A Swift Introduction to Group Theory

Cade McManus, Michael Moorman

December 2022

1 Introduction

The question "When will I ever use this?" is one that is likely most frequently asked in the math classroom and in the context of mathematics as a whole. This question is often difficult to answer because of the abstract nature of mathematics. In the context of Math 22a and linear algebra, this question is partially answered through the final projects at the end of the semester. Then, why is it that we are discussing group theory in this paper, another seemingly abstract mathematical concept? The answer is that group theory itself is a powerful tool that can be used to study the natural world in a mathematical, but still intuitive way. In this paper, we will explore the fundamental ideas of group theory and how they can be applied to the real world, both for the sake of understanding the world around us and for the sake of exploring math beyond 22a for the sake of math itself.

2 What is Group Theory?

In order to connect the ideas of group theory to the real world, we must first understand what a group is in relation to linear algebra. A group is a set of elements that are closed under a binary operation, similarly to how vector spaces are closed under addition and scalar multiplication, but with the generalization that groups need not be composed of only vectors, but any set of elements.

Definition 1: A **group** is a set S with a binary[1] operation $\cdot$ such that the following axioms hold:

1. Closure: For all $a, b \in S$, $a \cdot b \in S$.
2. Associativity: For all $a, b, c \in S$, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
3. Identity: There exists a unique element $e \in S$ such that for all $a \in S$, $a \cdot e = e \cdot a = a$.
4. Inverse: For all $a \in S$, there exists a unique element $a^{-1} \in S$ such that $a \cdot a^{-1} = a^{-1} \cdot a = e$.

Formally, the set S is known as the **underlying set** of the group, and the binary operation $\cdot$ is known as the **group operation**. Frequently, a group G is denoted by $G = \langle S, \cdot \rangle$. Additionally, abuses of notation are common, such that G referring to a group can be used to refer to the group itself or the underlying set, such that a statement like '$a \in G$' is meant to convey that a is an element in the underlying set of the group G.

Additionally, another special type of group we will see is an **abelian group**.

Definition 2: An **abelian group** is a group $\langle V, \cdot \rangle$ such that for all $u, v \in V$, $u \cdot v = v \cdot u$. This means that our group operation is commutative, and the order in which we add elements does not matter.

[1] **Binary** refers to the number of elements that the operation requires to be defined, specifically 2. Consider the binary operation of multiplication. The expression $5*$ does not make sense without a second input, making multiplication of real numbers a binary operation.

Example: This may seem confusing, but let's make it more concrete with an example: the real numbers together with the operation of addition as we know it. This group would be written $G = \langle \mathbb{R}, + \rangle$, and it is an abelian group. Let's verify our axioms to truly grasp what they mean through this example:

1. Closure: Let $x, y \in \mathbb{R}$. Because the sum of any two real numbers is defined as a real number, it must be the case that $x + y \in \mathbb{R}$.
2. Associativity: Let $x, y, z \in \mathbb{R}$. Because addition of real numbers does not depend on the way in which terms are associated, it is the case that $(x+y)+z = x+y+z$ and that $x+(y+z) = x+y+z$. Thus, $(x + y) + z = x + (y + z)$.
3. Identity: Let $e = 0$. $0 \in \mathbb{R}$, and the addition of 0 to any $x \in \mathbb{R}$ is equal to x itself. Thus, for all $x \in \mathbb{R}$, $e + x = x + e = x$.
4. Inverse: For all $x \in \mathbb{R}$, allow x^{-1}, the inverse of x, to be equal to $-x$, the negation of x. Thus, $-x + x = 0$ via simple addition, and $e = 0$. Therefore, for any choice of $x \in \mathbb{R}$, there exists an inverse element x^{-1} in $\mathbb{R}$ such that $x + x^{-1} = e$, the identity element.
5. Commutatvitiy: Let $x, y \in \mathbb{R}$. Because addition of real numbers is commutative, it is the case that $x + y = y + x$. (Note that this case is only necessary to prove that a group is abelian, not that it is a group.)

