism group $\mathrm{Out}(N)\cong\mathbb{Z}_2\times\mathbb{Z}_2$ with generators $\alpha,\beta$ acting by

$$
\alpha(x)=x^3,\alpha(y)=y \qquad\qquad \beta(x)=x,\beta(y)=yx \tag{15.653}
$$

Then there is is no group extension with group $G$ fitting in

$$
1\to N\to G\to\mathbb{Z}_2\to 1 \tag{15.654}
$$

[278] I learned these nice examples from Clay Cordova. They will appear in a forthcoming paper with Po-Shen Hsin and Francesco Benini.

inducing the homomorphism $\bar{\omega} : \mathbb{Z}_2 \to \mathrm{Out}(N)$ defined by $\bar{\omega}(\sigma) = \alpha \circ \beta$ where $\sigma$ is the nontrivial element of $\mathbb{Z}_2$. One way to prove this is to look up the list of groups of order 32 and search for those with maximal normal subgroup given by $N$. [279] There are five such. Then one computes $\bar{\omega}$ for each such extension and finds that it is never of the above type. A similar example can be constructed by taking $N$ to be a dihedral group of order 16.

♣Really should explain four-term sequences and crossed modules here.... ♣

**Remark** There is an interpretation of $H^3(Q, Z(N))$ as a classification of four-term exact sequences, and there are generalizations of this to higher degree. See

1. K. Brown, *Group Cohomology.*
2. C. A. Weibel, *An introduction to homological algebra*, chapter 6

### 15.9 Some references

Some online sources with links to further material are

1. http://en.wikipedia.org/wiki/Group-extension
2. http://ncatlab.org/nlab/show/group+extension

3 http://terrytao.wordpress.com/2010/01/23/some-notes-on-group-extensions/

4. Section 15.8.5, known as the Artin-Schreier theory, is based on a nice little note by P.J. Morandi,

http://sierra.nmsu.edu/morandi/notes/GroupExtensions.pdf

5. Jungmann, Notes on Group Theory

6. S. MacLane, "Topology And Logic As A Source Of Algebra," Bull. Amer. Math. Soc. 82 (1976), 1-4.

Textbooks:

1. K. Brown, Group Cohomology
2. Karpilovsky, The Schur Multiplier
3. C. A. Weibel, *An introduction to homological algebra*, chapter 6

## 16. Overview of general classification theorems for finite groups

In general if a mathematical object proves to be useful then there is always an associated important problem, namely the *classification* of these objects.

For example, with groups we can divide them into classes: finite and infinite, abelian and nonabelian producing a four-fold classication:

| | |
|---|---|
| Finite abelian | Finite nonabelian |
| Infinite abelian | Infinite nonabelian |

[279] See, for example, B. Shuster, "Morava K-theory of groups of order 32," Algebr. Geom. Topol. **11** (2011) 503-521.

But this is too rough, it does not give us a good feeling for what the examples really are.

Once we have a "good" criterion we often can make a nontrivial statement about the general structure of objects in a given class. Ideally, we should be able to construct all the examples algorithmically, and be able to distinguish the ones which are not isomorphic. Of course, finding such a "good" criterion is an art. For example, classification of infinite nonabelian groups is completely out of the question. But in Chapter *** we will see that an important class of infinite nonabelian groups, the simply connected compact simple Lie groups, have a very beautiful classification: There are four infinite sequences of classical matrix groups: $SU(n), Spin(n), USp(2n)$ and then five exceptional cases with names $G_2, F_4, E_6, E_7, E_8$. [280]

One might well ask: Can we classify finite groups? In this section we survey a little of what is known about this problem.

### 16.1 Brute force

If we just start listing groups of low order we soon start to appreciate what a jungle is out there.

But let us try, if only as an exercise in applying what we have learned so far. First, let us note that for groups of order $p$ where $p$ is prime we automatically have the unique possibility of the cyclic group $\mathbb{Z}/p\mathbb{Z}$. Similarly, for groups of order $p^2$ there are precisely two possibilities: $\mathbb{Z}/p^2\mathbb{Z}$ and $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. This gets us through many of the low order cases.

Given this remark the first nontrivial order to work with is $|G| = 6$. By Cauchy's theorem there are elements of order 2 and 3. Call them $b$, with $b^2 = 1$ and $a$ with $a^3 = 1$. Then $(bab)^3 = 1$, so either

1. $bab = a$ which implies $ab = ba$ which implies $G = \mathbb{Z}_2 \times \mathbb{Z}_3 = \mathbb{Z}_6$
2. $bab = a^{-1}$ which implies $G = D_3$.

This is the first place we meet a nonabelian group. It is the dihedral group, the first of the series we saw before

$$D_n = \langle a, b | a^n = b^2 = 1, bab = a^{-1} \rangle \tag{16.1}$$

and has order $2n$. There is a special isomorphism $D_3 \cong S_3$ with the symmetric group on three letters.

The next nontrivial case is $|G| = 8$. Here we can invoke Sylow's theorem: If $p^k || G|$ then $G$ has a subgroup of order $p^k$. Let us apply this to 4 dividing $|G|$. Such a subgroup has index two and hence must be a normal subgroup, and hence fits in a sequence

$$1 \to N \to G \to \mathbb{Z}_2 \to 1 \tag{16.2}$$

Now, $N$ is of order 4 so we know that $N \cong \mathbb{Z}_2 \times \mathbb{Z}_2$ or $N \cong \mathbb{Z}_4$. If we have

$$1 \to \mathbb{Z}_4 \to G \to \mathbb{Z}_2 \to 1 \tag{16.3}$$

[280] $Spin(n)$ double covers the classical matrix group $SO(n)$.

then we have $\alpha : \mathbb{Z}_2 \to \mathrm{Aut}(\mathbb{Z}_4) \cong \mathbb{Z}_2$ and there are exactly two such homomorphisms. Moreover, for a fixed $\alpha$ there are two possibilities for the square $\tilde{\sigma}^2 \in \mathbb{Z}_4$ where $\tilde{\sigma}$ is a lift of the generator of $\mathbb{Z}_2$. Altogether this gives four possibilities:

♣Need to explain more here. ♣

$$1 \to \mathbb{Z}_4 \to \mathbb{Z}_2 \times \mathbb{Z}_4 \to \mathbb{Z}_2 \to 1 \tag{16.4}$$

$$1 \to \mathbb{Z}_4 \to \mathbb{Z}_8 \to \mathbb{Z}_2 \to 1 \tag{16.5}$$

$$1 \to \mathbb{Z}_4 \to D_4 \to \mathbb{Z}_2 \to 1 \tag{16.6}$$

$$1 \to \mathbb{Z}_4 \to \widetilde{D}_2 \to \mathbb{Z}_2 \to 1 \tag{16.7}$$

Here we meet the first of the series of *dicyclic* or *binary dihedral* groups defined by

$$\widetilde{D_n} := \langle a, b | a^{2n} = 1, a^n = b^2, b^{-1}ab = a^{-1} \rangle \tag{16.8}$$

It has order $4n$. There is a special isomorphism of $\widetilde{D}_2$ with the quaternion group.

The other possibility for $N$ is $\mathbb{Z}_2 \times \mathbb{Z}_2$ and here one new group is found, namely $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$.

Thus there are 5 inequivalent groups of order 8.

The next few cases are trivial until we get to $|G| = 12$. By Cauchy's theorem there are subgroups isomorphic to $\mathbb{Z}_2$, so we can view $G$ as an extension of $D_3$ or $\mathbb{Z}_6$ by $\mathbb{Z}_2$. There is also a subgroup isomorphic to $\mathbb{Z}_3$ so we can view it as an extension of an order 4 group by an order 3 group. We skip the analysis and just present the 5 distinct order 12 groups. In this way we find the groups forming the pattern at lower order:

♣Check this reasoning is correct. You need to know the subgroups are normal to say there is an extension. ♣

$$\mathbb{Z}_{12}, \quad , \mathbb{Z}_2 \times \mathbb{Z}_6, \quad , D_6, \quad , \tilde{D}_3 \tag{16.9}$$

And we find one "new" group: $A_4 \subset S_4$.

We can easily continue the table until we get to order $|G| = 16$. At order 16 there are 14 inequivalent groups! So we will stop here. [281]

[281] See, however, M. Wild, "Groups of order 16 made easy," American Mathematical Monthly, Jan 2005

| Order | Presentation | name |
|---|---|---|
| 1 | $\langle a \mid a = 1\rangle$ | Trivial group |
| 2 | $\langle a \mid a^2 = 1\rangle$ | Cyclic $\mathbb{Z}/2\mathbb{Z}$ |
| 3 | $\langle a \mid a^3 = 1\rangle$ | Cyclic $\mathbb{Z}/3\mathbb{Z}$ |
| 4 | $\langle a \mid a^4 = 1\rangle$ | Cyclic $\mathbb{Z}/4\mathbb{Z}$ |
| 4 | $\langle a, b \mid a^2 = b^2 = (ab)^2 = 1\rangle$ | Dihedral $D_2 \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, Klein |
| 5 | $\langle a \mid a^5 = 1\rangle$ | Cyclic $\mathbb{Z}/5\mathbb{Z}$ |
| 6 | $\langle a, b \mid a^3 = 1, b^2 = 1, bab = a\rangle$ | Cyclic $\mathbb{Z}/6\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ |
| 6 | $\langle a, b \mid a^3 = 1, b^2 = 1, bab = a^{-1}\rangle$ | Dihedral $D_3 \cong S_3$ |
| 7 | $\langle a \mid a^7 = 1\rangle$ | Cyclic $\mathbb{Z}/7\mathbb{Z}$ |
| 8 | $\langle a \mid a^8 = 1\rangle$ | Cyclic $\mathbb{Z}/8\mathbb{Z}$ |
| 8 | $\langle a, b \mid a^2 = 1, b^4 = 1, aba = b\rangle$ | $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$ |
| 8 | $\langle a, b, c \mid a^2 = b^2 = c^2 = 1, [a, b] = [a, c] = [b, c] = 1\rangle$ | $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ |
| 8 | $\langle a, b \mid a^4 = 1, b^2 = 1, bab = a^{-1}\rangle$ | Dihedral $D_4$ |
| 8 | $\langle a, b \mid a^4 = 1, a^2 = b^2, b^{-1}ab = a^{-1}\rangle$ | Dicyclic $\widetilde{D_2} \cong Q$, quaternion |
| 9 | $\langle a \mid a^9 = 1\rangle$ | Cyclic $\mathbb{Z}/9\mathbb{Z}$ |
| 9 | $\langle a, b \mid a^3 = b^3 = 1, [a, b] = 1\rangle$ | $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ |
| 10 | $\langle a \mid a^{10} = 1\rangle$ | Cyclic $\mathbb{Z}/10\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$ |
| 10 | $\langle a \mid a^5 = b^2 = 1, bab = a^{-1}\rangle$ | Dihedral $D_5$ |
| 11 | $\langle a \mid a^{11} = 1\rangle$ | Cyclic $\mathbb{Z}/11\mathbb{Z}$ |
| 12 | $\langle a \mid a^{12} = 1\rangle$ | Cyclic $\mathbb{Z}/12\mathbb{Z} \cong \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ |
| 12 | $\langle a, b \mid a^2 = 1, b^6 = 1, [a, b] = 1\rangle$ | $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ |
| 12 | $\langle a, b \mid a^6 = 1, b^2 = 1, bab = a^{-1}\rangle$ | Dihedral $D_6$ |
| 12 | $\langle a, b \mid a^6 = 1, a^3 = b^2, b^{-1}ab = a^{-1}\rangle$ | Dicyclic $\widetilde{D_3}$ |
| 12 | $\langle a, b \mid a^3 = 1, b^2 = 1, (ab)^3 = 1\rangle$ | Alternating $A_4$ |
| 13 | $\langle a \mid a^{13} = 1\rangle$ | Cyclic $\mathbb{Z}/13\mathbb{Z}$ |
| 14 | $\langle a \mid a^{14} = 1\rangle$ | Cyclic $\mathbb{Z}/14\mathbb{Z} \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z}$ |
| 14 | $\langle a, b \mid a^7 = 1, b^2 = 1, bab = a^{-1}\rangle$ | Dihedral $D_7$ |
| 15 | $\langle a \mid a^{15} = 1\rangle$ | Cyclic $\mathbb{Z}/15\mathbb{Z} \cong \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$ |

**Remarks**:

1. Explicit tabulation of the isomorphism classes of groups was initiated by by Otto Holder who completed a table for $|G| \leq 200$ about 100 years ago. Since then there has been much effort in extending those results. For surveys see

   1. J.A. Gallan, "The search for finite simple groups," Mathematics Magazine, vol. 49 (1976) p. 149. (This paper is a bit dated.)

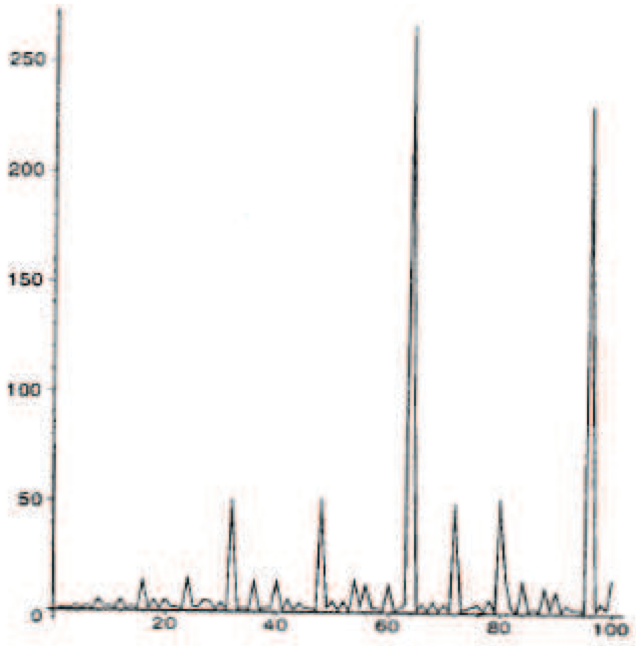



**Figure 45:** A plot of the number of nonisomorphic groups of order $n$. This plot was taken from the book by D. Joyner, *Adventures in Group Theory*.

2. H.U. Besche, B. Eick, E.A. O'Brian, "A millenium project: Constructing Groups of Small Order,"

2. There are also nice tables of groups of low order, in Joyner, *Adventures in Group Theory*, pp. 168-172, and Karpilovsky, *The Schur Multiplier* which go beyond the above table.

3. There are also online resources:
   1. http://www.gap-system.org/ for GAP
   2. http://hobbes.la.asu.edu/groups/groups.html for groups of low order.
   3. http://www.bluetulip.org/programs/finitegroups.html
   4. http://en.wikipedia.org/wiki/List-of-small-groups

4. The number of isomorphism types of groups jumps wildly. Apparently, there are $49,487,365,422$ isomorphism types of groups of order $2^{10} = 1024$. (Besche et. al. loc. cit.) The remarkable plot of Figure 45 from Joyner's book shows a plot of the number of isomorphism classes vs. order up to order 100. Figure 46 shows a log plot of the number of groups up to order 2000.

5. There is, however, a formula giving the asymptotics of the number $f(n,p)$ of isomorphism classes of groups of order $p^n$ for $n \to \infty$ for a fixed prime $p$. (Of course, there are $p(n)$ Abelian groups, where $p(n)$ the the number of partitions of $n$. Here we are

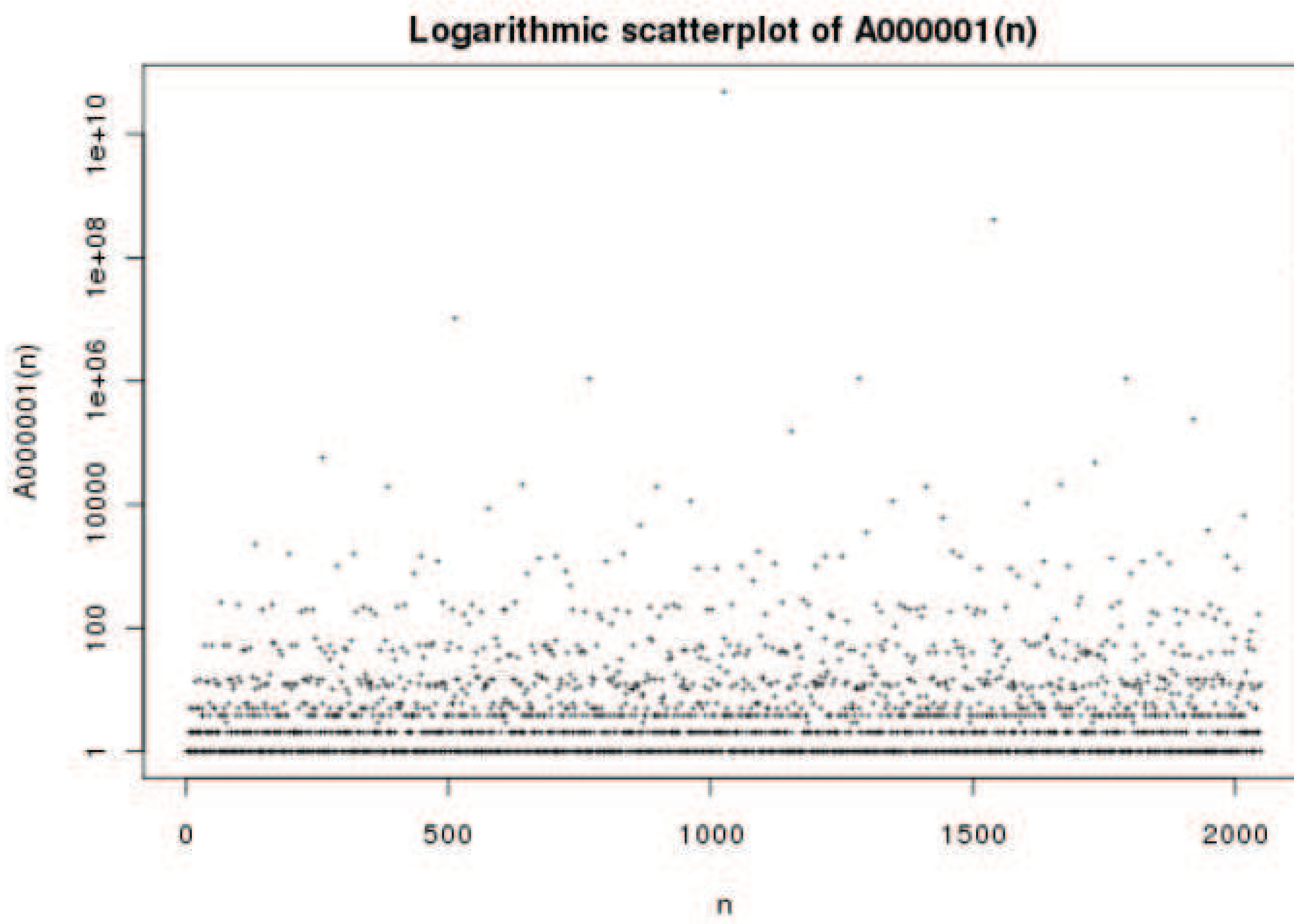



**Figure 46:** A logarithmic plot of the number of nonisomorphic groups of order $n$ out to $n \leq 2000$. This plot was taken from online encyclopedia of integer sequences, OEIS.

talking about the number of all groups.) This is due to G. Higman [282] and C. Sims [283] and the result states that:

$$f(n,p) \sim p^{\frac{2}{27}n^3} \tag{16.10}$$

Note that the asymptotics we derived for $p(n)$ before had a growth like $e^{const.n^{1/2}}$ so, unsurprisingly, most of the groups are nonabelian.

