Contents

1 Group Theory 4
  1.1 Definition of a group ......................................... 4
  1.2 Subgroups .................................................. 7
  1.3 Group Action ................................................ 9
  1.4 Useful definitions .......................................... 10
  1.5 Representation theory ...................................... 11
    1.5.1 Irreducible representations ............................ 12
  1.6 Lie Groups and Lie Algebras ............................... 13
    1.6.1 Pre-discussion of manifolds including physics ...... 15
    1.6.2 Differentiable Manifolds ............................... 16
    1.6.3 Why should we study Lie groups and Lie algebras? . . 19
    1.6.4 Lie Groups ............................................. 20
    1.6.5 Why bother with Lie Algebras? ......................... 21
    1.6.6 Lie Algebras .......................................... 22
    1.6.7 Structure Constants .................................. 24
    1.6.8 The Dynkin Index ..................................... 25
  1.7 The Poincaré Group and its Algebra ....................... 25
    1.7.1 The Lorentz Group and its Algebra .................... 25
    1.7.2 The Poincaré Group and its Algebra .................. 25
    1.7.3 Casimir Operators .................................... 25

2 Tensors 25
  2.1 Semi-Rigorous Mathematical Definition .................... 26
    2.1.1 Dual Vectors ......................................... 26
    2.1.2 Dual Spaces In Terms of Row and Column Vectors .... 27
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1.3</td>
<td>Tensors</td>
<td>28</td>
</tr>
<tr>
<td>2.1.4</td>
<td>A Brief Discussion on Tensor Products</td>
<td>29</td>
</tr>
<tr>
<td>2.1.5</td>
<td>Tensor Fields on Manifolds</td>
<td>31</td>
</tr>
<tr>
<td>2.2</td>
<td>Usual Definition as Seen in Physics Texts</td>
<td>32</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Index Gymnastics and Important Tensors</td>
<td>33</td>
</tr>
<tr>
<td>3</td>
<td>Various Mathematical Tools</td>
<td>33</td>
</tr>
<tr>
<td>3.1</td>
<td>Fourier Transforms</td>
<td>33</td>
</tr>
<tr>
<td>3.2</td>
<td>The Dirac Delta Distribution</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>Green’s Functions</td>
<td>33</td>
</tr>
<tr>
<td>3.4</td>
<td>Complex Analysis</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>Classical field theory</td>
<td>33</td>
</tr>
<tr>
<td>4.1</td>
<td>The need for quantum fields</td>
<td>33</td>
</tr>
<tr>
<td>4.2</td>
<td>Lagrangian and Hamiltonian Field Theory</td>
<td>34</td>
</tr>
<tr>
<td>4.3</td>
<td>Noether’s Theorem</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>The Klein-Gordon Field</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>The Dirac Field</td>
<td>37</td>
</tr>
<tr>
<td>7</td>
<td>Quantization of Free Fields</td>
<td>37</td>
</tr>
<tr>
<td>7.1</td>
<td>Fock Space</td>
<td>37</td>
</tr>
<tr>
<td>8</td>
<td>Perturbation Theory and Feynman Diagrams</td>
<td>37</td>
</tr>
<tr>
<td>9</td>
<td>Elementary QED Processes &amp; Functional Methods</td>
<td>37</td>
</tr>
<tr>
<td>10</td>
<td>Cross sections and the S-matrix; Feynman Rules for QED</td>
<td>37</td>
</tr>
<tr>
<td>11</td>
<td>Non-Abelian Gauge Invariance</td>
<td>37</td>
</tr>
</tbody>
</table>
12 Quantization of Non-Abelian Gauge Theories 37

13 Higgs Mechanism & Electroweak Theory 37
1 Group Theory

In short, group theory is the study of symmetry. If a geometrical shape is left invariant under some form of transformation we say that the shape is symmetric, e.g. the rotation of a sphere which is highly symmetric. The notion of symmetry does not only apply to geometrical objects, it also applies to more abstract things like functions. For example, any rearrangement of $x$, $y$ and $z$ leaves the function $x^2 + y^2 + z^2$ invariant.

1.1 Definition of a group

A group $(G, \circ)$ is a set $G$ together with a binary operation\(^1\) $(a, b) \mapsto a \circ b$ that satisfies the following axioms:

Definition 1.1. 
1. $a \circ b \in G$, $\forall a, b \in G$ (Closure: similar to linear algebra where the condition $(u + v) \in V$ is satisfied if $u, v \in V$)
2. $(a \circ b) \circ c = a \circ (b \circ c)$, $\forall a, b, c \in G$ (Associativity: does not imply the commutative property)
3. $\exists e$ such that $e \circ a = a \circ e = a$, $\forall a \in G$ (Identity element)
4. $\exists a^{-1}$ such that $a^{-1} \circ a = a \circ a^{-1} = e$, $\forall a \in G$ (Inverse element)

In order to understand why this abstract definition is of use to physicists we can take a look at a familiar concept, namely mappings between sets where we will look at transformation groups\(^1\). That is, the set of invertible mappings onto itself. This is not only important in physics, mappings are the most basic operations one can conduct on sets.

The rotations mentioned above is really a mapping of space and this also includes translations, so mappings are important in physical theories regarding symmetries and other elementary properties. Because we want to be able to find the composition of any two given mappings, we need to insist that the mappings operate on a single set. That is, if both $f$ and $g$ is a mapping $A \to B$ we cannot compose the two. If we have three sets $A, B$ and $C$ where $f : A \to B$ and $g : B \to C$ and if $f \circ g$ is defined then $g \circ f$ is not. In physics, tracing back the path of an object

\(^1\)A binary operation on a set $G$ is a function $G \times G \to G$ that assigns to each pair $(a, b) \in G \times G$ a unique element $a \circ b$ called the composition of $a$ and $b.$
to its original starting point is something we can do. This corresponds to the act of undoing a mapping, i.e. tracing back the original position in the initial set. If mappings have an inverse we can do this. And so, if we try to find the composition of such an inverse mapping with its original mapping, we’ll be left with the identity element. If these properties of mappings sound familiar, that’s good. If not, then it’s worth having another look at the definition of a group given above.