That's all! We have proven that the real numbers, along with the operation of addition as we know it, form an abelian group! Now that we're a bit more comfortable with what a group is, what does it have to do with linear algebra? Or even anything else at all?

3 Vector Spaces and their Connections to Groups

In class, we defined a vector space as a non-empty set V on which two operations, scalar multiplication and vector addition, are defined subject to the following axioms, where $u, v, w \in V$ and $c, d \in \mathbb{R}$:

1. $u + v \in V$
2. $u + v = v + u$
3. $(u + v) + w = u + (v + w)$
4. there exists a vector $0_v \in V$ such that $u + 0_v = u$
5. there exists a vector $-u \in V$ such that $u + (-u) = 0_v$
6. $cu \in V$
7. $c(u + v) = cu + cv$
8. $(c + d)u = cu + du$
9. $c(du) = (cd)u$
10. $1 \cdot u = u$

This is absolutely a correct definition, but we can introduce some additional notation as well as generalize $\mathbb{R}$. We will say that a vector space is a tuple $\langle V, K, +, \cdot \rangle$, where V is a non-empty set, K is a field[2], $+$ is our vector addition operator, and $\cdot$ is our scalar multiplication operator, and all of the previous axioms hold.

As we have seen, a group is a set of elements closed under a binary operation, which seems like a less restrictive version of a vector space, especially after having established this notation. It turns out that this is indeed the case, and that given our vector field $\langle V, K, +, \cdot \rangle$, we can always define an abelian group as simply $\langle V, + \rangle$.

[2]A **field** refers to a set on which addition, multiplication, subtraction, etc. are defined as we know them to work for the real numbers. This is not incredibly important here, but it is relevant to note that the field K is not necessarily the real numbers.

Theorem 1: Let $\langle V, K, +, \cdot \rangle$ be a vector space. Then, $\langle V, + \rangle$ is an abelian group.

Proof. In order to prove this theorem, we will verify all of the axioms of a group given that V is a non-empty set of a vector space, and $+$ is the vector addition operator from that same vector space.

1. Closure: Let $u, v \in V$. Because axiom 1 of a vector space states that $u + v \in V$ and we are using the same set V from our vector space and the same binary operator, this property holds for our group as well.
2. Associativity: Let $u, v, w \in V$. Because axiom 3 of a vector space states that $(u + v) + w = u + (v + w)$ and we are using the same set V from our vector space and the same binary operator, this property holds for our group as well.
3. Identity: Let u be an arbitrary element of V. Because axiom 4 of a vector space states that there exists a vector $0_v \in V$ such that $u + 0_v = u$ and we are using the same set V from our vector space and the same binary operator, this property holds for our group as well using the same identity element 0_v that we use for our vector space.
4. Inverse: Let u be an arbitrary element of V. Because axiom 5 of a vector space states that there exists a vector $-u \in V$ such that $u + (-u) = 0_v$ and we are using the same set V from our vector space and the same binary operator, this property holds for our group as well.

Additionally, this is an abelian group, which requires that our group operation is commutative. We know that this is the case because axiom 2 of a vector space states that $u + v = v + u$ for all $u, v \in V$, where V is our underlying set and $+$ is our group operator. Because we have shown that each of our axioms for a group holds given a non-empty set V from a vector space and the vector addition operator $+$ from that same vector space, and that our group operation is commutative, we have proven that $\langle V, + \rangle$ is an abelian group formed from any arbitrary vector space $\langle V, K, +, \cdot \rangle$. ■

So, there is an interesting connection between groups and vector spaces, but the similarities and connections do not stop there. We turn to examine **subgroups** and their connections to vector subspaces.

4 Subgroups

In similar fashion to the nice connection between groups and vector spaces, there also exists an analog of subspaces in the context of groups.

Definition 3: Let $\langle G, \cdot \rangle$ be a group. Let H be a non-empty subset of G. We say that H is a **subgroup** of G if H is closed under the group operation $\cdot$ and contains the identity element of G, or in other words forms its own group with the binary operation $\cdot$.