---

**Exercise** *Relating the binary dihedral and dihedral groups*

[282] G. Higman, "Enumerating p-Groups," Proc. London Math. Soc. 3) 10 (1960)

[283] C. Sims, "Enumerating p-Groups," Proc. London Math. Soc. (3) IS (1965) 151-66

Show that $\widetilde{D}_n$ is a double-cover of $D_n$ which fits into the exact sequence:

$$\begin{array}{ccccccccc}
 & & \mathbb{Z}_2 & = & \mathbb{Z}_2 & & & & \\
 & & \downarrow & & \downarrow & & & & \\
1 & \longrightarrow & \mathbb{Z}_{2n} & \longrightarrow & \widetilde{D}_n & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & 1 \\
 & & \| & & \downarrow & & \| & & \\
1 & \longrightarrow & \mathbb{Z}_n & \longrightarrow & D_n & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & 1
\end{array} \tag{16.11}$$

---

### 16.2 Finite Abelian Groups

The upper left box of our rough classification can be dealt with thoroughly, and the result is extremely beautiful.

In this subsection we will write our abelian groups *additively*.

Recall that we have shown that if $p$ and $q$ are positive integers then

$$0 \to \mathbb{Z}/gcd(p,q)\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/q\mathbb{Z} \to \mathbb{Z}/lcm(p,q)\mathbb{Z} \to 0 \tag{16.12}$$

and in particular, if $p, q$ are relatively prime then

$$\mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/q\mathbb{Z} \cong \mathbb{Z}/pq\mathbb{Z}. \tag{16.13}$$

It thus follows that if $n$ has prime decomposition

$$n = \prod_i p_i^{e_i} \tag{16.14}$$

then

$$\mathbb{Z}/n\mathbb{Z} \cong \oplus_i \mathbb{Z}/p_i^{e_i}\mathbb{Z} \tag{16.15}$$

This decomposition has a beautiful generalization to an arbitrary finite abelian group:

**Kronecker Structure Theorem**. Any finite abelian group is a direct product of cyclic groups of order a prime power. That is, we firstly have the decomposition:

$$\begin{aligned} G &= G_2 \oplus G_3 \oplus G_5 \oplus G_7 \oplus \cdots \\ &= \oplus_{p \text{ prime}} G_p \end{aligned} \tag{16.16}$$

where $G_p$ has order $p^n$ for some $n \geq 0$ ($n$ can depend on $p$, and for all but finitely many $p$, $G_p = \{0\}$.) And, secondly, each nonzero factor $G_p$ can be written:

$$G_p = \oplus_i \mathbb{Z}/(p^{n_i}\mathbb{Z}) \tag{16.17}$$

for some finite collection of positive integers $n_i$ (depending on $p$).

*Proof*: The proof proceeds in two parts. The first, easy, part shows that we can split $G$ into a direct sum of "$p$-groups" (defined below). The second, harder, part shows that an arbitrary abelian $p$-group is a direct sum of cyclic groups.

For part 1 of the proof let us consider an arbitrary finite abelian group $G$. We will write the group multiplication additively. Suppose $n$ is an integer so that $ng = 0$ for all $g \in G$. To fix ideas let us take $n = |G|$. Suppose $n = m_1 m_2$ where $m_1, m_2$ are relatively prime integers. Then there are integers $s_1, s_2$ so that

$$s_1 m_1 + s_2 m_2 = 1 \tag{16.18}$$

Therefore any element $g$ can be written as

$$g = s_1(m_1 g) + s_2(m_2 g) \tag{16.19}$$

Now $m_1 G$ and $m_2 G$ are subgroups and we claim that $m_1 G \cap m_2 G = \{0\}$. If $a \in m_1 G \cap m_2 G$ then $m_1 a = 0$ and $m_2 a = 0$ and hence (16.19) implies $a = 0$. Thus,

$$G = m_1 G \oplus m_2 G \tag{16.20}$$

Moreover, we claim that $m_1 G = \{g \in G | m_2 g = 0\}$. It is clear that every element in $m_1 G$ is killed by $m_2$. Suppose on the other hand that $m_2 g = 0$. Again applying (16.19) we see that $g = s_1 m_1 g = m_1(s_1 g) \in m_1 G$.

Thus, we can decompose

$$G = \oplus p \ \mathrm{prime} G_p \tag{16.21}$$

where $G_p$ is the subgroup of $G$ of elements whose order is a power of $p$.

If $p$ is a prime number then a *$p$-group* is a group all of whose elements have order a power of $p$. Now for part 2 of the proof we show that any abelian $p$-group is a direct sum of the form (16.17). The proof of this statement proceeds by induction and is based on a systematic application of Cauchy's theorem: If $p$ divides $|G|$ then there is an element of $G$ of order precisely $p$. (Recall we proved this theorem in Section 9.

Now, note that any $p$-group $G$ has an order which is a power $p^n$ for some $n$. If not, then $|G| = p^n m$ where $m$ is relatively prime to $p$. But then - by Cauchy's theorem - there would have to be an element of $G$ whose order is a prime divisor of $m$.

Next we claim that if an abelian $p$-group has a *unique* subgroup $H$ of order $p$ then $G$ itself is cyclic.

To prove this we again proceed by induction on $|G|$. Consider the subgroup defined by:

$$H = \{g | pg = 0\} \tag{16.22}$$

From Cauchy's theorem we see that $H$ cannot be the trivial group, and hence this must be the unique subgroup of order $p$. On the other hand, $H$ is manifestly the kernel of the homomorphism $\phi : G \to G$ given by $\phi(g) = pg$. Again by Cauchy, $\phi(G)$ has a subgroup of order $p$, but this must also be a subgroup of $G$, which contains $\phi(G)$, and hence $\phi(G)$ has a unique subgroup of order $p$. By the induction hypothesis, $\phi(G)$ is cyclic. But now $\phi(G) \cong G/H$, so let $g_0 + H$ be a generator of the cyclic group $G/H$. Next we claim that $H \subset \langle g_0 \rangle$. Since $G$ is a $p$-group the subgroup $\langle g_0 \rangle$ is a $p$-group and hence contains a subgroup of order $p$ (by Cauchy) but (by hypothesis) there is a unique such subgroup in $G$ and any subgroup of $\langle g_0 \rangle$ is a subgroup of $G$, so $H \subset \langle g_0 \rangle$. But now take any element $g \in G$. On the one hand it must project to an element $[g] \in G/H$. Thus must be of the form $[g] = kg_0 + H$, since $g_0 + H$ generates $G/H$. That means $g = kg_0 + h$, $h \in H$, but since $H \subset \langle g_0 \rangle$ we must have $h = \ell g_0$ for some integer $\ell$. Therefore $G = \langle g_0 \rangle$ is cyclic.

The final step proceeds by showing that if $G$ is a finite abelian $p$-group and $M$ is a cyclic subgroup of maximal order then $G = M \oplus N$ for some subgroup $N$. Once we have established this the desired result follows by induction.

So, now suppose that that $G$ has a cyclic subgroup of maximal order $M$. If $G$ is cyclic then $N = \{0\}$. If $G$ is not cyclic then we just proved that there must be at least two distinct subgroups of order $p$. One of them is in $M$. Choose another one, say $K$. Note that $K$ must not be in $M$, because $M$ is cyclic and has a unique subgroup of order $p$. Therefore $K \cap M = \{0\}$. Therefore $(M+K)/K \cong M$. Therefore $(M+K)/K$ is a cyclic subgroup of $G/K$. Any element $g + K$ has an order which divides $|g|$, and $|g| \le |M|$ since $M$ is a maximal cyclic subgroup. Therefore the cyclic subgroup $(M+K)/K$ is a maximal order cyclic subgroup of $G/K$. Now the inductive hypothesis implies $G/K = (M+K)/K \oplus H/K$ for some subgroup $K \subset H \subset G$. But this means $(M+K) \cap H = K$ and hence $M \cap H = \{0\}$ and hence $G = M \oplus H$. ♠

For other proofs see

1. S. Lang, *Algebra*, ch. 1, sec. 10.
2. I.N. Herstein, Ch. 2, sec. 14.
3. J. Stillwell, *Classical Topology and Combinatorial Group Theory.*
4. Our proof is based on G. Navarro, "On the fundamental theorem of finite abelian groups," Amer. Math. Monthly, Feb. 2003, vol. 110, p. 153.

One class of examples where we have a finite Abelian group, but it's Kronecker decomposition is far from obvious is the following: Consider the Abelian group $\mathbb{Z}^d$. Choose a set of $d$ vectors $v_i \in \mathbb{Z}^d$, linearly independent as vectors in $\mathbb{R}^d$.

$$L := \{\sum_{i=1}^{d} n_i v_i | n_i \in \mathbb{Z}\} \tag{16.23}$$

is a subgroup. Then

$$A = \mathbb{Z}^d / L \tag{16.24}$$

is a finite Abelian group. For example if $v_i = ke_i$ where $e_i$ is the standard unit vector in the $i^{th}$ direction then obviously $A \cong (\mathbb{Z}/k\mathbb{Z})^d$. But for a general set of vectors the decomposition is not obvious.

So, here is an algorithm for giving the Kronecker decomposition of a finite Abelian group:

1. Compute the orders of the various elements.

2. You need only consider the elements whose order is a prime power. (By the Bezout manipulation all the others will be sums of these.)

3. Focusing on one prime at a time. Take the element $g_1$ whose order is maximal. Then $G_p = \langle g_1 \rangle \oplus N$. ♣Have to say how to get $N$. ♣

4. Repeat for $N$.

---

**Exercise**

Show that an alternative of the structure theorem is the statement than any finite abelian group is isomorphic to

$$\mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_k} \tag{16.25}$$

where

$$n_1 | n_2 \quad \& \quad n_2 | n_3 \quad \& \quad \cdots \quad \& \quad n_{k-1} | n_k \tag{16.26}$$

Write the $n_i$ in terms of the prime powers in (16.17).

---

---

**Exercise** *p-groups*

a.) Show that $\mathbb{Z}_4$ is not isomorphic to $\mathbb{Z}_2 \oplus \mathbb{Z}_2$.

b.) Show more generally that if $p$ is prime $\mathbb{Z}_{p^n}$ and $\mathbb{Z}_{p^{n-m}} \oplus \mathbb{Z}_{p^m}$ are not isomorphic if $0 < m < n$.

c.) How many nonisomorphic abelian groups have order $p^n$?

---

---

**Exercise**

Suppose $e_1, e_2 \in \mathbb{Z}^2$ are two linearly independent vectors (over $\mathbb{Q}$). Let $\Lambda = \langle e_1, e_2 \rangle \subset \mathbb{Z}^2$ be the sublattice generated by these vectors. Then $\mathbb{Z}^2/\Lambda$ is a finite abelian group. Compute its Kronecker decomposition in terms of the coordinates of $e_1, e_2$.

---

### 16.3 Finitely Generated Abelian Groups

It is hopeless to classify all infinite abelian groups, but a "good" criterion that leads to an interesting classification is that of *finitely generated* abelian groups.

Any abelian group has a canonically defined subgroup known as the *torsion subgroup*, and denoted Tors$(G)$. This is the subgoup of elements of *finite order*:

$$\mathrm{Tors}(G) := \{g \in G | \exists n \in \mathbb{Z} \qquad ng = 0\} \tag{16.27}$$

where we are writing the group $G$ additively, so $ng = g + \cdots + g$.

One can show that any *finitely generated abelian group* fits in an exact sequence

$$0 \to \mathrm{Tors}(G) \to A \to \mathbb{Z}^r \to 0 \tag{16.28}$$

where $\mathrm{Tors}(G)$ is a *finite abelian group*.

For a proof, see, e.g., S. Lang, *Algebra* .

Moreover (16.28) is a split extension, that is, it is isomorphic to

$$\mathbb{Z}^r \oplus \mathrm{Tors}(G) \tag{16.29}$$

The integer $r$, called the *rank of the group*, and the finite abelian group $\mathrm{Tors}(G)$ are invariants of the finitely generated abelian group. Since we have a general picture of the finite abelian groups we have now got a general picture of the finitely generated abelian groups.

**Remark**:

**Remarks**

1. The groups $\mathbb{C}, \mathbb{R}, \mathbb{Q}$ under addition are abelian but not finitely generated. This is obvious for $\mathbb{C}$ and $\mathbb{R}$ since these are uncountable sets. To see that $\mathbb{Q}$ is not finitely generated consider any finite set of fractions $\{\frac{p_1}{q_1}, \dots, \frac{p_s}{q_s}\}$. This set will will only generate rational numbers which, when written in lowest terms, have denominator at most $q_1 q_2 \cdots q_s$.

2. Note that a torsion abelian group need not be finite in general. For example $\mathbb{Q}/\mathbb{Z}$ is entirely torsion, but is not finite.

3. A rich source of finitely generated abelian groups are the integral cohomology groups $H^n(X; \mathbb{Z})$ of smooth compact manifolds.

4. We must stress that the presentation (16.29) of a finitely generated abelian group is <u>not canonical</u>! There are many distinct splittings of (16.28). They are in 1-1 correspondence with the group homomorphisms $\mathrm{Hom}(\mathbb{Z}^r, \mathrm{Tors}(G))$. For a simple example consider $\mathbb{Z}^d/\Lambda$ where $\Lambda$ is a general sublattice of rank less than $d$.

5. In a nonabelian group the product of two finite-order elements can very well have infinite order. Examples include free products of cyclic groups and simple rotations by $2\pi/n$ around different axes in $SO(3)$. So, there is no straightforward generalization of $\mathrm{Tors}(G)$ to the case of nonabelian groups.

---

**Exercise**

Consider the finitely generated Abelian group [284]

$$L = \{(x_1, x_2, x_3, x_4) \in \mathbb{Z}^4 | \sum_i x_i = 0\text{mod}2\} \tag{16.30}$$

and consider the subgroup $S$ generated by

$$\begin{aligned} v_1 &= (1,1,1,1) \\ v_2 &= (1,1,-1,-1) \end{aligned} \tag{16.31}$$

a.) What is the torsion group of $L/S$ ?

b.) Find a splitting of the sequence (16.28) and compare with the one found by other students in the course. Are they the same?

---

---

**Exercise**

Given a set of finite generators of an Abelian group $A$ try to find an algorithm for a splitting of the sequence (16.28).

---

### 16.4 The Classification Of Finite Simple Groups

Kronecker's structure theorem is a very satisfying, beautiful and elegant answer to a classification question. The generalization to nonabelian groups is very hard. It turns out that a "good" criterion is that a finite group be a *simple* group. This idea arose from the Galois demonstration of (non)solvability of polynomial equations by radicals.

A key concept in abstract group theory is provided by the notion of a *composition series*. This is a sequence of subgroups

$$1 = G_{s+1} \triangleleft G_s \triangleleft \cdots \triangleleft G_2 \triangleleft G_1 = G \tag{16.32}$$

which have the property that $G_{i+1}$ is a maximal normal subgroup of $G_i$. (Note: $G_{i+1}$ need not be normal in $G$. Moreover, there might be more than one maximal normal subgroup in $G_i$. ) As a simple example we shall see that we have

♣should give an example of this.... ♣

$$1 = G_4 \triangleleft G_3 = \mathbb{Z}_2 \times \mathbb{Z}_2 \triangleleft G_2 = A_4 \triangleleft G_1 = S_4 \tag{16.33}$$

but

$$G_3 = 1 \triangleleft G_2 = A_n \triangleleft G_1 = S_n \qquad\qquad n \geq 5 \tag{16.34}$$

Not every group admits a composition series. For example $G = \mathbb{Z}$ does not admit a composition series. (Explain why!) However, it can be shown that every <u>finite</u> group admits a composition series.

♣Give a reference. ♣

[284] It is the root lattice of $\mathfrak{so}(8)$.

It follows that in a composition series the quotient groups $G_i/G_{i+1}$ are *simple groups*: By definition, a simple group is one whose only normal subgroups are 1 and itself. From what we have learned above, that means that a simple group has no nontrivial homomorphic images. It also implies that the center is trivial or the whole group.

Let us prove that the $G_i/G_{i+1}$ are simple: In general, if $N \triangleleft G$ is a normal subgroup then there is a 1-1 correspondence:

*Subgroups H between N and G:* $N \subset H \subset G \Leftrightarrow$ *Subgroups of* $G/N$

Moreover, under this correspondence:

*Normal subgroups of* $G/N \Leftrightarrow$ *Normal subgroups* $N \subset H \triangleleft G$. If $H/G_{i+1} \subset G_i/G_{i+1}$ were normal and $\neq 1$ then $G_{i+1} \subset H \subset G_i$ would be normal and and properly contain $G_{i+1}$, contradicting maximality of $G_{i+1}$. ♠

♣Make this an exercise in an earlier section. ♣

A composition series is a nonabelian generalization of the Kronecker decomposition. It is not unique (see exercise below) but the the following theorem, known as the Jordan-Hölder theorem states that there are some invariant aspects of the decomposition:

**Theorem**: Suppose there are two different composition series for $G$:

$$1 = G_{s+1} \triangleleft G_s \triangleleft \cdots \triangleleft G_2 \triangleleft G_1 = G \tag{16.35}$$

$$1 = G'_{s'+1} \triangleleft G'_s \triangleleft \cdots \triangleleft G'_2 \triangleleft G'_1 = G \tag{16.36}$$

Then $s = s'$ and there is a permutation $i \to i'$ so that $G_i/G_{i+1} \cong G'_{i'}/G'_{i'+1}$. That is: The length and the unordered set of quotients are both invariants of the group and do not depend on the particular composition series.

For a proof see Jacobsen, Section 4.6.

The classification of all finite groups is reduced to solving the extension problem in general, and then classifying finite simple groups. The idea is that if we know $G_i/G_{i+1} = S_i$ is a finite simple group then we construct $G_i$ from $G_{i+1}$ and the extension:

$$1 \to G_{i+1} \to G_i \to S_i \to 1 \tag{16.37}$$

We have discussed the extension problem thoroughly above. One of the great achievements of 20th century mathematics is the complete classification of finite simple groups, so let us look at the finite simple groups:

First consider the abelian ones. These cannot have nontrivial subgroups and hence must be of the form $\mathbb{Z}/p\mathbb{Z}$ where $p$ is prime.

So, now we search for the nonabelian finite simple groups. A natural source of nonabelian groups are the symmetric groups $S_n$. Of course, these are not simple because $A_n \subset S_n$ are normal subgroups. Could the $A_n$ be simple? The first nonabelian example is $A_4$ and it is not a simple group! Indeed, consider the cycle structures $(2)^2$. There are three nontrivial elements: (12)(34), (13)(24), and (14)(23), they are all involutions, and

$$((12)(34)) \cdot ((13)(24)) = ((13)(24)) \cdot ((12)(34)) = (14)(23) \tag{16.38}$$

and therefore together with the identity they form a subgroup $K \subset A_4$ isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$. Since cycle-structure is preserved under conjugation, this is obviously a normal subgroup of $A_4$!. After this unpromising beginning you might be surprised to learn:

**Theorem** $A_n$ is a simple group for $n \geq 5$.

*Sketch of the proof*:

We first observe that $A_n$ is generated by cycles of length three: $(abc)$. The reason is that $(abc) = (ab)(bc)$, so any word in an even number of distinct transpositions can be rearranged into a word made from a product of cycles of length three. Therefore, the strategy is to show that any normal subgroup $K \subset A_n$ which is larger than 1 must contain at least one three-cycle $(abc)$. WLOG let us say it is $(123)$. Now we claim that the entire conjugacy class of three-cycles must be in $K$. We consider a permutation $\phi$ which takes

$$\phi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \cdots \\ i & j & k & l & m & \cdots \end{pmatrix} \tag{16.39}$$

Then $\phi(123)\phi^{-1} = (ijk)$. If $\phi \in A_n$ we are done, since $K$ is normal in $A_n$ so then $(ijk) \in K$. If $\phi$ is an odd permutation then $\tilde{\phi} = \phi(45)$ is even and $\tilde{\phi}(123)\tilde{\phi}^{-1} = (ijk)$.

Thus, we need only show that some 3-cycle is in $K$. For $n = 5$ this can be done rather explicitly. See the exercise below. Once we have established that $A_5$ is simple we can proceed by induction as follows.

We first establish a lemma: If $n \geq 5$ then for any $\sigma \in A_n$, $\sigma \neq 1$ there is a conjugate element (in $A_n$) $\sigma'$ with $\sigma' \neq \sigma$ such that there is an $i \in \{1, \dots, n\}$ so that $\sigma(i) = \sigma'(i)$.

To prove the lemma choose any $\sigma \neq 1$ and for $\sigma$ choose a cycle of maximal length, say $r$ so that $\sigma = (12\dots r)\pi$ with $\pi$ fixing $\{1, \dots, r\}$. If $r \geq 3$ then consider the conjugate:

$$\sigma' = (345)\sigma(345)^{-1} = (345)(123\cdots)\pi(354) \tag{16.40}$$

We see that $\sigma(1) = \sigma'(1) = 2$, while $\sigma(2) = 3$ and $\sigma'(2) = 4$. We leave the case $r = 2$ to the reader.

Now we proceed by induction: Suppose $A_j$ is simple for $5 \leq j \leq n$. Consider $A_{n+1}$ and let $N \triangleleft A_{n+1}$. Then choose $\sigma \in N$ and using the lemma consider $\sigma' \in A_{n+1}$ with $\sigma' \neq \sigma$ and $\sigma'(i) = \sigma(i)$ for some $i$. Let $H_i \subset A_{n+1}$ be the subgroup of permutations fixing $i$. It is isomorphic to $A_n$. Now, $\sigma' \in N$ since it is a conjugate of $\sigma \in N$ and $N$ is assumed to be normal. Therefore $\sigma^{-1}\sigma' \in N$, and $\sigma^{-1}\sigma' \neq 1$. Therefore $N \cap H_i \neq 1$. But $N \cap H_i$ must be normal in $H_i$. Since $H_i \cong A_n$ it follows that $N \cap H_i = H_i$. But $H_i$ contains 3-cycles. Therefore $N$ contains 3-cycles and hence $N \cong A_{n+1}$. ♠

**Remark**: For several other proofs of the same theorem and other interesting related facts see

http://www.math.uconn.edu/kconrad/blurbs/grouptheory/Ansimple.pdf.

**Digressive Remark**: A group is called *solvable* if the $G_i/G_{i+1}$ are abelian (and hence $\mathbb{Z}/p\mathbb{Z}$ for some prime $p$). The term has its origin in Galois theory, which in turn was the original genesis of group theory. Briefly, in Galois theory one considers a polynomial $P(x)$ with coefficients drawn from a field $F$. (e.g. consider $F = \mathbb{Q}$ or $\mathbb{R}$). Then the roots of the polynomial $\theta_i$ can be adjoined to $F$ to produce a bigger field $K = F[\theta_i]$. The *Galois group of the polynomial* $Gal(P)$ is the group of automorphisms of $K$ fixing $F$. Galois theory sets up a beautiful 1-1 correspondence between subgroups $H \subset Gal(P)$ and subfields $F \subset K_H \subset K$. The intuitive notion of solving a polynomial by radicals corresponds to finding a series of subfields $F \subset F_1 \subset F_2 \subset \cdots \subset K$ so that $F_{i+1}$ is obtained from $F_i$ by adjoining the solutions of an equation $y^d = z$. Under the Galois correspondence this translates into a composition series where $Gal(P)$ is a solvable group - hence the name. If we take $F = \mathbb{C}[a_0, \dots, a_{n-1}]$ for an $n^{th}$ order polynomial

$$P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \tag{16.41}$$

then the roots $\theta_i$ are such that $a_j$ are the $j^{th}$ elementary symmetric polynomials in the $\theta_i$ (See Chapter 2 below). The Galois group is then $S_n$. For $n \geq 5$ the only nontrivial normal subgroup of $S_n$ is $A_n$, and this group is simple, hence certainly not solvable. That is why there is no solution of an $n^{th}$ order polynomial equation in radicals for $n \geq 5$.

Returning to our main theme, we ask: What other finite simple groups are there? The full list is known. The list is absolutely fascinating: [285]

1. $\mathbb{Z}/p\mathbb{Z}$ for $p$ prime.

2. The subgroup $A_n \subset S_n$ for $n \geq 5$.

3. "Simple Lie groups over finite fields."

4. 26 "sporadic oddballs"

We won't explain example 3 in great detail, but it consists of a few more infinite sequences of groups, like 1,2 above. To get a flavor of what is involved note the following: The additive group $\mathbb{Z}/p\mathbb{Z}$ where $p$ is prime has more structure: One can multiply elements, and if an element is nonzero then it has a multiplicative inverse, in other words, it is a *finite field*. One can therefore consider the group of invertible matrices over this field $GL(n,p)$, and its subgroup $SL(n,p)$ of matrices of unit determinant. Since $\mathbb{Z}/p\mathbb{Z}$ has a finite number of elements it is a finite group. This group is not simple, because it has a nontrivial center, in general. For example, if $n$ is even then the group $\{\pm 1\}$ is a normal subgroup isomorphic to $\mathbb{Z}_2$. If we divide by the center the we get a group $PSL(n,p)$ which, it turns out, is indeed a simple group. This construction can be generalized in a few directions. First, there is a natural generalization of $\mathbb{Z}/p\mathbb{Z}$ to finite fields $\mathbb{F}_q$ of order a prime power $q = p^k$. Then we can similarly define $PSL(n,q)$ and it turns out these are simple groups except for some low values of $n, q$. Just as the Lie groups $SL(n,\mathbb{C})$ have counterparts $O(n), Sp(n)$ etc. one can generalize this construction to groups of type $B, C, D, E$. This construction can be used to obtain the third class of finite simple groups.

♣Double check. Does this figure leave out a subgroup relation? ♣

[285] See the *Atlas of Finite Simple Groups*, by Conway and Norton

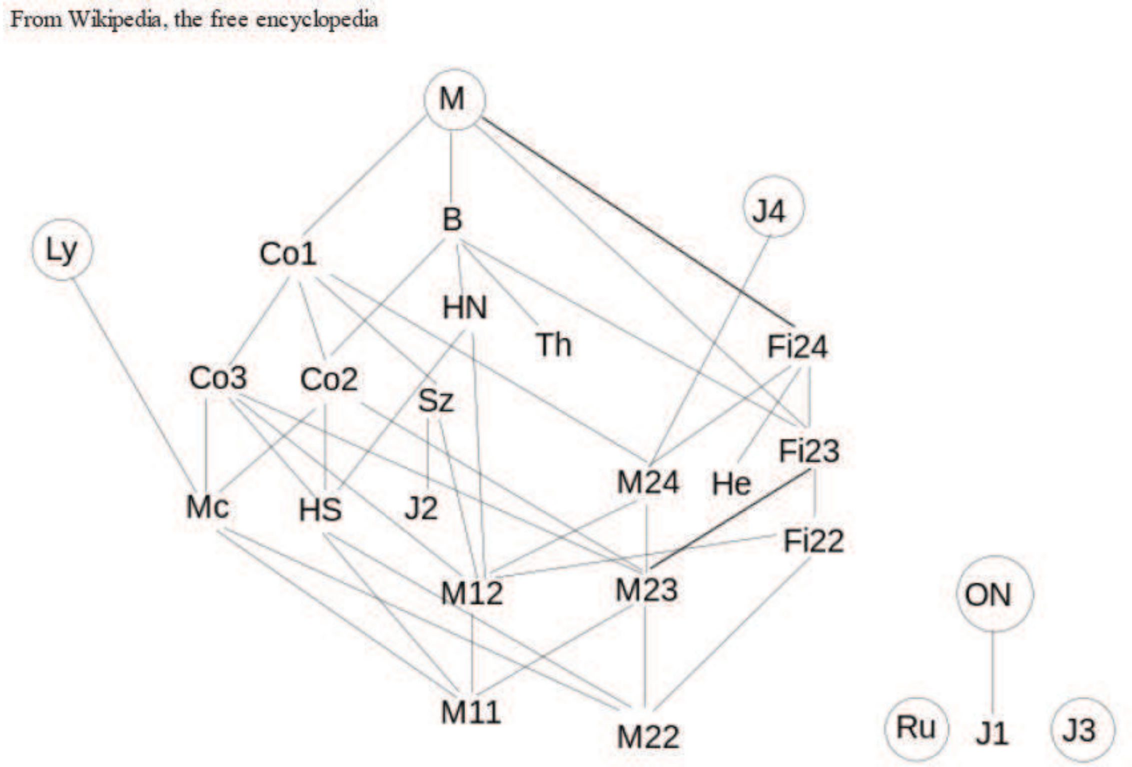



**Figure 47:** A table of the sporadic groups including subgroup relations. Source: Wikipedia.

It turns out that there are exactly 26 oddballs, known as the "sporadic groups." Some relationships between them are illustrated in Figure 47. The sporadic groups first showed up in the $19^{th}$ century via the Mathieu groups

$$M_{11}, M_{12}, M_{22}, M_{23}, M_{24}. \tag{16.42}$$

$M_n$ is a subgroup of the symmetric group $S_n$. $M_{11}$, which has order $|M_{11}| = 7920$ was discovered in 1861. We met $M_{12}$ when discussing card-shuffling. The last group $M_{24}$, with order $\sim 10^9$ was discovered in 1873. All these groups may be understood as automorphisms of certain combinatorial objects called "Steiner systems."

It was a great surprise when Janko constructed a new sporadic group $J_1$ of order $175,560$ in 1965, roughly 100 years after the discovery of the Mathieu groups. The list of sporadic groups is now thought to be complete. The largest sporadic group is called the Monster group and its order is:

$$\begin{aligned} |Monster| &= 2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \\ &= 808017424794512875886459904961710757005754368000000000 \\ &\cong 8.08 \times 10^{53} \end{aligned} \tag{16.43}$$

but it has only 194 conjugacy classes! (Thus, by the class equation, it is "very" nonabelian. The center is trivial and $Z(g)$ tends to be a small order group.)

The history and status of the classification of finite simple groups is somewhat curious: [286]

1. The problem was first proposed by Hölder in 1892. Intense work on the classification begins during the 20th century.

2. Feit and Thompson show (1963) that any finite group of odd order is solvable. In particular, it cannot be a simple group.

3. Janko discovers (1965) the first new sporadic group in almost a century.

4. Progress is then rapid and in 1972 Daniel Gorenstein (of Rutgers University) announces a detailed outline of a program to classify finite simple groups.

5. The largest sporadic group, the Monster, was first shown to exist in 1980 by Fischer and Griess. It was explicitly constructed (as opposed to just being shown to exist) by Griess in 1982.

6. The proof is completed in 2004. It uses papers from hundreds of mathematicians between 1955 and 2004, and largely follows Gorenstein's program. The proof entails tens of thousands of pages. Errors and gaps have been found, but so far they are just "local."

Compared to the simple and elegant proof of the classification of simple Lie algebras (to be covered in Chapter **** below) the proof is obviously terribly unwieldy.

It is conceivable that physics might actually shed some light on this problem. The simple groups are probably best understood as automorphism groups of some mathematical, perhaps even geometrical object. For example, the first nonabelian simple group, $A_5$ is the group of symmetries of the icosahedron, as we will discuss in detail below. A construction of the monster along these lines was indeed provided by Frenkel, Lepowsky, Meurman, (at Rutgers) using vertex operator algebras, which are important in the description of perturbative string theory. More recently the mystery has deepened with interesting experimental discoveries linking the largest Mathieu group $M_{24}$ to nonlinear sigma models with K3 target spaces. For more discussion about the possible role of physics in this subject see:

1. Articles by Griess and Frenkel et. al. in *Vertex Operators in Mathematics and Physics*, J. Lepowsky, S. Mandelstam, and I.M. Singer, eds.

2. J. Harvey, "Twisting the Heterotic String," in *Unified String Theories*, Green and Gross eds.

[286] Our source here is the Wikipedia article on the classification of finite simple groups. See also: Solomon, Ronald, "A brief history of the classification of the Finite simple groups," American Mathematical Society. Bulletin. New Series, 38 (3): 315-352 (2001).

3. L.J. Dixon, P.H. Ginsparg, and J.A. Harvey, "Beauty And The Beast: Superconformal Symmetry In A Monster Module," Commun.Math.Phys. 119 (1988) 221-241

4. M.C.N. Cheng, J.F.R. Duncan, and J.A. Harvey, "Umbral Moonshine," e-Print: arXiv:1204.2779 [math.RT]

---

**Exercise** *Completing the proof that $A_5$ is simple*

Show that any nontrivial normal subgroup of $A_5$ must contain a 3-cycle as follows:

a.) If $N \triangleleft A_5$ is a normal subgroup containing no 3-cycles then the elements must have cycle type $(ab)(cd)$ or $(abcde)$.

b.) Compute the group commutators ($a, b, c, d, e$ are all distinct):

$$[(abe), (ab)(cd)] = (aeb) \tag{16.44}$$

$$[(abc), (abcde)] = (abd) \tag{16.45}$$

c.) Use these facts to conclude that $N$ must contain a 3-cycle.

Legend has it that Galois discovered this theorem on the night before his fatal duel.

---

---

**Exercise** *Conjugacy classes in $A_n$*

Note that conjugacy classes in $A_n$ are different from conjugacy classes in $S_n$. For example, (123) and (132) are not conjugate in $A_3$.

Describe the conjugacy classes in $A_n$.

---

---

**Exercise** *Jordan-Hölder decomposition*

Work out JH decompositions for the order 8 quaternion group $\widetilde{D}_2$ and observe that there are several maximal normal subgroups.

---

---

**Exercise** *The simplest of the Chevalley groups*

a.) Verify that $SL(2, \mathbb{Z}/p\mathbb{Z})$ is a group.

b.) Show that the order of $SL(2, \mathbb{Z}/p\mathbb{Z})$ is $p(p^2 - 1)$. [287]

[287] Break up the cases into $d = 0$ and $d \neq 0$. When $d = 0$ you can solve $ad - bc = 1$ for $a$. When $d = 0$ you can have arbitrary $a$ but you must have $bc = -1$.

c.) Note that the scalar multiples of the $2 \times 2$ identity matrix form a normal subgroup of $SL(2, \mathbb{Z}/p\mathbb{Z})$. Show that the number of such matrices is the number of solutions of $x^2 = 1 \text{mod} p$. Dividing by this normal subgroup produces the group $PSL(2, \mathbb{Z}/p\mathbb{Z})$. Jordan proved that these are simple groups for $p \neq 2, 3$.