Recall how we defined a subspace of a vector space in class, in particular that a subspace is a subset of a vector space that is closed under scalar multiplication and vector addition and contains the zero vector.

Consider how similarly these are defined, the only difference being that we do not have a notion of scalar multiplication in our group. In fact, these are so similar that there exists a very nice continuation of Theorem 1 with regard to subgroups.

Theorem 2: Let $B = \langle W, K, +, \cdot \rangle$ be a supspace of vector space $A = \langle V, K, +, \cdot \rangle$. Then, $S = \langle W, + \rangle$ is a subgroup of $G = \langle V, + \rangle$.

Proof. In order to show this is true, we must show that $W \subseteq V$, that W is closed under the group operation $+$, and that $0_v \in W$, all of the conditions necessary for S to be a subgroup of G.

1. $W \subseteq V$: This is true because B is a subspace of A, which means by definition that W is a subset of V, the underlying set of A.

2. W is closed under the group operation $+$: This is true because B is a subspace of A, which means by definition that W is closed under the vector addition operator $+$.

3. $0_v \in W$: This is true because B is a subspace of A, which means by definition that $0_v \in W$, the zero vector of V. By the proof of Theorem 2, we know that this is the identity element of G and therefore of S.

■

We've already shown that each vector space can be transformed into a group, so it naturally followed already that each vector subspace can be transformed into a group. However, now we know that these groups are subgroups of the group formed from the original vector space! This is a very interesting result, and further ties together the concepts of groups and vector spaces. Besides demonstrating these structural similarities, we can turn to examine the similarities between functions between groups, known as **group homomorphisms**, and linear transformations between vector spaces.

5 Group Homomorphisms and Isomorphisms

Definition 3: A **group homomorphism** is a function $f : G \to H$ from a group $\langle G, \cdot \rangle$ to a group $\langle H, * \rangle$ such that f preserves the group structure of G and H. That is, f must satisfy the following property:

1. $f(g \cdot h) = f(g) * f(h)$

where $g, h \in G$. From this property, it can be deduced that $f(e_G) = e_H$, where e_G is the identity element of G and e_H is the identity element of H, and that $f(g^{-1}) = (f(g))^{-1}$. We leave these proofs as an exercise to the reader.

This property may look familiar, and it should. This is exactly how we define a linear transformation with respect to vector addition. In fact, we can generalize this as follows.

Theorem 3: Let $\langle V, K, +, \cdot \rangle$ and $\langle W, L, \oplus, \odot \rangle$ be vector spaces. Then, a linear transformation $T : V \to W$ is a group homomorphism from $\langle V, + \rangle$ to $\langle W, \oplus \rangle$.

Proof. One of the two axioms[3] of a linear transformation is that $T(u+v) = T(u) \oplus T(v)$ for all $u, v \in V$, when T maps V to W. Well, because our binary operator for our group $\langle V, + \rangle$ is $+$, and our binary operator for our group $\langle W, \oplus \rangle$ is $\oplus$, our sole property for a group homomorphism is satisfied simply by this one axiom of our linear transformation, that $T(u + v) = T(u) \oplus T(v)$ for all $u, v \in V$. ■

Thus, group homomorphisms are generalizations of linear transformations between vector spaces, just like how groups are generalizations of vector spaces. We're not done just yet, though. We now introduce the concept of a **group isomorphism**.

Definition 4: A **group isomorphism** is a special type of group homomorphism that is bijective. That is, a group isomorphism is a function $f : G \to H$ from a group $\langle G, \cdot \rangle$ to a group $\langle H, * \rangle$ such that there exists a function $g : H \to G$ such that $f(g(h)) = h$ and $g(f(g)) = g$ for all $h \in H$. g is known as the **inverse homomorphism** of f. So, a group isomorphism is a group homomorphism that has an inverse homomorphism.

We can generalize this to linear transformations quite nicely as well.

Theorem 4: Let $\langle V, K, +, \cdot \rangle$ and $\langle W, L, \oplus, \odot \rangle$ be vector spaces. Then, a linear transformation $T : V \to W$ between vector spaces is a group isomorphism from $\langle V, + \rangle$ to $\langle W, \oplus \rangle$ if and only if T is bijective.

Proof. We will prove this by showing that if T is bijective, then T is a group isomorphism, and that if T is a group isomorphism, then T is bijective.

[3] In class, we present this axiom as $T(u + v) = T(u) + T(v)$, but it can really be written as $T(u + v) = T(u) \oplus T(v)$, as our binary operator for vector addition may be different in our W vector space.

1. If T is bijective, then T is a group isomorphism. Let $f : V \to W$ be the function $f(v) = T(v)$ for all $v \in V$. Then, f is a group homomorphism from $\langle V, + \rangle$ to $\langle W, \oplus \rangle$ by Theorem 3. Let $g : W \to V$ be the function $g(w) = T^{-1}(w)$ for all $w \in W$. We know this inverse exists because T is bijective. Then, g is the inverse homomorphism of f because $f(g(w)) = T(T^{-1}(w)) = w$ and $g(f(g)) = T^{-1}(T(g)) = g$ for all $w \in W$. Thus, T is a group isomorphism.

2. If T is a group isomorphism, then T is bijective. Let $f : V \to W$ be the function $f(v) = T(v)$ for all $v \in V$. Then, f is a group homomorphism from $\langle V, + \rangle$ to $\langle W, \oplus \rangle$ by Theorem 3. Let $g : W \to V$ be the function $g(w) = T^{-1}(w)$ for all $w \in W$. We know that this inverse homomorphism exists because T is a group isomorphism. Then, g is the inverse homomorphism of f because $f(g(w)) = T(T^{-1}(w)) = w$ and $g(f(g)) = T^{-1}(T(g)) = g$ for all $w \in W$. Thus, g is a group homomorphism from $\langle W, \oplus \rangle$ to $\langle V, + \rangle$. Because $f(g(w)) = w$ for all $w \in W$ and therefore $T(T^{-1}(w)) = w$ for all $w \in W$, we know that T must be bijective[4] because it has a defined inverse for all $w \in W$.

■

By now, we've laid the general groundwork for group theory. We've seen connections to vector spaces, linear transformations, subspaces, and more. In the way that linear algebra provides the tools to study abstract vector spaces for plenty of applications, the same is true for group theory. In the next section, we will see how group theory can be used to study symmetry, in both abstract mathematical terms and in the context of physical systems!

6 Groups of Functions, Symmetry, and Group Actions

We have seen how vector spaces can satisfy the group axioms, so a natural area of inquiry might be to dig further into the connection between these objects. For example, instead of investigating questions about groups directly, might we instead be able to study some equivalent vector space / linear algebraic object and then leverage the tools that we have been studying this semester to answer our original question? As it turns out this is indeed possible. For the purposes of this paper (and simplicity) feel free to assume the relevant objects are finite linear groups.

To find the explicit connection we will first introduce the notion of **group actions** in the context of vector spaces. Suppose we have a real vector space $V = \langle \mathbb{R}^n, K, +, \cdot \rangle$, and the set of n x n invertible matrices[5], $M_{n\text{x}n}$, in that space. This set of matrices is known as the **general linear group** of V, $\text{GL}(V)$, and is indeed a group under matrix multiplication as the binary operator.

Theorem 5: The set of n x n invertible matrices, $M_{n\text{x}n}$, in $\mathbb{R}^n$ forms a group under matrix multiplication as the binary operator, written $\langle M_{n\text{x}n}, \circ \rangle$.

Proof. We will prove this by showing that $M_{n\text{x}n}$ is closed under matrix multiplication, that the identity matrix is an identity element, and that the inverse of a matrix is indeed its inverse.

1. $M_{n\text{x}n}$ is closed under matrix multiplication. Let $A, B \in M_{n\text{x}n}$. Then, A and B are invertible n x n matrices. Then, $\det(A) \neq 0$ and $\det(B) \neq 0$ by the IMT. Because $\det(AB)$=$\det(A)\det(B)$, we know that the matrix AB does not have a zero determinant, so it is invertible. Then, $AB \in M_{n\text{x}n}$.