It turns out that $PSL(2, \mathbb{Z}_5) \cong A_5$. (Check that the orders match.) Therefore the next simple group in the series is $PSL(2, \mathbb{Z}_7)$. It has many magical properties.

d.) Show that $PSL(2, \mathbb{Z}_7)$ has order 168.

---

## 17. Categories: Groups and Groupoids

A rather abstract notion, which nevertheless has found recent application in string theory and conformal field theory is the language of categories. Many physicists object to the high level of abstraction entailed in the category language. Some mathematicians even refer to the subject as "abstract nonsense." (Others take it very seriously.) However, it seems to be of increasing utility in the further formal development of string theory and supersymmetric gauge theory. It is also essential for reading any of the literature on topological field theory.

We briefly illustrate some of that language here. Our main point here is to introduce a different viewpoint on what groups are that leads to a significant generalization: groupoids. Moreover, this point of view also provides some very interesting insight into the meaning of group cohomology. Related constructions have been popular in condensed matter physics and topological field theory.

**Definition** A *category* $\mathcal{C}$ consists of

a.) A set $Ob(\mathcal{C})$ of "objects"

b.) A collection $Mor(\mathcal{C})$ of sets $\text{hom}(X, Y)$, defined for any two objects $X, Y \in Ob(\mathcal{C})$. The elements of $\text{hom}(X, Y)$ are called the "morphisms from $X$ to $Y$." They are often denoted as arrows:

$$X \overset{\phi}{\rightarrow} Y \tag{17.1}$$

c.) A composition law:

$$\text{hom}(X, Y) \times \text{hom}(Y, Z) \rightarrow \text{hom}(X, Z) \tag{17.2}$$

$$(\psi_1, \psi_2) \mapsto \psi_2 \circ \psi_1 \tag{17.3}$$

Such that

1. A morphism $\phi$ uniquely determines its source $X$ and target $Y$. That is, $\text{hom}(X, Y)$ are disjoint for distinct ordered pairs $(X, Y)$.

2. $\forall X \in Ob(\mathcal{C})$ there is a distinguished morphism, denoted $1_X \in \text{hom}(X, X)$ or $\text{Id}_X \in \text{hom}(X, X)$, which satisfies:

$$1_X \circ \phi = \phi \qquad \psi \circ 1_X = \psi \tag{17.4}$$

for all morphisms $\phi \in \text{hom}(Y, X)$ and $\psi \in \text{hom}(X, Y)$ for all $Y \in Ob(\mathcal{C})$. [288]

[288] As an exercise, show that these conditions uniquely determine the morphism $1_X$.

3. Composition of morphisms is associative:

$$(\psi_1 \circ \psi_2) \circ \psi_3 = \psi_1 \circ (\psi_2 \circ \psi_3) \tag{17.5}$$

An alternative definition one sometimes finds is that a category is defined by two sets $\mathcal{C}_0$ (the objects) and $\mathcal{C}_1$ (the morphisms) with two maps $p_0 : \mathcal{C}_1 \to \mathcal{C}_0$ and $p_1 : \mathcal{C}_1 \to \mathcal{C}_0$. The map $p_0(f) = x_1 \in \mathcal{C}_0$ is the *range* map and $p_1(f) = x_0 \in \mathcal{C}_0$ is the *domain* map. In this alternative definition a category is then defined by a composition law on the set of *composable morphisms*

$$\mathcal{C}_2 = \{(f,g) \in \mathcal{C}_1 \times \mathcal{C}_1 | p_0(f) = p_1(g)\} \tag{17.6}$$

which is sometimes denoted $\mathcal{C}_{1p_1} \times_{p_0} \mathcal{C}_1$ and called the *fiber product*. The composition law takes $\mathcal{C}_2 \to \mathcal{C}_1$ and may be pictured as the composition of arrows. If $f : x_0 \to x_1$ and $g : x_1 \to x_2$ then the composed arrow will be denoted $g \circ f : x_0 \to x_2$. The composition law satisfies the axioms

1. For all $x \in X_0$ there is an identity morphism in $X_1$, denoted $1_x$, or $Id_x$, such that $1_x f = f$ and $g 1_x = g$ for all suitably composable morphisms $f, g$.
2. The composition law is associative. If $f, g, h$ are 3-composable morphisms then $(hg)f = h(gf)$.

**Remarks**:

1. When defining composition of arrows one needs to make an important notational decision. If $f : x_0 \to x_1$ and $g : x_1 \to x_2$ then the composed arrow is an arrow $x_0 \to x_2$. We will write $g \circ f$ when we want to think of $f, g$ as functions and $fg$ when we think of them as arrows.

♣Is this dual notation really a good idea?? ♣

2. It is possible to endow the data $X_0, X_1$ and $p_0, p_1$ with additional structures, such as topologies, and demand that $p_0, p_1$ have continuity or other properties.
3. A morphism $\phi \in \hom(X, Y)$ is said to be *invertible* if there is a morphism $\psi \in \hom(Y, X)$ such that $\psi \circ \phi = 1_X$ and $\phi \circ \psi = 1_Y$. If $X$ and $Y$ are objects with an invertible morphism between them then they are called *isomorphic objects*. One key reason to use the language of categories is that objects can have nontrivial automorphisms. That is, $\hom(X, X)$ can have invertible elements other than just $1_X$ in it. When this is true then it is tricky to speak of "equality" of objects, and the language of categories becomes very helpful. As a concrete example you might ponder the following question: "Are all real vector spaces of dimension $n$ *the same*?"

Here are some simple examples of categories:

1. **SET**: The category of sets and maps of sets. [289]

[289] We take an appropriate collection of sets and maps to avoid the annoying paradoxes of set theory.

2. **TOP**: The category of topological spaces and continuous maps.

3. **TOPH**: The category of topological spaces and homotopy classes of continuous maps.

4. **MANIFOLD**: The category of manifolds and suitable maps. We could take topological manifolds and continuous maps of manifolds. Or we could take smooth manifolds and smooth maps as morphisms. The two choices lead to two (very different!) categories.

5. **BORD**$(n)$: The bordism category of $n$-dimensional manifolds. Roughly speaking, the objects are $n$-dimensional manifolds without boundary and the morphisms are bordisms. A bordism $Y$ from an $n$-manifold $M_1$ to and $n$-manifold $M_2$ is an $(n+1)$-dimensional manifold with a decomposition of its boundary $\partial Y = (\partial Y)_{in} \amalg (\partial Y)_{out}$ together with diffeomorphisms $\theta_1 : (\partial Y)_{in} \to M_1$ and $\theta_2 : (\partial Y)_{out} \to M_2$.

6. **GROUP**: the category of groups and homomorphisms of groups. Note that here if we took our morphisms to be isomorphisms instead of homomorphisms then we would get a very different category. All the pairs of objects in the category with nontrivial morphism spaces between them would be pairs of isomorphic groups.

7. **AB**: The (sub) category of abelian groups.

8. Fix a group $G$ and let **G-SET** be the category of $G$-sets, that is, sets $X$ with a $G$-action. For simplicity let us just write the $G$-action $\Phi(g,x)$ as $g \cdot x$ for $x$ a point in a $G$-set $X$. We take the morphisms $f : X_1 \to X_2$ to satisfy satisfy $f(g \cdot x_1) = g \cdot f(x_1)$.

9. **VECT**$_\kappa$: The category of finite-dimensional vector spaces over a field $\kappa$ with morphisms the linear transformations.

One use of categories is that they provide a language for describing precisely notions of "similar structures" in different mathematical contexts. When discussed in this way it is important to introduce the notion of "functors" and "natural transformations" to speak of interesting relationships between categories.

In order to state a relation between categories one needs a "map of categories." This is what is known as a functor:

**Definition** A *functor* between two categories $\mathcal{C}_1$ and $\mathcal{C}_2$ consists of a pair of maps $F_{\rm obj} : Obj(\mathcal{C}_1) \to Obj(\mathcal{C}_2)$ and $F_{\rm mor} : Mor(\mathcal{C}_1) \to Mor(\mathcal{C}_2)$ so that if

$$x \xrightarrow{f} y \ \in \hom(x,y) \tag{17.7}$$

then

$$F_{\rm obj}(x) \xrightarrow{F_{\rm mor}(f)} F_{\rm obj}(y) \ \in \hom(F_{\rm obj}(x), F_{\rm obj}(y)) \tag{17.8}$$

and moreover we require that $F_{\text{mor}}$ should be compatible with composition of morphisms: There are two ways this can happen. If $f_1, f_2$ are composable morphisms then we say $F$ is a *covariant functor* if

$$F_{\text{mor}}(f_1 \circ f_2) = F_{\text{mor}}(f_1) \circ F_{\text{mor}}(f_2) \tag{17.9}$$

and we say that $F$ is a *contravariant functor* if

$$F_{\text{mor}}(f_1 \circ f_2) = F_{\text{mor}}(f_2) \circ F_{\text{mor}}(f_1) \tag{17.10}$$

In both cases we also require [290]

$$F_{\text{mor}}(\text{Id}_X) = \text{Id}_{F(X)} \tag{17.11}$$

We usually drop the subscript on $F$ since it is clear what is meant from context.

---

**Exercise**

Using the alternative definition of a category in terms of data $p_{0,1} : X_1 \to X_0$ define the notion of a functor writing out the relevant commutative diagrams.

---

---

**Exercise** *Opposite Category*

If $\mathcal{C}$ is a category then the *opposite category* $\mathcal{C}^{\text{opp}}$ is defined by just reversing all arrows. More formally: The set of objects is the same and

$$\text{hom}_{\mathcal{C}^{\text{opp}}}(X, Y) := \text{hom}_{\mathcal{C}}(Y, X) \tag{17.12}$$

so for every morphism $f \in \text{hom}_{\mathcal{C}}(Y, X)$ we associate $f^{\text{opp}} \in \text{hom}_{\mathcal{C}^{\text{opp}}}(X, Y)$ such that

$$f_1 \circ_{\mathcal{C}^{\text{opp}}} f_2 = (f_2 \circ_{\mathcal{C}} f_1)^{\text{opp}} \tag{17.13}$$

a.) Show that if $F : \mathcal{C} \to \mathcal{D}$ is a contravariant functor then one can define in a natural way a covariant functor $F : \mathcal{C}^{\text{opp}} \to \mathcal{D}$.

b.) Show that if $F : \mathcal{C} \to \mathcal{D}$ is a covariant functor then we can naturally define another covariant functor $F^{\text{opp}} : \mathcal{C}^{\text{opp}} \to \mathcal{D}^{\text{opp}}$

---

**Example 1**: Every category has a canonical functor to itself, called the identity functor $Id_{\mathcal{C}}$.

**Example 2**: There is an obvious functor, the *forgetful functor* from **GROUP** to **SET**. This idea extends to many other situations where we "forget" some mathematical structure and map to a category of more primitive objects.

[290] Although we do have $F_{\text{mor}}(\text{Id}_X) \circ F_{\text{mor}}(f) = F_{\text{mor}}(f)$ for all $f \in \text{hom}(Y, X)$ and $F_{\text{mor}}(f) \circ F_{\text{mor}}(\text{Id}_X) = F_{\text{mor}}(f)$ for all $f \in \text{hom}(X, Y)$ this is not the same as the statement that $F_{\text{mor}}(\text{Id}_X) \circ \phi = \phi$ for all $\phi \in \text{hom}(F(Y), F(X))$, so we need to impose this extra axiom.

**Example 3**: Since **AB** is a subcategory of **GROUP** there is an obvious functor $\mathcal{F}$ : **AB** $\rightarrow$ **GROUP**.

**Example 4**: In an exercise below you are asked to show that the abelianization of a group defines a functor $\mathcal{G}$ : **GROUP** $\rightarrow$ **AB**.

**Example 5**: Fix a group $G$. Then in the notes above we have on several occasions used the functor

$$F_G : \mathbf{SET} \rightarrow \mathbf{GROUP} \tag{17.14}$$

by observing that if $X$ is a set, then $F_G(X) = Maps[X \rightarrow G]$ is a group. Check this is a contravariant functor: If $f : X_1 \rightarrow X_2$ is a map of sets then

$$F_G(X_1) \xleftarrow{F_G(f)} F_G(X_2) \tag{17.15}$$

The map $F_G(f)$ is usually denoted $f^*$ and is known as the *pull-back*. To be quite explicit: If $\Psi$ is a map of $X_2 \rightarrow G$ then $f^*(\Psi) := \Psi \circ f$ is a map $X_1 \rightarrow G$.

This functor is used in the construction of certain *nonlinear sigma models* which are quantum field theories where the target space is a group $G$. The viewpoint that we are studying the representation theory of an infinite-dimensional group of maps to $G$ has been extremely successful in a particular case of the *Wess-Zumino-Witten* model, a certain two dimensional quantum field theory that enjoys conformal invariance (and more).

**Example 6**: Now let us return to the category **G-SET**. Now fix any set $Y$. Then in the notes above we have on several occasions used the functor

$$F_{G,Y} : \mathbf{G\text{-}SET} \rightarrow \mathbf{G\text{-}SET} \tag{17.16}$$

by observing that if $X$ is a $G$-set, then $F_Y(X) = Maps[X \rightarrow Y]$ is also a $G$-set. To check this is a contravariant functor we write:

$$\begin{aligned}
[g \cdot (f^*\Psi)](x_1) &= (f^*\Psi)(g^{-1} \cdot x_1) \\
&= \Psi(f(g^{-1} \cdot x_1)) \\
&= \Psi(g^{-1} \cdot (f(x_1))) \\
&= (g \cdot \Psi)(f(x_1)) \\
&= (f^*(g \cdot \Psi))(x_1)
\end{aligned} \tag{17.17}$$

and hence $\Psi \rightarrow g \cdot \Psi$ is a morphism of $G$-sets.

This functor is ubiquitous in quantum field theory: If a spacetime enjoys some symmetry (for example rotational or Poincaré symmetry) then the same group will act on the space of fields defined on that spacetime.

**Example 7**: Fix a nonnegative integer $n$ and a group $G$. Then the group cohomology we discussed above (take the trivial twisting $\omega_g = \mathrm{Id}_A$ for all $g$) defines a covariant functor

$$H^n(G, \bullet) : \mathbf{AB} \rightarrow \mathbf{AB} \tag{17.18}$$

To check this is really a functor we need to observe the following: If $\varphi : A_1 \to A_2$ is a homomorphism of Abelian groups then there is an induced homomorphim, usually denoted

$$\varphi_* : H^n(G, A_1) \to H^n(G, A_2) \tag{17.19}$$

You have to check that $\mathrm{Id}_* = \mathrm{Id}$ and

$$(\varphi_1 \circ \varphi_2)_* = (\varphi_1)_* \circ (\varphi_2)_* \tag{17.20}$$

Strictly speaking we should denote $\varphi_*$ by $H^n(G, \varphi)$, but this is too fastidious for the present author.

**Example 8**: Fix a nonnegative integer $n$ and any group $A$. Then the group cohomology we discussed above (take the trivial twisting $\omega_g = \mathrm{Id}_A$ for all $g$) defines a contravariant functor

$$H^n(\bullet, A) : \mathbf{GROUP} \to \mathbf{AB} \tag{17.21}$$

To check this is really a functor we need to observe the following: If $\varphi : G_1 \to G_2$ is a homomorphism of Abelian groups then there is an induced homomorphim, usually denoted $\varphi^*$

$$\varphi^* : H^n(G_2, A) \to H^n(G_1, A) \tag{17.22}$$

**Example 9**: *Topological Field Theory.* The very definition of topological field theory is that it is a functor from a bordism category of manifolds to the category of vector spaces and linear transformations. For much more about this one can consult a number of papers. Two online resources are

http://www.physics.rutgers.edu/∼gmoore/695Fall2015/TopologicalFieldTheory.pdf
https://www.ma.utexas.edu/users/dafr/bordism.pdf

Note that in example 2 there is no obvious functor going the reverse direction. When there are functors both ways between two categories we might ask whether they might be, in some sense, "the same." But saying precisely what is meant by "the same" requires some care.

**Definition** If $\mathcal{C}_1$ and $\mathcal{C}_2$ are categories and $F_1 : \mathcal{C}_1 \to \mathcal{C}_2$ and $F_2 : \mathcal{C}_1 \to \mathcal{C}_2$ are two functors then a *natural transformation* $\tau : F_1 \to F_2$ is a rule which, for every $X \in Obj(\mathcal{C}_1)$ assigns an arrow $\tau_X : F_1(X) \to F_2(X)$ so that, for all $X, Y \in Obj(\mathcal{C}_1)$ and all $f \in \hom(X, Y)$,

$$\tau_Y \circ F_1(f) = F_2(f) \circ \tau_X \tag{17.23}$$

Or, in terms of diagrams.

$$\begin{array}{ccc} F_1(X) & \xrightarrow{F_1(f)} & F_1(Y) \\ \Big\downarrow{\tau_X} & & \Big\downarrow{\tau_Y} \\ F_2(X) & \xrightarrow{F_2(f)} & F_2(Y) \end{array} \tag{17.24}$$

**Example 1**: *The evaluation map.* Here is another tautological construction which nevertheless can be useful. Let $S$ be any set and define a functor

$$F_S : \mathbf{SET} \to \mathbf{SET} \tag{17.25}$$

by saying that on objects we have

$$F_S(X) := Map[S \to X] \times S \tag{17.26}$$

and if $\varphi : X_1 \to X_2$ is a map of sets then

$$F_S(\varphi) : Map[S \to X_1] \times S \to Map[S \to X_2] \times S \tag{17.27}$$

is defined by $F_S(\varphi) : (f, s) \mapsto (\varphi \circ f, s)$. Then we claim there is a natural transformation to the identity functor. For every set $X$ we have

$$\tau_X : F_S(X) = Map[S \to X] \times S \to \mathrm{Id}(X) = X \tag{17.28}$$

It is defined by $\tau_X(f, s) := f(s)$. This is known as the "evaluation map." Then we need to check

$$\begin{array}{ccc} F_S(X) & \xrightarrow{\tau_X} & X \\ \downarrow{\scriptstyle F_S(\varphi)} & & \downarrow{\scriptstyle \varphi} \\ F_S(Y) & \xrightarrow{\tau_Y} & Y \end{array} \tag{17.29}$$

commutes. If you work it out, it is just a tautology.