2. The identity matrix is an identity element. Let $A \in M_{n\text{x}n}$. Because $AI_n = A$ and $I_nA = A$, we know that I_n is an identity element for $M_{n\text{x}n}$.

3. The inverse of a matrix is its inverse group element. Let $A \in M_{n\text{x}n}$. Then, A is an invertible n x n matrix. Then, A^{-1} is also an invertible n x n matrix. Because $AA^{-1} = I_n$, we know that for all $A \in M_{n\text{x}n}$ there exists an A^{-1} such that $AA^{-1} = I_n$.

Thus, $M_{n\text{x}n}$ is a group under matrix multiplication as the binary operator. ■

[4]It suffices to say that T is bijective because it has a defined inverse. The proof for this statement is elementary, and is left as an exercise for the reader in order to keep this proof both short and on topic.

[5]Typically, $M_{n\text{x}n}$ denotes the set of all $n\text{x}n$ matrices, but we will use this notation to denote the set of invertible matrices for simplicity and because we will not use non-invertible matrices in this paper.

This may seem trivial, but recognize how this is different from our previous examples of groups. Matrices as we know them are not only objects, but linear transformations themselves. This opens the door for not only studying groups of objects like vectors, but also for bijective functions, in this example over vector spaces. Additionally, consider that matrix multiplication is simply the composition of matrices. Thus, we can think of the group of matrices as a group of functions (linear transformations), with the binary operator of function composition over a certain domain.

With it in mind that groups can take the form of sets of functions and transformations over vector spaces, consider how we could extend this to other domains besides vector spaces. Additionally, consider how these matrices are all invertible, and preserve the structure of the vector space. This is an important property, and is highly related to the notion of symmetry when groups like these are extended to physical systems.

This idea is important, so consider the example of a square (consider the location of its vertices, sides, or any defining property to be our domain). If you were to rotate that square by 90 degrees, it would look exactly the same. Its vertices may be in different locations, but they were mapped to each other in a such a way that the structure we care about preserving (square-ness) is kept intact. Also, notice that there are many different symmetries that one could exploit to perform a symmetric action, such as flipping the square over. Because of the nature of a "symmetric operation", whenever one considers the set of symmetric operations on an object, there is an identity element, an inverse, and the binary operator of composition. The underlying structure always looks the same from these operations, so composing symmetric operations will also produce the same structure, so the set of symmetric operations is closed under composition, which makes the set of symmetric operations a group, known as a **symmetry group**.

Definition 6: A **symmetry group** of an object X is a group G of transformations of X that preserves the structure of X with the binary operator of composition. It is commonly denoted as $G = \mathrm{Sym}(X)$.

The upshot is this: groups have a deep connection to symmetry, and **group actions** acting on a set are homomorphisms from a given group into the group of transformations of that set. When we can perform group actions on a vector space V, we can often map group elements to subgroups of $\mathrm{GL}(V)$, allowing us to use linear algebra. We now formalize what we mean by a 'group action'.

Definition N: If G is a group with identity element ϵ and X is a set, then a function $\alpha : G \times X \to X$ is a **group action** if it satisfies the following two axioms:

1. Identity: $\alpha(\epsilon, x) = x$ for all $x \in X$
2. Compatability: $\alpha(g, \alpha(h, x)) = \alpha(gh, x)$ for all $x \in X$

Remember that we have 3 objects under consideration. Group G, the set of automorphism (or equivalently, symmetries) of X, denoted $\mathrm{Sym}(X)$, and X itself. As elements in G correspond to symmetry operations of X. We interpret $\alpha(g, x) \text{where} g \in G, x \in X$ as outputting where the symmetry operation corresponding to element g takes element x when applied to set X. Back in the square example, X would the set of vertices, x would be the top right vertex, and g would correspond to, for example, a $90°$ clockwise rotation, and $\alpha(g, x)$ would output the bottom right vertex.