**Example 2**: *The determinant.* [291] Let **COMMRING** be the category of commutative rings with morphisms the ring morphisms. (So, $\varphi : R_1 \to R_2$ is a homomorphism of Abelian groups and moreover $\varphi(r \cdot s) = \varphi(r) \cdot \varphi(s)$.) Let us consider two functors

$$\mathbf{COMMRING} \to \mathbf{GROUP} \tag{17.30}$$

The first functor $F_1$ maps a ring $R$ to the multiplicative group $U(R)$ of multiplicatively invertible elements. This is often called the group of units in $R$. If $\varphi$ is a morphism of rings and $r \in U(R_1)$ then $\varphi(r) \in U(R_2)$ and the map $\varphi_* : U(R_1) \to U(R_2)$ defined by

$$\varphi_* : r \mapsto \varphi(r) \tag{17.31}$$

is a group homomorphism. So $F_1$ is a functor. The second functor $F_2$ maps a ring $R$ to the matrix group $GL(n, R)$ of $n \times n$ matrices such that there exists an inverse matrix with values in $R$. Again, if $\varphi : R_1 \to R_2$ is a morphism then applying $\varphi$ to each matrix element defines a group homomorphism $\varphi_* : GL(n, R_1) \to GL(n, R_2)$. Now consider the determinant of a matrix $g \in GL(n, R)$. The usual formula

$$\det(g) := \sum_{\sigma \in S_n} \epsilon(\sigma) g_{1,\sigma(1)} \cdot g_{2,\sigma(2)} \cdots g_{n,\sigma(n)} \tag{17.32}$$

[291] This example uses some terms from linear algebra which can be found in the "User's Manual," Chapter 2 below.

makes perfect sense for $g \in GL(n, R)$. Moreover,

$$\det(g_1 g_2) = \det(g_1)\det(g_2) \tag{17.33}$$

Now we claim that the determinant defines a natural transformation $\tau : F_1 \to F_2$. For each object $R \in Ob(\mathbf{COMMRING})$ we assign the morphism

$$\tau_R : GL(n, R) \to U(R) \tag{17.34}$$

defined by $\tau_R(g) := \det(g)$. Thanks to (17.33) this is indeed a morphism in the category **GROUP**, that is, it is a group homomorphism. Moreover, it satisfies the required commutative diagram because if $\varphi : R_1 \to R_2$ is a morphism of rings then

$$\varphi_*(\det(g)) = \det(\varphi_*(g)). \tag{17.35}$$

**Example 3**: *Natural transformations in cohomology theory.* Cohomology groups provide natural examples of functors, as we have stressed above. There are a number of interesting natural transformations between these different cohomology-group functors.

♣Can we explain an elementary example with group cohomology as developed so far??? ♣

**Definition** Two categories are said to be *equivalent* if there are functors $F : \mathcal{C}_1 \to \mathcal{C}_2$ and $G : \mathcal{C}_2 \to \mathcal{C}_1$ together with isomorphisms (via natural transformations) $FG \cong Id_{\mathcal{C}_2}$ and $GF \cong Id_{\mathcal{C}_1}$. (Note that $FG$ and $Id_{\mathcal{C}_2}$ are both objects in the category of functors $\mathrm{FUNCT}(\mathcal{C}_2, \mathcal{C}_2)$ so it makes sense to say that they are isomorphic.)

♣Should explain example showing category of finite-dimensional vector spaces over a field is equivalent to the catetgory of nonnegative integers. ♣

Many important theorems in mathematics can be given an elegant and concise formulation by saying that two seemingly different categories are in fact equivalent. Here is a (very selective) list: [292]

**Example 1**: Consider the category with one object for each nonnegative integer $n$ and the morphism space $GL(n, \kappa)$ of invertible $n \times n$ matrices over the field $\kappa$. These categories are equivalent. That is one way of saying that the only invariant of a finite-dimensional vector space is its dimension.

**Example 2**: The basic relation between Lie groups and Lie algebras the statement that the functor which takes a Lie group $G$ to its tangent space at the identity, $T_1 G$ is an equivalence of the category of connected and simply-connected Lie groups with the category of finite-dimensional Lie algebras. One of the nontrivial theorems in the theory is the existence of a functor from the category of finite-dimensional Lie algebras to the category of connected and simply-connected Lie groups. Intuitively, it is given by exponentiating the elements of the Lie algebra.

**Example 3**: Covering space theory is about an equivalence of categories. On the one hand we have the category of coverings of a pointed space $(X, x_0)$ and on the other hand

[292] I thank G. Segal for a nice discussion that helped prepare this list.

the category of topological spaces with an action of the group $\pi_1(X, x_0)$. Closely related to this, Galois theory can be viewed as an equivalence of categories.

**Example 4**: The category of unital commutative $C^*$-algebras is equivalent to the category of compact Hausdorff topological spaces. This is known as Gelfand's theorem.

**Example 5**: Similar to the previous example, an important point in algebraic geometry is that there is an equivalence of categories of commutative algebras over a field $\kappa$ (with no nilpotent elements) and the category of affine algebraic varieties.

**Example 6**: Pontryagin duality is a nontrivial self-equivalence of the category of locally compact abelian groups (and continuous homomorphisms) with itself.

**Example 7**: A generalization of Pontryagin duality is Tannaka-Krein duality between the category of compact groups and a certain category of linear tensor categories. (The idea is that, given an abstract tensor category satisfying certain conditions one can construct a group, and if that tensor category is the category of representations of a compact group, one recovers that group.)

**Example 8**: The Riemann-Hilbert correspondence can be viewed as an equivalence of categories of flat connections (a.k.a. linear differential equations, a.k.a. D-modules) with their monodromy representations.

♣This needs a lot more explanation. ♣

In physics, the statement of "dualities" between different physical theories can sometimes be formulated precisely as an equivalence of categories. One important example of this is mirror symmetry, which asserts an equivalence of ($A_\infty$)-) categories of the derived category of holomorphic bundles on $X$ and the Fukaya category of Lagrangians on $X^\vee$. But more generally, nontrivial duality symmetries in string theory and field theory have a strong flavor of an equivalence of categories.

---

**Exercise**

As we noted above, there is a functor $\mathbf{AB} \to \mathbf{GROUP}$ just given by inclusion.

a.) Show that the abelianization map $G \to G/[G, G]$ defines a functor $\mathbf{GROUP} \to \mathbf{AB}$.

b.) Show that the existence of nontrivial perfect groups, such as $A_5$, implies that this functor cannot be an equivalence of categories.

---

---

**Exercise** *Playing with natural transformations* a.) Given two categories $\mathcal{C}_1, \mathcal{C}_2$ show that the natural transformations allow one to define a category FUNCT$(\mathcal{C}_1, \mathcal{C}_2)$ whose objects are functors from $\mathcal{C}_1$ to $\mathcal{C}_2$ and whose morphisms are natural transformations. For this reason natural transformations are often called "morphisms of functors." [293]

b.) Write out the meaning of a natural transformation of the identity functor $Id_{\mathcal{C}}$ to itself. Show that $End(Id_{\mathcal{C}})$, the set of all natural transformations of the identity functor to itself is a monoid.

---

---

**Exercise** *Freyd's theorem*

A "practical" way to tell if two categories are equivalent is the following:

By definition, a *fully faithful functor* is a functor $F : \mathcal{C}_1 \to \mathcal{C}_2$ where $F_{\rm mor}$ is a bijection on all the hom-sets. That is, for all $X, Y \in Obj(\mathcal{C}_1)$ the map

$$F_{\rm mor} : \hom(X, Y) \to \hom(F_{\rm obj}(X), F_{\rm obj}(Y)) \tag{17.36}$$

is a bijection.

Show that $\mathcal{C}_1$ is equivalent to $\mathcal{C}_2$ iff there is a fully faithful functor $F : \mathcal{C}_1 \to \mathcal{C}_2$ so that any object $\alpha \in Obj(\mathcal{C}_2)$ is isomorphic to an object of the form $F(X)$ for some $X \in Obj(\mathcal{C}_1)$.

---

---

**Exercise** *Yoneda Lemma*

One of the most important results about natural transformations is the *Yoneda lemma*. In this exercise we work through it.

a.) Suppose we are given a category $\mathcal{C}$ and an object $X \in \mathcal{C}_0$. Show that there are canonical covariant and contravariant functors functor $F_X : \mathcal{C} \to \mathbf{SET}$ and $F^X : \mathcal{C} \to \mathbf{SET}$, respectively. [294]

b.) Now suppose we are given one more piece of data, namely a functor $G : \mathcal{C} \to \mathbf{SET}$. We can ask about the natural transformations from $F_X$ to $G$. Recall that this means that for all $Y, Z \in \mathcal{C}_0$ we have

$$\begin{array}{ccc} F_X(Y) & \xrightarrow{\tau_Y} & G(Y) \\ {\scriptstyle F_X(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F_X(Z) & \xrightarrow[\tau_Z]{} & G(Z) \end{array} \tag{17.37}$$

[293] One can go on to define a "2-category" of categories, whose objects are categories, morphisms are functors, and 2-morphisms are natural transformations.

[294] *Answer.* Let $F_X(Y) := Hom(X, Y)$ on objects. If $f : Y \to Z$ is a morphism then $F_X(f) : F_X(Y) \to F_X(Z)$ is the map of sets $F_X(f) : Hom(X, Y) \to Hom(X, Z)$ that takes $g \in Hom(X, Y)$ to $f \circ g \in Hom(X, Z)$. Similarly, $F^X(Y) = Hom(Y, X)$, etc.

Show that such a natural transformation automatically picks out a distinguished element $p \in G(X)$. [295]

c.) Show that for any object $Z \in \mathcal{C}_0$ and any $f \in Hom(X, Z)$ we must have [296]

$$\tau_Z(f) = G(f)(\tau_X(Id_X)) \tag{17.38}$$

d.) Conclude the Yoneda lemma which states: *There is a natural bijective correspondence between the set of natural transformations from $F_X$ to $G$ with the set $G(X)$.*

---

In addition to the very abstract view of categories we have just sketched, very concrete objects, like groups, manifolds, and orbifolds can profitably be viewed as categories.

One may always picture a category with the objects constituting points and the morphisms directed arrows between the points as shown in Figure 48.

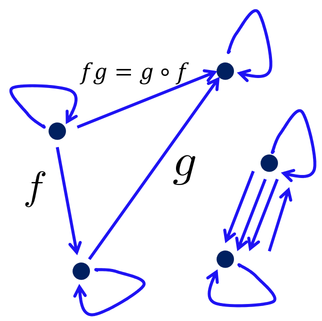



**Figure 48:** Pictorial illustration of a category. The objects are the black dots. The arrows are shown, and one must give a rule for composing each arrow and identifying with one of the other arrows. For example, given the arrows denoted $f$ and $g$ it follows that there must be an arrow of the type denoted $f \circ g$. Note that every object $x$ has at least one arrow, the identity arrow in $Hom(x, x)$.

As an extreme example of this let us consider a category with only *one object*, but we allow the possibility that there are several morphisms. For such a category let us look carefully at the structure on morphisms $f \in Mor(\mathcal{C})$. We know that there is a binary operation, with an identity 1 which is associative.

---

[295] *Answer*: Consider the case $Y = X$. There is a distinguished morphism $Id_X \in Hom(X, X)$. Therefore $\tau_X(Id_X) \in G(X)$ is a distinguished element of $G(X)$.

[296] *Answer*: Apply the defining diagram of a natural transformation to the case $Y = X$. Note that $F_X(f)(Id_X) = f$. Note that $f$ is simultaneously playing two roles here: First as a morphism so we can evaluated $F_X(f)$ to produce a morphism of sets, but also, at the same time, $f$ is an object in the set $Hom(X, Y)$. This is one of the key ideas in the lemma.

But this is just the definition of a monoid!

If we have in addition inverses then we get a group. Hence:

**Definition** A *group* is a category with one object, all of whose morphisms are invertible.

To see that this is equivalent to our previous notion of a group we associate to each morphism a group element. Composition of morphisms is the group operation. The invertibility of morphisms is the existence of inverses.

We will briefly describe an important and far-reaching generalization of a group afforded by this viewpoint. Then we will show that this viewpoint leads to a nice geometrical construction making the formulae of group cohomology a little bit more intuitive.

*************************************

CONSTRUCT EXERCISE HERE EXAMINING HOW CONCEPTS OF FUNCTORS AND NATURAL TRANSFORMATIONS TRANSLATE INTO GROUP THEORY LANGUAGE WHEN SPECIALIZED TO THE CATEGORIES CORRESPONDING TO GROUPS

*************************************

### 17.1 Groupoids

**Definition** A *groupoid* is a category all of whose morphisms are invertible.

Note that for any object $x$ in a groupoid, $\hom(x, x)$ is a group. It is called the *automorphism group* of the object $x$.

**Example 1**. Any equivalence relation on a set $X$ defines a groupoid. The objects are the elements of $X$. The set $\mathrm{Hom}(a, b)$ has one element if $a \sim b$ and is empty otherwise. The composition law on morphisms then means that $a \sim b$ with $b \sim c$ implies $a \sim c$. Clearly, every morphism is invertible.

**Example 2**. Consider time evolution in quantum mechanics with a time-dependent Hamiltonian. There is no sense to time evolution $U(t)$. Rather one must speak of unitary evolution $U(t_1, t_2)$ such that $U(t_1, t_2)U(t_2, t_3) = U(t_1, t_3)$. Given a solution of the Schrodinger equation $\Psi(t)$ we may consider the state vectors $\Psi(t)$ as objects and $U(t_1, t_2)$ as morphisms. In this way a solution of the Schrodinger equation defines a groupoid. ♣Clarify this remark. ♣

**Example 3**. Let $X$ be a topological space. The fundamental groupoid $\pi_{\leq 1}(X)$ is the category whose objects are points $x \in X$, and whose morphisms are homotopy classes of paths $f : x \to x'$. These compose in a natural way. Note that the automorphism group of a point $x \in X$, namely, $\hom(x, x)$ is the fundamental group of $X$ based at $x$, $\pi_1(X, x)$.

**Example 4**. Gauge theory: Objects = connections on a principal bundle. Morphisms = gauge transformations. This is the right point of view for thinking about some more exotic (abelian) gauge theories of higher degree forms which arise in supergravity and string theories.

**Example 5**. In the theory of string theory orbifolds and orientifolds spacetime must be considered to be a groupoid. Suppose we have a right action of $G$ on a set $X$, so we have a map

$$\Phi : X \times G \to X \tag{17.39}$$

such that

$$\Phi(\Phi(x, g_1), g_2) = \Phi(x, g_1 g_2) \tag{17.40}$$

$$\Phi(x, 1_G) = x \tag{17.41}$$

for all $x \in X$ and $g_1, g_2 \in G$. We can just write $\Phi(x, g) := x \cdot g$ for short. We can then form the category $X//G$ with

$$\begin{aligned} Ob(X//G) &= X \\ Mor(X//G) &= X \times G \end{aligned} \tag{17.42}$$

We should think of a morphism as an arrow, labeled by $g$, connecting the point $x$ to the point $x \cdot g$. The target and source maps are: ♣FIGURE NEEDED HERE! ♣

$$p_0((x, g)) := x \cdot g \qquad\qquad p_1((x, g)) := x \tag{17.43}$$

The composition of morphisms is defined by

$$(xg_1, g_2) \circ (x, g_1) := (x, g_1 g_2) \tag{17.44}$$

or, in the other notation (better suited to a right-action):

$$(x, g_1)(xg_1, g_2) := (x, g_1 g_2) \tag{17.45}$$

Note that $(x, 1_G) \in \hom(x, x)$ is the identity morphism, and the composition of morphisms makes sense because we have a group action. Also note that $pt//G$ where $G$ has the trivial action on a point realizes the group $G$ as a category, as sketched above.

**Example 6**. In the theory of string theory orbifolds and orientifolds spacetime must be considered to be a groupoid. (This is closely related to the previous example.)

---

**Exercise**

For a group $G$ let us define a groupoid denoted $G//G$ (for reasons explained later) whose objects are group elements $Obj(G//G) = G$ and whose morphisms are arrows defined by

$$g_1 \xrightarrow{h} g_2 \tag{17.46}$$

iff $g_2 = h^{-1} g_1 h$. This is the groupoid of principal $G$-bundles on the circle.

Draw the groupoid corresponding to $S_3$.

---

---

**Exercise** *The Quotient Groupoid*

a.) Show that whenever $G$ acts on a set $X$ one can canonically define a groupoid: The objects are the points $x \in X$. The morphisms are pairs $(g, x)$, to be thought of as arrows $x \xrightarrow{g} g \cdot x$. Thus, $X_0 = X$ and $X_1 = G \times X$.

b.) What is the automorphism group of an object $x \in X$.

This groupoid is commonly denoted as $X/\!/G$.

---

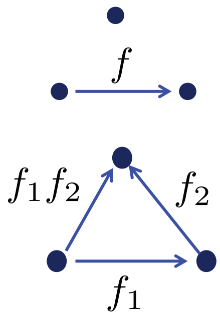



**Figure 49:** Elementary $0, 1, 2$ simplices in the simplicial space $|\mathcal{C}|$ of a category

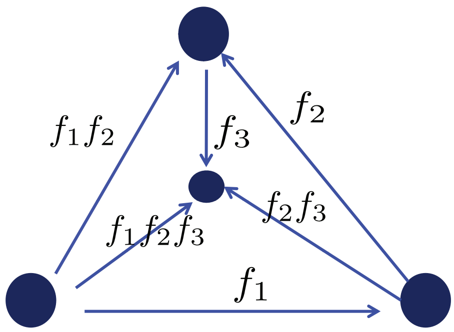



**Figure 50:** An elementary 3-simplex in the simplicial space $|\mathcal{C}|$ of a category

### 17.2 The topology behind group cohomology

Now, let us show that this point of view on the definition of a group can lead to a very nontrivial and beautiful structure associated with a group.

An interesting construction that applies to any category is its associated simplicial space $|\mathcal{C}|$.

This is a space made by gluing together simplices [297] whose simplices are:

[297] Technically, it is a *simplicial space*.

- 0-simplices = objects
- 1-simplices = $\Delta_1(f)$ associated to each morphism $f : x_0 \to x_1 \in X_1$.
- 2-simplices: $\Delta(f_1, f_2)$ associated composable morphisms

$$(f_1, f_2) \in X_2 := \{(f_1, f_2) \in X_1 \times X_1 | p_0(f_1) = p_1(f_2)\} \tag{17.47}$$

- 3-simplices: $\Delta(f_1, f_2, f_3)$ associated to 3 composable morphisms, i.e. elements of:

$$X_3 = \{(f_1, f_2, f_3) \in X_1 \times X_1 \times X_1 | p_0(f_i) = p_1(f_{i+1}), i = 1, 2\} \tag{17.48}$$

- and so on. There are infinitely many simplices of arbitarily high dimension because we can keep composing morphisms as long as we like.

And so on. See Figures 49 and 50. The figures make clear how these simplices are glued together:

$$\partial\Delta_1(f) = x_1 - x_0 \tag{17.49}$$

$$\partial\Delta_2(f_1, f_2) = \Delta_1(f_1) - \Delta_1(f_1 f_2) + \Delta_1(f_2) \tag{17.50}$$

and for Figure 50 view this as looking down on a tetrahedron. Give the 2-simplices of Figure 49 the counterclockwise orientation and the boundary of the 3-simplex the induced orientation from the outwards normal. Then we have

$$\partial\Delta(f_1, f_2, f_3) = \Delta(f_2, f_3) - \Delta(f_1 f_2, f_3) + \Delta(f_1, f_2 f_3) - \Delta(f_1, f_2) \tag{17.51}$$

Note that on the three upper faces of Figure 50 the induced orientation is the ccw orientation for $\Delta(f_1, f_2 f_3)$ and $\Delta(f_2, f_3)$, but with the cw orientation for $\Delta(f_1 f_2, f_3)$. On the bottom fact the inward orientation is ccw and hence the outward orientation is $-\Delta(f_1, f_2)$.

Clearly, we can keep composing morphisms so the space $|\mathcal{C}|$ has simplices of arbitrarily high dimension, that is, it is an infinite-dimensional space.

Let look more closely at this space for the case of a group, regarded as a category with one object. Then in the above pictures we identify all the vertices with a single vertex.

For each group element $g$ we have a one-simplex $\Delta_1(g)$ beginning and ending at this vertex.

For each ordered pair $(g_1, g_2)$ we have an oriented 2-simplex $\Delta(g_1, g_2)$, etc. We simply replace $f_i \to g_i$ in the above formulae, with $g_i$ now interpreted as elements of $G$:

$$\partial\Delta(g) = 0 \tag{17.52}$$

$$\partial\Delta(g_1, g_2) = \Delta_1(g_1) + \Delta_1(g_2) - \Delta_1(g_1 g_2) \tag{17.53}$$

$$\partial\Delta(g_1, g_2, g_3) = \Delta(g_2, g_3) - \Delta(g_1 g_2, g_3) + \Delta(g_1, g_2 g_3) - \Delta(g_1, g_2) \tag{17.54}$$

See Figure 50.

Let us construct this topological space a bit more formally:

We begin by defining $n+1$ maps from $G^n \to G^{n-1}$ for $n \geq 1$ given by

$$\begin{aligned}
d^0(g_1,\dots,g_n) &= (g_2,\dots,g_n)\\
d^1(g_1,\dots,g_n) &= (g_1g_2,g_3,\dots,g_n)\\
d^2(g_1,\dots,g_n) &= (g_1,g_2g_3,g_4,\dots,g_n)\\
&\cdots\cdots\\
&\cdots\cdots\\
d^{n-1}(g_1,\dots,g_n) &= (g_1,\dots,g_{n-1}g_n)\\
d^n(g_1,\dots,g_n) &= (g_1,\dots,g_{n-1})
\end{aligned} \tag{17.55}$$

On the other hand, we can view an $n$-simplex $\Delta_n$ as

$$\Delta_n := \{(t_0,t_1,\dots,t_n)|t_i \geq 0 \qquad \& \qquad \sum_{i=0}^n t_i = 1\} \tag{17.56}$$

Now, there are also $(n+1)$ *face maps* which map an $(n-1)$-simplex $\Delta_{n-1}$ into one of the $(n+1)$ faces of the $n$-simplex $\Delta_n$:

$$\begin{aligned}
d_0(t_0,\dots,t_{n-1}) &= (0,t_0,\dots,t_{n-1})\\
d_1(t_0,\dots,t_{n-1}) &= (t_0,0,t_1,\dots,t_{n-1})\\
&\cdots\cdots\\
&\cdots\cdots\\
d_n(t_0,\dots,t_{n-1}) &= (t_0,\dots,t_{n-1},0)
\end{aligned} \tag{17.57}$$

$d_i$ embeds the $(n-1)$ simplex into the face $t_i = 0$ which is opposite the $i^{th}$ vertex $t_i = 1$ of $\Delta_n$.

Now we identify [298]

$$(\amalg_{n=0}^\infty \Delta_n \times G^n) \,/ \sim$$

via

$$(d_i(\vec{t}),\vec{g}) \sim (\vec{t}, d^i(\vec{g})). \tag{17.58}$$

The space we have constructed this way has a homotopy type denoted $BG$. This homotopy type is known as the *classifying space of the group* $G$. It can be characterized as the homotopy type of a topological space which is both contractible and admits a free $G$-action.

Note that for all $g \in G$, $\partial\Delta_1(g) = 0$, so for each group element there is a closed loop. On the other hand

$$\Delta_1(1_G) = \partial(\Delta_2(1_G,1_G)) \tag{17.59}$$

so $\Delta_1(1_G)$ is a contractible loop. But all other loops are noncontractible. (Show this!) Therefore:

$$\pi_1(BG,*) \cong G \tag{17.60}$$

[298] This means we take the set of equivalence classes and impose the weakest topology on the set of equivalence classes so that the projection map is continuous.

Moreover, if $G$ is a finite group one can show that all the higher homotopy groups of $BG$ are contractible. So then $BG$ is what is known as an *Eilenberg MacLane space* $K(G, 1)$.

Even for the simplest nontrivial group $G = \mathbb{Z}/2\mathbb{Z}$ the construction is quite nontrivial and $BG$ has the homotopy type of $\mathbb{R}P^{\infty}$.

Now, an $n$-cochain in $C^n(G, \mathbb{Z})$ (here we take $A = \mathbb{Z}$ for simplicity) is simply an assignment of an integer for each $n$-simplex in $BG$. Then the coboundary and boundary maps are related by

$$\langle d\phi_n, \Delta\rangle = \langle \phi_n, \partial\Delta\rangle \tag{17.61}$$

and from the above formulae we recover, rather beautifully, the formula for the coboundary in group cohomology.

**Remarks**:

1. When we defined group cohomology we also used homogeneous cochains. This is based on defining $G$ as a groupoid from its left action and considering the mapping of groupoids $G//G \to pt//G$. ♣Explain more here? ♣

2. A Lie group is a manifold and hence has its own cohomology groups as a manifold, $H^n(G; \mathbb{Z})$. There is a relation between these: There is a group homomorphism

$$H^{n+1}_{\text{group cohomology}}(G; \mathbb{Z}) \to H^{n}_{\text{topological space cohomology}}(G; \mathbb{Z}) \tag{17.62}$$

3. One can show that $H^n(BG; \mathbb{Z})$ is always a finite abelian group if $G$ is a finite group. [GIVE REFERENCE].

4. The above construction of $BG$ is already somewhat nontrivial even for the trivial group $G = \{1_G\}$. Indeed, following it through for the 2-cell, we need to identify the three vertices of a triangle to one vertex, and the three edges to a single edge, embedded as a closed circle. If you do this by first identifying two edges and then try to identify the third edge you will see why it is called the "dunce's cap." It is true, but hard to visualize, that this is a contractible space. Things only get worse as we go to higher dimensions. A better construction, due to Milnor, is to construct what is known as a "simplicial set," and then collapse all degenerate simplices to a point. This gives a nicer realization of $BG$, but one which is homotopy equivalent to the one we described above. For the trivial category with one object and one morphism one just gets a topological space consisting of a single point. [299]

5. The "space" $BG$ is really only defined up to homotopy equivalence. For some $G$ there are very nice realizations as infinite-dimensional homogeneous spaces. This is useful for defining things like "universal connections." For example, one model for $B\mathbb{Z}$ is as the humble circle $\mathbb{R}/\mathbb{Z} = S^1$. This generalizes to lattices $B\mathbb{Z}^d = T^d$, the $d$-dimensional torus. On the other hand $B\mathbb{Z}_2$ must be infinite-dimensional but it can be realized as $\mathbb{RP}^{\infty}$, the quotient of the unit sphere in a real infinite-dimensional separable Hilbert space by the antipodal map. Similarly, $BU(1)$ is $\mathbb{CP}^{\infty}$, realized as

[299] I thank G. Segal for helpful remarks on this issue.

the quotient of the unit sphere in a complex infinite-dimensional separable Hilbert space by scaling vectors by a phase: $\psi \to e^{i\theta}\psi$.

## 18. Lattice Gauge Theory

As an application of some of the general concepts of group theory we discuss briefly lattice gauge theory.

Lattice gauge theory can be defined on any graph: There is a set of unoriented edges $\bar{\mathcal{E}}$. Each edge can be given either orientation and we denote the set of oriented edges by $\mathcal{E}$. The set of vertices is denoted $\mathcal{V}$ and source and target maps that tells us the vertex at the beginning and end of each oriented edge:

$$s : \mathcal{E} \to \mathcal{V} \qquad\qquad t : \mathcal{E} \to \mathcal{V} \tag{18.1}$$

We will view the union of edges $\bar{\mathcal{E}}$ (i.e. forgetting the orientation) as a topological space and denote it as $\Gamma$.

The original idea of Ken Wilson was that we could formulate Yang-Mills theory on a "lattice approximation to Euclidean spacetime" which we visualize as a cubic lattice in $\mathbb{R}^d$ for some $d$. Then, the heuristic idea is, that as the bond lengths are taken to zero we get a good approximation to a field theory in the continuum. Making this idea precise is highly nontrivial! For example, just one of the many issues that arise is that important symmetries such as Euclidean or Poincaré symmetries of the continuum models we wish to understand are broken, in this formulation, to crystallographic symmetries.

### 18.1 Some Simple Preliminary Computations

A rather trivial part of the idea is to notice the following: Suppose we have a field theory on $\mathbb{R}^d$ of fields

$$\phi : \mathbb{R}^d \to \mathcal{T} \tag{18.2}$$

where $\mathcal{T}$ is some "target space." Then if we consider the embedded hypercubic lattice:

$$\Lambda_a := \{a(n_1, \dots, n_d) \in \mathbb{R}^d | n_i \in \mathbb{Z}\} \tag{18.3}$$

and we restrict $\phi$ to $\Lambda_a$ then at neighboring vertices the value of $\phi$ will converge as $a \to 0$:

$$\lim_{a\to 0} \phi(\vec{x}_0 + a\hat{e}_\mu) = \phi(\vec{x}_0) \tag{18.4}$$

where $\hat{e}_\mu$, $\mu = 1, \dots, d$ is a unit vector in the $\mu^{th}$ direction. Moreover, if $\phi : \mathbb{R}^d \to \mathcal{T}$ is differentiable and $\mathcal{T}$ is a linear space then

$$\lim_{a\to 0} a^{-1}(\phi(\vec{x}_0 + a\hat{e}_\mu) - \phi(\vec{x}_0)) = \partial_\mu \phi(\vec{x}_0) \tag{18.5}$$

and so on.

In lattice field theory we attempt to go the other way: We assume that we have fields defined on a sequence of lattices $\Lambda_a \subset \mathbb{R}^d$ and try to take an $a \to 0$ limit to define a continuum field theory.

Here is a simple paradigm to keep in mind: [300] Consider the one-dimensional lattice $\mathbb{Z}$, but it is embedded in the real line so that bond-length is $a$, so $\Lambda_a = \{an|n \in \mathbb{Z}\} \subset \mathbb{R}$. Our degrees of freedom will be a real number $\phi_\ell(n)$ at each lattice site $n \in \mathbb{Z}$, and it will evolve in time to give a motion $\phi_\ell(n,t)$ according to the action:

$$S = \int_{\mathbb{R}} dt \sum_{n\in\mathbb{Z}} \left( \frac{m}{2}\dot{\phi}_\ell(n,t)^2 - \frac{k}{2}\left(\phi_\ell(n,t) - \phi_\ell(n+1,t)\right)^2 \right) \tag{18.6}$$

We can think of this as a system of particles of mass $m$ fixed at the vertices of $\Lambda_a$ with neighboring particles connected by a spring with spring constant $k$. For the action to have proper units, $\phi_\ell(n,t)$ should have dimensions of length, suggesting it measures the displacement of the particle in some orthogonal direction to the real line. The equations of motion are of course:

$$m\frac{d^2}{dt^2}\phi_\ell(n,t) = k(\phi_\ell(n+1,t) - 2\phi_\ell(n,t) + \phi_\ell(n-1,t)) \tag{18.7}$$

Now we wish to take the $a \to 0$ limit. We assume that there is some differentiable function $\phi_{cont}(x,t)$ such that

$$\phi_{cont}(x,t)|_{x=an} = \phi_\ell(n,t) \tag{18.8}$$

so by Taylor expansion

$$\phi_\ell(n+1,t) - 2\phi_\ell(n,t) + \phi_\ell(n-1,t) = a^2\frac{d^2}{dx^2}\phi_{cont}|_{x=an} + \mathcal{O}(a^3) \tag{18.9}$$

Now suppose we scale the parameters of the Lagrangian so that

$$m = aT \qquad k = \frac{v^2 T}{a} \tag{18.10}$$

then, if the limits really exist, the continuum function $\phi_{cont}(x,t)$ must satisfy the wave equation:

$$\frac{d^2}{dt^2}\phi_{cont} - v^2\frac{d^2}{dx^2}\phi_{cont} = 0 \tag{18.11}$$

whose general solution is

$$\Phi_{left}(x+vt) + \Phi_{right}(x-vt) \tag{18.12}$$

The general solution is described by arbitrary wavepackets traveling to the left and right along the real line. (We took $v > 0$ here.) We can also see this at the level of the Lagrangian since if $\phi_\ell(n,t)$ is well-approximated by a continuum function $\phi_{cont}(x,t)$ then

$$S \to T\int_{\mathbb{R}} dt \int_{\mathbb{R}} dx \left[ \frac{1}{2}\left(\frac{d}{dt}\phi_{cont}\right)^2 - \frac{v^2}{2}\left(\frac{d}{dx}\phi_{cont}\right)^2 \right] + \mathcal{O}(a) \tag{18.13}$$

**Remarks**

[300] Here we will just latticize the spatial dimension of a $1+1$ dimensional field theory. In the rest of the section we latticize spacetime with Euclidean signature.

1. In the lattice theory there will certainly be sequences of field configurations $\phi_{lattice}(n,t)$ that have no good continuum limit. The idea is that these are unimportant to the physics because they have huge actions whose contributions to the path integral is unimportant in the continuum limit.

2. Keeping in mind the interpretation of $\phi_{cont}(x,t)$ as a height in a direction orthogonal to the real axis, we see that we are describing a string of tension $T$.

### 18.2 Gauge Group And Gauge Field

In lattice gauge theory we choose a group $G$ - known as the *gauge group*. For the moment it can be any group. The dynamical degree of freedom is a *gauge field*, or more precisely, the dynamical object is the *gauge equivalence class* or isomorphism class of the gauge field. This will be defined below.

In mathematics, a gauge field is called a *connection*.

To give the definition of a connection let $\mathcal{P}$ be the set of all connected open paths in $\Gamma$. For example, we can think of it as the set of continuous maps $\gamma : [0,1] \to \Gamma$. Since we are working on a graph you can also think of a path $\gamma$ as a sequence of edges $e_1, e_2, \dots, e_k$ such that

$$t(e_i) = s(e_{i+1}) \qquad 1 \le i \le k-1 \tag{18.14}$$

(We also allow for the trivial path $\gamma_v(t) = v$ for some fixed vertex $v$ which has no edges.) However, the former definition is superior because it generalizes to connections on other topological spaces.

Now, by definition, a connection is just a map

$$\mathbb{U} : \mathcal{P} \to G, \tag{18.15}$$

which satisfies the composition law: If we concatenate two paths $\gamma_1$ and $\gamma_2$ to make a path $\gamma_1 \circ \gamma_2$, so that the concatenated path begins at $\gamma_1(0)$ and ends at $\gamma_2(1)$ and such that $\gamma_1(1) = \gamma_2(0)$, that is, the end of $\gamma_1$ is the beginning of $\gamma_2$, then we must have:

$$\mathbb{U}(\gamma_1 \circ \gamma_2) = \mathbb{U}(\gamma_1)\mathbb{U}(\gamma_2) \tag{18.16}$$

If our path is the trivial path then

$$\mathbb{U}(\gamma_v) = 1_G \tag{18.17}$$

and if $\gamma^{-1}(t) = \gamma(1-t)$ is the path run backwards then

$$\mathbb{U}(\gamma^{-1}) = (\mathbb{U}(\gamma))^{-1} \tag{18.18}$$

♣Are these really independent conditions? ♣

Note that if the path $\gamma$ is made by concatenating edges $e_1, e_2, \dots, e_k$ then

$$\mathbb{U}(\gamma) = \mathbb{U}(e_1)\mathbb{U}(e_2)\cdots\mathbb{U}(e_k) \tag{18.19}$$

so, really, in lattice gauge theory it suffices to know the $\mathbb{U}(e)$ for the edges. If $e^{-1}$ is the edge $e$ with the opposite orientation then

$$\mathbb{U}(e^{-1}) = \mathbb{U}(e)^{-1} \tag{18.20}$$

We will denote the space of all connections by $\mathcal{A}(\Gamma)$.