While the linear case of group actions may not always apply, Cayley's theorem tells us that every group is isomorphic to a subgroup of some symmetric group (though this may be trivial, such as the identity symmetry group and or a permutation group of the group itself).
Proof: Later sorry :/

7 Representations

Now that we have done the groundwork we are ready to make explicit this intuitive connection between groups and vector spaces using **Representation Theory**. A **representation** is a group action onto $\mathrm{GL}(V)$, where V is a vector space. The elements of a group are mapped to invertible matrices in this case.

Definition N + 1: A **representation** of group G on a vector space V over field K, is a group homomorphism from G to $\mathrm{GL}(V)$, $\rho : G \to GL(V)$ such that $\rho(g, h) = \rho(gh)$

We've been a bit abstract so lets look at an example of a group, and show how one might generate such a representation. Lets consider an equilateral triangle as our object, and its corners the elements we are tracking. It turns out that (including the do-nothing identity) there are 6 symmetry operations we can perform to leave the triangle looking the same. For now lets just consider the clockwise 120° rotation, denoted as q. This is notational abuse, as given our previous description, q should be an element of a group we are mapping to the set of symmetry operations. In any case, we want to find an invertible matrix that captures what this element q does. That ought to sound familiar. When we wanted to find a matrix for a linear function, we looked at where it took our basis vectors - so let's choose a basis corresponding to the elements we are tracking. Let t, l, r correspond to the top, left, and right vertex respectively.

$$t = \begin{bmatrix}1\\0\\0\end{bmatrix},\ l = \begin{bmatrix}0\\1\\0\end{bmatrix},\ r = \begin{bmatrix}0\\0\\1\end{bmatrix}$$

Note that the standard basis was chosen merely for convience, and any basis could function equivalently (this will come up a bit later). Now let's observe where $q \cdot q$, l, and r takes our basis vectors. As a reminder, q is a 120°. rotation. For simplicity sake we shorten $\rho(q, x)$ to $q \cdot x$

$$q \cdot t = \begin{bmatrix}0\\0\\1\end{bmatrix},\ q \cdot l = \begin{bmatrix}1\\0\\0\end{bmatrix},\ q \cdot r = \begin{bmatrix}0\\1\\0\end{bmatrix}$$

So, just as we have done in class we can, with respect to our chosen basis, write an explicit matrix for q.

$$q = \begin{bmatrix}0 & 1 & 0\\0 & 0 & 1\\1 & 0 & 0\end{bmatrix}$$

In many applications of group theory, for example in chemistry, we may be interested in the "simplest possible" representation of a group. We call these representations irreducible (sometimes called irreps). Helpful for intuition is the notion of **orbits**, where an orbit of an element x of set X under group G is just all elements in X that can be mapped to (or from, by definition) x. It is sort of like its own little world within the bigger world of X. A similar idea applies to representations, where of course we are working with vector spaces instead of general sets. Here the elements being acted upon by the group are vectors, and sets of those vectors where every vector is mapped onto another in the set by the operations in G would form an orbit.

Definition N + 2: Let $\rho : G \to GL(V)$ be a representation. A G-invariant subspace W is a subspace of V such that for all $w \in W$ and $g \in G$, $g \cdot w \in W$. If we restrict the representation to just W then W forms a **sub-representation** of V, and thus V would not be irreducible.

Definition N + 3: An irreducible representation V is one with only trivial subrepresentations (0, and V).

Didn't have time to finish, but to spoil the ending, the notion of eigenvectors comes into play here.

References

[1] M. Aschbacher. *Finite Group Theory*. 2nd ed. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2000. DOI: 10.1017/CBO9781139175319.

[2] David S. Dummit and Richard M. Foote. *Abstract Algebra*. Ed. by Jennifer Battista. Third Edition. Chapter 1. John Wiley and Sons, Inc., 2004.

[3] Willard Miller. *Symmetry Groups and Their Applications*. Academic Press, 1972.

[4] James S. Milne. *Group Theory (v4.00)*. Chapters 1 and 4. Available at www.jmilne.org/math/. 2021.