**Remark**: *Background heuristics*: For those who know something about gauge fields in field theory we should think of $\mathbb{U}(e)$ as the parallel transport (in some trivialization of our principal bundle) along the edge $e$. From these parallel transports along edges we can recover the components of the gauge field. To explain more let us assume for simplicity that $G = U(N)$ is a unitary group, or some matrix subgroup of $U(N)$.

Recall some elementary ideas from the theory of Lie groups: If $\alpha$ is any anti-Hermitian matrix then $\exp[\alpha]$ is a unitary matrix. Moreover, if $\alpha$ is "small" then $\exp[\alpha]$ is close to the identity. Conversely, if $U$ is "close" to the identity then it can be uniquely written in the form $U = \exp[\alpha]$ for a "small" anti-Hermitian matrix $\alpha$. Put more formally: The tangent space to $U(N)$ at the identity is the (real!) vector space of $N \times N$ anti-Hermitian matrices. (This vector space is a real Lie algebra, because the commutator of anti-Hermitian matrices is an anti-Hermitian matrix.) Moreover, the exponential map gives a good coordinate chart in some neighborhood of the identity of the topological group $U(N)$.

The poor man's way of understanding the relation between Lie algebras and Lie groups is to use the very useful Baker-Campbell-Hausdorf formula: If $A, B$ are $n \times n$ matrices then the formula gives an expression for an $n \times n$ matrix $C$ so that

$$e^A e^B = e^C \tag{18.21}$$

The formula is a (very explicit) infinite set of terms all expressed in terms of multiple commutators. The first few terms are:

$$\boxed{C = A + B + \frac{1}{2}[A,B] + \frac{1}{12}[A,[A,B]] + \frac{1}{12}[B,[B,A]] + \frac{1}{24}[A,[B,[A,B]]] + \cdots} \tag{18.22}$$

The series is convergent as long as $A, B$ are small enough (technically, such that the characteristic values of $\mathrm{Ad}(A)$ and $\mathrm{Ad}(B)$ are less than $2\pi$ in magnitude). See Chapter 8 for a full explanation. Note in particular that if we expand in small parameters $\epsilon_1, \epsilon_2$ then

$$e^{\epsilon_1 A} e^{\epsilon_2 B} e^{-\epsilon_1 A} e^{-\epsilon_2 B} = e^{\epsilon_1 \epsilon_2 [A,B] + \cdots} \tag{18.23}$$

Now, returning to lattice gauge theory: In the usual picture of "approximating" Euclidean $\mathbb{R}^d$ by $\mathbb{Z}^d$ with bond-length $a$ we can write a fundamental edge $e_\mu(\vec{n})$ as the straight line in $\mathbb{R}^d$ from $\vec{n}$ to $\vec{n} + a\hat{e}_\mu$. If $a$ is small and we have some suitable continuity then $\mathbb{U}(e_\mu(\vec{n}))$ will be near the identity and we can write:

$$\mathbb{U}(e_\mu(\vec{n})) = \exp[a A_\mu^{lattice}(\vec{n})] \tag{18.24}$$

for some anti-Hermitian matrix $A_\mu(a\vec{n})$. In lattice gauge theory, the connections with a good continuum limit are those such that there is a locally defined 1-form valued in $N \times N$ anti-Hermitian matrices $A_\mu^{cont}(\vec{x}) dx^\mu$ so that $A_\mu^{cont}(a\vec{n}) = A_\mu^{lattice}(\vec{n})$.

Now, the gauge field $\mathbb{U}$ has redundant information in it. The reason it is useful to include this redundant information is that many aspects of locality become much clearer when working with $\mathcal{A}(\Gamma)$ as we will see when trying to write actions below. The redundant information is reflected in a *gauge transformation* which is simply a map

$$f : \mathcal{V} \to G \tag{18.25}$$

The idea is that if $\gamma$ is a path then the gauge fields $\mathbb{U}$ and $\mathbb{U}'$ related by the rule

$$\mathbb{U}'(\gamma) = f(s(\gamma))\mathbb{U}(\gamma)f(t(\gamma))^{-1} \tag{18.26}$$

are deemed to be gauge equivalent, i.e. isomorphic. We denote the set of gauge transformations by $\mathcal{G}(\Gamma)$. Note that, being a function space whose target is a group, this set is a group in a natural way. It is called *the group of gauge transformations.* [301] The group of gauge transformations $\mathcal{G}(\Gamma)$ acts on $\mathcal{A}(\Gamma)$. The moduli space of gauge inequivalent fields is the set of equivalence classes: $\mathcal{A}(\Gamma)/\mathcal{G}(\Gamma)$. Mathematicians would call these isomorphism classes of connections.

It might seem like there is no content here. Can't we always choose $f(s(\gamma))$ to set $\mathbb{U}'(\gamma)$ to 1? Yes, in general, except when $s(\gamma) = t(\gamma)$, that is, when $\gamma$ is a closed loop based at a vertex, say $v_0$. For such closed loops we are stuck, all we can do by gauge transformations is conjugate:

$$\mathbb{U}'(\gamma) = g\mathbb{U}(\gamma)g^{-1} \tag{18.27}$$

where $g$ is the gauge transformation at $v_0$. Moreover, if we start the closed loop at another vertex on the loop then the parallel transport is again in the same conjugacy class. Thus there is gauge invariant information associated to a loop $\gamma$: The conjugacy class of the $\mathbb{U}(\gamma)$. That is: The *holonomy function*:

$$\mathrm{Hol}_{\mathbb{U}} : L\Gamma \to \mathrm{Conj}(G) \tag{18.28}$$

that maps the loops in $\Gamma$ to the conjugacy class:

$$\mathrm{Hol}_{\mathbb{U}} : \gamma \mapsto C(\mathbb{U}(\gamma)) \tag{18.29}$$

is gauge invariant: If $\mathbb{U}' \sim \mathbb{U}$ are gauge equivalent then

$$\mathrm{Hol}_{\mathbb{U}'} = \mathrm{Hol}_{\mathbb{U}} \tag{18.30}$$

In fact, one can show that $\mathrm{Hol}_{\mathbb{U}}$ is a complete invariant, meaning that we have the converse: If $\mathrm{Hol}_{\mathbb{U}'} = \mathrm{Hol}_{\mathbb{U}}$ then $\mathbb{U}'$ is gauge equivalent to $\mathbb{U}$. Put informally:

*The gauge invariant information in a gauge field, or connection, is encoded in the set of conjugacy classes associated to the closed loops in $\Gamma$.*

---

**Exercise**

[301] AND IS NOT TO BE CONFUSED WITH THE gauge group $G$!!!!

Show that if $\gamma$ is a closed loop beginning and ending at $v_0$ and if $v_1$ is another vertex on the path $\gamma$ then if $\gamma'$ describes the "same" loop but starting at $v_1$ then $\mathbb{U}(\gamma)$ and $\mathbb{U}(\gamma')$ are in the same conjugacy class in $G$.

---

---

**Exercise**

Consider a graph $\Gamma$ which forms a star: There is one central vertex, and $r$ "legs" each consisting or $N_i$ edges radiating outward, where $i = 1, \dots, r$.

a.) Show explicitly that any gauge field can be gauged to $\mathbb{U} = 1$.

b.) What is the unbroken subgroup of the group of gauge transformations? (That is, what is the automorphism group of the gauge field $\mathbb{U} = 1$? )

---

---

**Exercise**

Consider a $d$-dimensional hypercubic lattice with periodic boundary conditions, so that we are "approximating a torus" which is a product of "circles" of length $Na$.

What is the maximal number of edges so that we can set $\mathbb{U}(e) = 1$?

---

### 18.3 Defining A Partition Function

Next, to do physics, we need to define a gauge invariant action. At the most general level this is simply a function $F : \mathcal{A}(\Gamma)/\mathcal{G}(\Gamma) \to \mathbb{C}$ so that we can define a "partition function":

$$Z = \sum_{[\mathbb{U}] \in \mathcal{A}(\Gamma)/\mathcal{G}(\Gamma)} F([\mathbb{U}]) \tag{18.31}$$

If $\Gamma$ is finite and $G$ is finite this sum is just a finite sum. If $\Gamma$ is finite and $G$ is a finite-dimensional Lie group then $\mathcal{A}(\Gamma)/\mathcal{G}(\Gamma)$ is a finite-dimensional topological space and the "sum" needs to be interpreted as some kind of integral. Since a connection on $\Gamma$ is completely determined by its values on the elementary edges (for a single orientation) we can, noncanonically, identity the space of all connections as

$$\mathcal{A}(\Gamma) \cong G^{|\bar{\mathcal{E}}|}. \tag{18.32}$$

Similarly

$$\mathcal{G}(\Gamma) \cong G^{|\mathcal{V}|} \tag{18.33}$$

Now we need a way of integrating over the group. If $G$ is a finite group and $F : G \to \mathbb{C}$ is a function then

$$\int_G F d\mu := \frac{1}{|G|} \sum_{g \in G} F(g) \tag{18.34}$$

This basic idea can be generalized to Lie groups. A Lie group is a manifold and we define a measure on it $d\mu$. (If $G$ is a simple Lie group then there is a canonical choice of measure up to an overall scale.) As a simple example, coonsider $U(1) = \{e^{i\theta}\}$ then the integration is

$$\int_0^{2\pi} F(e^{i\theta}) \frac{d\theta}{2\pi} \tag{18.35}$$

In all cases, the crucial property of the group integration is that, for all $h$ we have

$$\int_G F(gh) d\mu(g) = \int_G F(hg) d\mu(g) = \int_G F(g) d\mu(g) \tag{18.36}$$

This property defines what is called a *left-right-invariant measure*. It is also known as the *Haar measure*.

In general the Haar measure is only defined up to an overall scale. In the above examples we chose the normalization so that the volume of the group is 1.

Now, choosing a left-right-invariant measure we can define:

$$Z = \frac{1}{\mathrm{vol}\,(\mathcal{G}(\Gamma))} \int_{\mathcal{A}(\Gamma)} \hat{F}(\mathbb{U}) d\mu_{\mathcal{A}(\Gamma)} \tag{18.37}$$

where $\hat{F}$ is a lifting of $F$ to a $\mathcal{G}(\Gamma)$-invariant function on $\mathcal{A}(\Gamma)$ and $d\mu_{\mathcal{G}(\Gamma)}$ is the Haar measure on $G^{|\bar{\mathcal{E}}|}$ induced by a choice of Haar measure on $G$. It is gauge invariant because the Haar measure is left- and right- invariant.

If we want to impose locality then it is natural to have $\hat{F}(\mathbb{U})$ depend only on the local gauge invariant data. This motivates us to consider "small" loops and consider a *class function*.

In general, a class function on a group $G$ is a function $F : G \to \mathbb{C}$ such that $F(hgh^{-1}) = F(g)$ for all $h \in G$. We should clearly take $\hat{F}$ to be some kind of class function. A natural source of class functions are traces in representations, for if $\rho : G \to GL(N, \mathbb{C})$ is a matrix representation then $\chi(g) := \mathrm{Tr}\rho(g)$ is a class function by cyclicity of the trace. (This class function is called the *character of the representation*.)

The smallest closed loops we can make are the "plaquettes." For $\Lambda_a \subset \mathbb{R}^d$ these would be labeled by a pair of directions $\mu, \nu$ with $\mu \neq \nu$ and would be the closed loop

$$a\vec{n} \to a\vec{n} + a\hat{e}_\mu \to a\vec{n} + a\hat{e}_\mu + a\hat{e}_\nu \to a\vec{n} + a\hat{e}_\nu \to a\vec{n} \tag{18.38}$$

Let us denote this plaquette as $p_{\mu\nu}(\vec{n})$. ♣FIGURE NEEDED HERE! ♣

Given a class function $F : G \to \mathbb{C}$ we can form a partition function by taking

$$\hat{F}(\mathbb{U}) := e^{-S(\mathbb{U})} := e^{-\sum_p S(p)} \tag{18.39}$$

where we have summed over all plaquettes in the exponential to make this look more like a discrete approximation to a field theory path integral, and the action $S(p)$ of a plaquette $p$ is some class function applied to $\mathbb{U}(p)$. If $G$ is a continuous group then we need to interpret the sum over $\mathcal{A}(\Gamma)/\mathcal{G}(\Gamma)$ as some kind of integral, as discussed above.

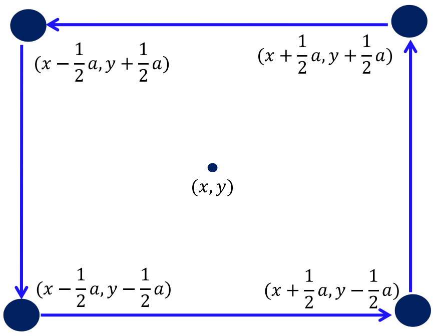



**Figure 51:** A small plaquette, centered on a surface element in a tangent plane with coordinates $(x, y)$ and centered on a point with coordinates $(x, y)$. The holonomy around the plaquette, to leading order in an expansion in small values of bond-length $a$ is governed by the curvature tensor evaluated on that area element.

**Remark**: *More background heuristics*: For those who know something about gauge fields in field theory we should think of the parallel transport $\mathbb{U}(p)$ around a plaquette $p$ as defining the components of the curvature on a small area element $dx^\mu \wedge dx^\nu$ at some point $\vec{x}_0 = a\vec{n}$ (in some framing). Indeed, using the idea that

$$\mathbb{U}(e_\mu(\vec{n})) \sim \exp[aA_\mu^{cont}|_{\vec{x}=a\vec{n}}] \tag{18.40}$$

we can try to take a "limit" where $a \to 0$. The plaquette $p_{\mu\nu}(\vec{n})$ is two-dimensional so, temporarily choosing coordinates so that $\mu = 1$ and $\nu = 2$ we can write the plaquette gauge group element as

$$e^{aA_1(x,y-\frac{1}{2}a)} e^{aA_2(x+\frac{1}{2}a,y)} e^{-aA_1(x,y+\frac{1}{2}a)} e^{-aA_2(x-\frac{1}{2}a,y)} \tag{18.41}$$

See Figure 51. Now, using the BCH formula [302] we define the *fieldstrength of the gauge field* or, equivalently, the *curvature of the connection* by

$$\mathbb{U}(p_{\mu\nu}(\vec{n})) = \exp[a^2 F_{\mu\nu} + \mathcal{O}(a^4)] \tag{18.44}$$

Here in the continuum we would have the relation:

$$F_{\mu\nu}(\vec{x}) = \partial_\mu A_\nu(\vec{x}) - \partial_\nu A_\mu(\vec{x}) + [A_\mu(\vec{x}), A_\nu(\vec{x})] \tag{18.45}$$

[302] Warning: If you are not careful the algebra can be extremely cumbersome here! Taylor expansion in to order $a^2$ gives:

$$e^{aA_1 - \frac{a^2}{2}\partial_2 A_1} e^{aA_2 + \frac{a^2}{2}\partial_1 A_2} e^{-aA_1 - \frac{a^2}{2}\partial_2 A_1} e^{-aA_2 + \frac{a^2}{2}\partial_1 A_2} \tag{18.42}$$

We only need to keep the first commutator term in the BCH formula if we are working to order $a^2$ so we get

$$e^{a^2(\partial_1 A_2 - \partial_2 A_1 + [A_1, A_2]) + \mathcal{O}(a^3)} \tag{18.43}$$

A standard action used in lattice gauge theory in the literature is constructed as follows:

First, choose a finite-dimensional unitary representation of $G$, that is, a group homomorphism

$$\rho : G \to U(r) \tag{18.46}$$

Next, define the action for a plaquette to be

$$S(p) = K(r - \mathrm{Re}[\mathrm{Tr}\rho(\mathbb{U}(p))]) \tag{18.47}$$

for some constant $K$. Note that the trivial gauge field has action $S(p) = 0$. Moreover, every unitary matrix can be diagonalized, by the spectral theorem, with eigenvalues $e^{\mathrm{i}\theta_i(p)}$, $i = 1, \dots, r$ and then

$$S(p) = K \sum_{i=1}^{r} (1 - \cos\theta_i(p)) = 2K \sum_{i=1} \sin^2(\theta_i(p)/2) \tag{18.48}$$

is clearly positive definite for $K > 0$. This is good for unitarity (or its Euclidean counterpart - "reflection positivity.")

**Remarks**:

1. *Correlation Functions*: The typical physical quantities we might want to compute are expectation values of products of gauge invariant operators. In view of our discussion of gauge equivalence classes of gauge fields above one very natural way to make such gauge invariant operators is via *Wilson loop operators.* For these one chooses a matrix representation $R : G \to GL(N, \mathbb{C})$ of $G$ (totally unrelated to the choice we made in defining the action) and a particular loop $\gamma$ and defines:

$$W(R, \gamma)(\mathbb{U}) := \mathrm{Tr}_{\mathbb{C}^N} R(\mathbb{U}(\gamma)) \tag{18.49}$$

So, $W(R, \gamma)$ should be regarded as a gauge invariant function

$$W(R, \gamma) : \mathcal{A}(\Gamma) \to \mathbb{C} \tag{18.50}$$

and therefore we can consider the expectation values:

$$\langle \prod_i W(R_i, \gamma_i) \rangle := \frac{\int_{\mathcal{A}(\Gamma)} \prod_i W(R_i, \gamma_i) e^{-S(\mathbb{U})} d\mu_{\mathcal{A}(\Gamma)}}{\int_{\mathcal{A}(\Gamma)} e^{-S(\mathbb{U})} d\mu_{\mathcal{A}(\Gamma)}} \tag{18.51}$$

2. *Yet more background heuristics*: For those who know something about gauge fields in field theory we can begin to recognize something like the Yang-Mills action if we use (18.44) and write

$$S(p) = K \sum_{p_{\mu\nu}(\vec{n})} (r - \mathrm{Re}[\mathrm{Tr}\rho(\mathbb{U}(p_{\mu\nu}(\vec{n})))]) \to -\frac{1}{2} K a^4 \sum_{\vec{n} \in \mathbb{Z}^d} \sum_{\mu \neq \nu} \mathrm{Tr}\rho(F_{\mu\nu}(a\vec{n}))^2 \tag{18.52}$$

♣There is a bit of a cheat here since you did not work out the plaquette to order $a^4$. ♣

The heuristic limit (18.52) is to be compared with the Yang-Mills action

$$S_{YM} = -\frac{1}{2g_0^2}\int_X d^dx\sqrt{\det g}g^{\mu\lambda}g^{\nu\rho}\mathrm{Tr}F_{\mu\nu}F_{\lambda\rho} \tag{18.53}$$

where here we wrote it in Euclidean signature on a Riemannian manifold $M$. The trace is in some suitable representation and the normalization of the trace can be absorbed in a rescaling of the coupling constant $g_0$. If we use the representation $\rho : G \to U(r)$ then

$$\frac{1}{g_0^2} = Ka^{4-d} \qquad \Rightarrow \qquad K = \frac{a^{d-4}}{g_0^2} \tag{18.54}$$

The constant $K$ must be dimensionless so that $d = 4$ dimensions is selected as special. For $d = 4$ the Yang-Mills coupling $g_0^2$ is dimensionless. It has dimensions of length to a positive power for $d > 4$ and length to a negative power for $d < 4$. To take the continuum limit we should hold $g_0^2$ fixed and scale $K$ as above as $a \to 0$.

3. *Very important subtlety in the case* $d = 4$ Actually, if one attempts to take the limit more carefully, the situation becomes more complicated in $d = 4$, because in quantum mechanics there are important effects known as *vacuum fluctuations*. What is expected to happen (based on continuum field theory) is that, if we replace $K$ by $g^{-2}(a)$ and allow $a$-dependence then we can get a good limit of, say, correlation functions of Wilson loop vev's if we scale $g^2(a)$ so that

$$\frac{8\pi^2}{g^2(a_1)} = \frac{8\pi^2}{g^2(a_2)} + \beta\log\frac{a_1}{a_2} + \mathcal{O}(g^2(a_2)) \tag{18.55}$$

where there are higher order terms in the RHS in an expansion in $g^2(a_2)$. Here $\beta$ is a constant, depending on the gauge group $G$ and other fields in the theory. For $G = SU(n)$ we have the renowned result of D. Gross and F. Wilczek, and of D. Politzer that

$$\beta = -\frac{11}{3}n \tag{18.56}$$

As long as $\beta < 0$ we see that $g^2(a_2) \to 0$ as $a_2 \to 0$. This is known as *asymptotic freedom*. It has the good property that as we attempt to take $a_2 \to 0$ the higher order terms on the RHS are at least formally going to zero.

4. One can therefore ask, to what extent is this continuum limit rigorously defined and how rigorously has (18.55) been established from the lattice gauge theory approach. My impression is that it is still open. Two textbooks on this subject are:

   1. C. Itzykson and J.-M. Drouffe, *Statistical Field Theory*, Cambridge

   2. M. Creutz, *Quarks, gluons, and lattices*, Cambridge

5. *Phases and confinement.* Many crucial physical properties can be deduced from Wilson loop vev's. In Yang-Mills theory a crucial question is whether, for large planar loops $\gamma$ $\langle W(R,\gamma)\rangle$ decays like $\exp[-T Area(\gamma)]$ or $\exp[-\mu Perimeter(\gamma)]$. If it decays like the area one can argue that quarks will be confined. For a nice explanation see S. Coleman, *Aspects Of Symmetry*, for a crystal clear explanation.

6. *Including quarks and QCD.* The beta function is further modified if there are "matter fields" coupling to the gauge fields. If we introduce $n_f$ Dirac fermions in the fundamental representation of $SU(n)$ then (18.56) is modified to:

$$\beta = -\left(\frac{11}{3}n - \frac{2}{3}n_f\right) \tag{18.57}$$

The theory of the strong nuclear force between quarks and gluons is based on $n = 3$ and $n_f = 6$. Actually, there is a strong hierarchy of quark masses so for low energy questions $n_f = 2$ (for "up" and "down" quarks) is more relevant.

7. There are very special situations in which $\beta = 0$ and in fact all the higher terms on the RHS of the "renormalization group equation" (18.55) vanish. These lead to scale-invariant theories, and in good cases, to conformal field theories. In the modern viewpoint on field theory, these conformal field theories are the basic building blocks of all quantum field theories.

## 18.4 Hamiltonian Formulation

EXPLAIN HILBERT SPACE FOR 1+1 CASE IS $L^2(G)$.

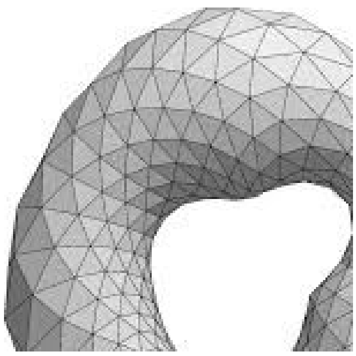


**Figure 52:** A triangulated surface. Figure from Wikipedia.

## 18.5 Topological Gauge Theory

A very popular subject in discussions of topological phases of matter is a set of models known as "topological gauge theories." In general, topological field theories are special classes of field theories that are independent of distances in spacetime. They focus on the topological aspects of physics. A formal mathematical definition is that it is a functor from some bordism category to, say, the category $\mathbf{Vect}_\kappa$.

If $G$ is a finite group and we are working on a smooth manifold then there can be no curvature tensor, so all gauge fields are "flat." They can still be nontrivial since $\mathbb{U}(\gamma)$ can still be nontrivial for homotopically nontrivial loops. The simplest example would be $0+1$ dimensional Yang-Mills theory on a circle. If the action is literally zero then the partition function is just

$$Z = \frac{1}{|G|}\sum_{g\in G} 1 \tag{18.58}$$

Recalling our discussion of the class equation we recognize that the partition function can be written as:

$$Z = \sum_{c.c.} \frac{1}{|Z(g)|} \tag{18.59}$$

where we sum over conjugacy classes in the group and weight each class by one over the order of the centralizer of some (any) representative of that class. This second form of the sum can be interpreted as a sum over the isomorphism classes of principal $G$-bundles over the circle, weighted by one over the automorphism group of the bundle.

For those who know something about gauge theory note that this illustrates a very general principle: *In the partition function of a gauge theory we sum over all the isomorphism classes of bundles with connection: We weight the bundle with connection by a gauge invariant functional divided by the order of the automorphism group of the bundle with connection.*

It is also worth remarking that, quite generally in field theory, the partition function on a manifold of the form $X \times S^1$ can be interpreted as a trace in a Hilbert space. With proper boundary conditions for the "fields" around $S^1$ we simply have

$$Z(X \times S^1) = \mathrm{Tr}_{\mathcal{H}(X)} e^{-\beta H} \tag{18.60}$$

where $\beta$ is the length of the circle. In a topological theory the Hamiltonian $H = 0$, so we just get the dimension of the Hilbert space associated to the spatial slice $X$. In the case of Yang-Mills in $0+1$ dimensions we see that the Hilbert space associated to a point is just $\mathcal{H} = \mathbb{C}$.

In lattice models of topological gauge theories in higher dimensions we insert the gauge-invariant function

$$\prod_p \delta(\mathbb{U}(p)) \tag{18.61}$$

where $\delta(g)$ is the Dirac delta function relative to the measure $d\mu(g)$ we chose on $G$, and is concentrated at $g = 1_G$. Here we take the product over all plaquettes that are meant to be "filled in" in the continuum limit. That means that the parallel transport around "small" loops defined by plaquettes will be trivial. This does not mean that the gauge field is trivial! For example if we consider a triangulation of a compact surface or higher dimensional manifold with nontrivial fundamental group then there can be nontrivial holonomy around homotopically nontrivial loops. In general, a connection, or gauge field, such that $\mathbb{U}(\gamma) = 1$ for homotopically nontrivial loops (this is equivalent to the vanishing of the curvature 2-form $F_{\mu\nu}$) is known as a *flat connection* or *flat gauge field.* In topological gauge theories we sum over (isomorphism classes of) flat connections.

Note that (18.61) is just part of the definition of a topological gauge theory. We want to do this so that physical quantities only depend on topological aspects of the theory. In standard Yang-Mills theory $\langle W(R, \gamma)\rangle$ will depend on lots of details of $\gamma$. Indeed, one definition of the curvature is how $W(R, \gamma)$ responds to small deformations of $\gamma$. In topological gauge theories we want

$$\langle \prod_i W(R_i, \gamma_i)\rangle \tag{18.62}$$

to be independent of (nonintersecting!) $\gamma_i$ under homotopy. Therefore, our measure should be concentrated on flat gauge fields, at least in some heuristic sense. In lattice topological gauge theory we do this by hand.

**Remark**: In general, flat gauge fields for a group $G$ on a manifold $M$ are classified, up to gauge equivalence by the conjugacy classes of homomorphisms $\mathrm{Hom}(\pi_1(M, x_0), G)$.

For a flat gauge field, the standard Wilson action we discussed above will simply vanish. We can get a wider class of models by using group cocycles. This was pointed out in the paper

R. Dijkgraaf and E. Witten, "Topological Gauge Theory And Group Cohomology," Commun.Math.Phys. 129 (1990) 39.

and topological gauge theories that make use of group cocycles for the action are now known as *Dijkgraaf-Witten models.*

For simplicity we now take our group $G$ to be a finite group. Let us start with a two-dimensional model. We can view $\Gamma$ as a triangulation of an oriented surface $M$ as in Figure 52. We want a local action, so let us restrict to a flat gauge field on a triangle as in Figure 49. We want to assign the local "Boltzman weight." It will be a function:

$$W : G \times G \to \mathbb{C}^* \tag{18.63}$$

(If we wish to match to some popular physical theories we might take it to be $U(1)$-valued. The distinction will not matter for anything we discuss here.) Now referring to Figure 49 we assign the weight

$$W(g_1, g_2) \tag{18.64}$$

to this triangle. But now we have to decide if we are to use this, or $W(g_2, (g_1g_2)^{-1})$ or $W((g_1g_2)^{-1}, g_1)$. In general these complex numbers will not be equal to each other. So we number the vertices $1, \dots, |\mathcal{V}|$ and then for any triangle $T$ we start with the vertices with the two smallest numbers. Call this $W(T)$. This will define an orientation that might or might not agree with that on the surface $M$. Let $\epsilon(T) = +1$ if it agrees and $\epsilon(T) = -1$ if it does not. Then the Boltzman weight for a flat gauge field configuration $\mathbb{U}$ on the entire surface is defined to be

$$W(\mathbb{U}) := \prod_{T} W(T)^{\epsilon(T)} \tag{18.65}$$

Now, if this weight is to be at all physically meaningful we definitely want the dependence on all sorts of choices to drop out.

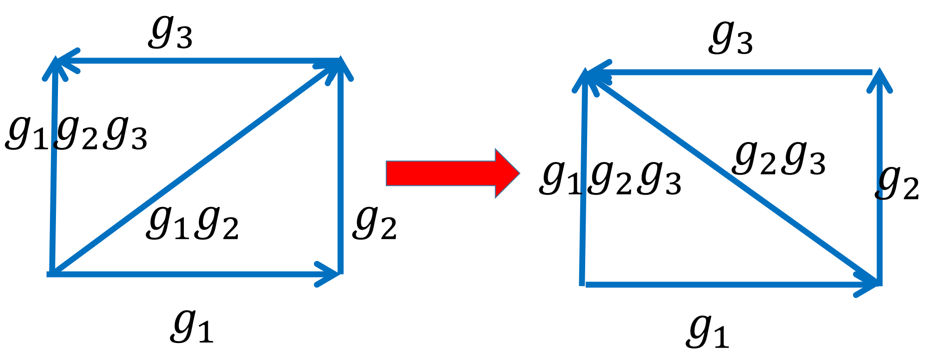



**Figure 53:** A local change of triangulation of type I.

Now, one thing we definitely want to have is independence of the choice of triangulation. A theorem of combinatorial topology states that any two triangulations can be related by a sequence of local changes of type I and type II illustrated in Figure 53 and 54, respectively. We see that the invariance of the action under type I requires:

$$W(g_1, g_2)W(g_1g_2, g_3) = W(g_1, g_2g_3)W(g_2, g_3) \tag{18.66}$$

and this is the condition that $W$ should be a 2-cocycle. Similarly, the change of type II doesn't matter provided

$$W(g_1, g_2) = W(g_1, g_2g_3)W(g_2, g_3)W(g_1g_2, g_3)^{-1} \tag{18.67}$$

which is again guaranteed by the cocycle equation! This strongly suggests we can get a good theory by using a 2-cocycle, and that is indeed the case. But we need to check some things first:

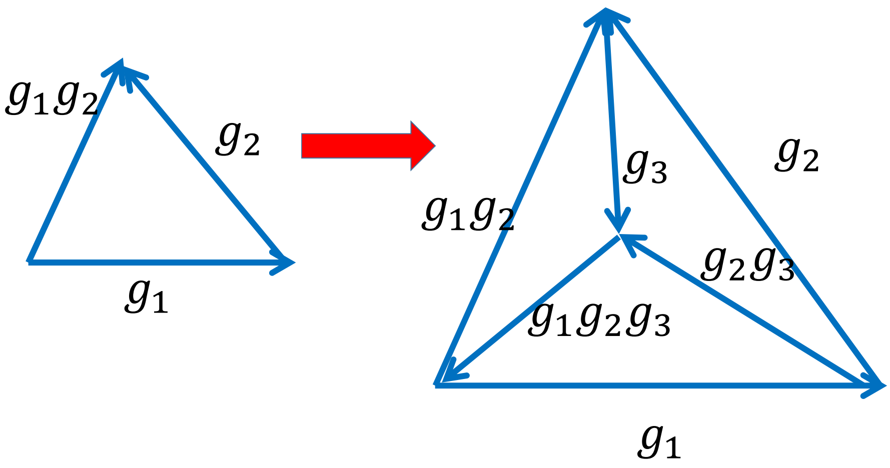



**Figure 54:** A local change of triangulation of type II.

1. The dependence on the labeling of the vertices drops out using an argument based on topology we haven't covered. This can be found in the Dijkgraaf-Witten paper. Similarly, if $W$ is changed by a coboundary then we modify

$$W(g_1, g_2) \to W(g_1, g_2)\frac{t(g_1)t(g_2)}{t(g_1 g_2)} \tag{18.68}$$

that is, we modify the weight by a factor based on a product around the edges. When multiplying the contributions of the individual triangles to get the total weight (18.65) the edge factors will cancel out from the two triangles sharing a common edge.

2. The action is not obviously gauge invariant, since it is certainly not true in general that $W(g_1, g_2)$ is equal to

$$W(h(v_1)^{-1}g_1 h(v_2), h(v_2)^{-1}g_2 h(v_3)) \tag{18.69}$$

for all group elements $h(v_1), h(v_2), h(v_3) \in G$. The argument that, nevertheless, the total action (18.65) is invariant is given (for the $d = 3$ case) in the Dijkgraaf-Witten paper around their equation (6.29).

♣Cop out. Give a better argument. Explain that Chern-Simons actions change by boundary terms and it is too much to hope for exact local gauge invariance. ♣

3. The idea above generalizes to define a topological gauge theory on oriented manifolds in $d$-dimensions for any $d$, where one uses a $d$-cocycle on $G$ with values in $\mathbb{C}^*$ (or $U(1)$). These topological gauge theories are known as "Dijkgraaf-Witten theories." The Boltzmann weight $W$ represents a topological term in the action that exists and is nontrivial even for flat gauge fields.

4. The invariance under the change of type II in Figure 54, which can be generalized to all dimensions is particularly interesting. It means that the action is an "exact renormalization group invariant" in the sense reminiscent of block spin renormalization. [303] This fits in harmoniously with the alleged the metric-independence of the topological gauge theory.

5. The case $d = 3$ is of special interest, and was the main focus of the original Dijkgraaf-Witten paper. In this case we have constructed a "lattice Chern-Simons invariant," and the theory with a cocycle $[W] \in H^3(BG, U(1)) = H^3_{\rm groupcohomology}(G, U(1))$ is a Chern-Simons theory for gauge group $G$. In the case of $G$ finite one can show that $H^3(BG, U(1)) \cong H^4(BG; \mathbb{Z})$. In general the level of a Chern-Simons theory is valued in $H^4(BG; \mathbb{Z})$ for all compact Lie groups $G$.

ALSO DISCUSS HAMILTONIAN VIEWPOINT! Check out D. Harlow and H. Ooguri, Appendix F of https://arxiv.org/pdf/1810.05338.pdf

## 19. Example: Symmetry Protected Phases Of Matter In $1+1$ Dimensions

[303] The idea of block spin renormalization, invented by Leo Kadanoff, is that we impose some small lattice spacing $a$ as a UV cutoff and try to describe an effective theory at ever larger distances. So, we block spins together in some way, define an effective spin, and then an effective action

$$e^{-S_{eff}} := \sum_{fixed-effective-spins} e^{-S(spins)} \tag{18.70}$$

The hope is that at long distances, with ever larger blocks, the "relevant" parts of $S_{eff}$ converge to a useful infrared field theory description.