An important distinction of groups is whether they are finite or infinite. Finite groups have a finite amount of elements \( g_i \in G \), and thus infinite groups have an infinite amount of elements\(^2\). The number of elements in a group is called its order, denoted by \( |G| \). Let’s look at some examples of sets that have the properties of a group under some binary operation.

**Example 1.1.**

1. The countable infinite set of integers, \( \mathbb{Z} \), is a group under the binary operation of addition. The identity element is 0 while the inverse of \( n \in \mathbb{Z} \) is \(-n\).

2. An example of an uncountable infinite set is \( \mathbb{R} \). Like \( \mathbb{Z} \), \( \mathbb{R} \) forms a group under the operation of addition where the identity element is 0 and the inverse of \( r \) is \(-r\).

3. The uncountable infinite set \( V \) of vectors in a vector space forms a group under addition whose identity element is the zero vector and the inverse of \( |a| \) is \(-|a|\).

4. The uncountable infinite set \( \mathbb{C} \setminus \{0\} \) forms a group under multiplication where the identity element is 1 and the inverse of \( z \) is \(1/z\).

5. The set \( \{A\} \) of invertible \( n \times n \) matrices forms a group under multiplication with the identity element given by the \( n \times n \) identity matrix. The inverse of \( A \) is \( A^{-1} \). This group, as the one in 4., is also uncountably infinite.

Abelian groups are exactly like ordinary groups, but with a small twist. The binary operation of two elements in an abelian group is **commutative**. That is, \( a \circ b = b \circ a \).

**Definition 1.2.** A group \((G, \circ)\) is called **abelian** or **commutative** if \( a \circ b = b \circ a \), \( \forall a, b \in G \). It is common to denote the binary operation of an abelian group by +.

We’ve already seen a few examples of such groups! The groups in ’Example 1.1.’ are all abelian except the last one containing the matrices. That is because matrix multiplication does not necessarily commute. When the binary operation does not

\(^2\)The infinite groups can have countable or continuous number of elements.
commute, we call the group non-Abelian.

**Example 1.2.** Let $A$ be a vector potential that gives rise to a magnetic field $B$. If we transform $A$ somehow and it still gives rise to the same magnetic field $B$, we say that the transformation leaves $A$ invariant. The set of these kinds of transformations is in fact an abelian group. If you remember your lectures on classical electrodynamics, this is the transformation where we add the gradient of some function $\chi$ to $A$.

**Definition 1.3.** Let $(G, \ast)$ and $(H, \circ)$ be groups. A homomorphism $f : G \to H$ is a map such that

$$f(a \ast b) = f(a) \circ f(b), \ \forall a, b \in G.$$  

An isomorphism is a homomorphism that is also a bijection (one-to-one and onto). Two groups are isomorphic, denoted $G \cong H$, if there is an isomorphism $f : G \to H$. An isomorphism of a group onto itself is called an automorphism.

In other words, a homomorphism is a mapping that preserves the algebraic structure from $G$ to $H$. The isomorphism is exactly what we wrote in the definition, namely that the group $G$ is really $H$ in 'disguise'. That is, an isomorphism is a homomorphism that is one-to-one and onto. Homomorphisms preserve the operations, but elements that look sufficiently alike might be 'collapsed' into one element. Homomorphisms can however be perfect, and thereby isomorphisms! Although, for our purposes it suffices to think of homomorphisms as a map that preserves the algebraic structure of a group.

Earlier we talked about the group of invertible mappings, and an interesting special case of this is when the set is a vector space $V$ and the maps are all linear. Here we're introduced to a new funny word, namely an endomorphism, which is simply a homomorphism that is onto.

**Definition 1.4.** The **general linear group** of a vector space $V$, denoted $GL(V)$, is the set of all invertible endomorphisms of $V$. In particular, when $V = \mathbb{C}^n$, we usually write $GL(n, \mathbb{C})$ instead of $GL(\mathbb{C}^n)$ with similar notation for $\mathbb{R}$.

In plain language, the general linear group is the set of all invertible $n \times n$ matrices together with the binary operation of ordinary matrix multiplication. This is because we're really talking about linear transformations when we talk about the set of all invertible endomorphisms of the vector space $V$, and we know that all linear transformations can be represented by matrices!
1.2 Subgroups

Given some subset $S \subset G$, it is not far fetched to say that $S$ is a subgroup of $G$ if it is a group under the very same binary operation under which $G$ is a group.

**Definition 1.5.** A subset $S$ of a group $G$ is a subgroup of $G$ if it is a group in its own right under the binary operation of $G$, i.e., if it contains the inverse of all its elements as well as the product of any pair of its elements.

Lets look at some examples.

**Example 1.3.**

1. For any $G$, the subset $\{e\}$, consisting of the identity alone, is a subgroup of $G$ called the trivial subgroup of $G$.

2. Because $\mathbb{Z} \subset \mathbb{R}$, it is clear that $(\mathbb{Z}, +)$ is a subgroup of $(\mathbb{R}, +)$.

3. The set $\{2n\}$ where $n \in \mathbb{Z}$, i.e. the set of even numbers, is a subgroup of $(\mathbb{Z}, +)$. It turns out that all subgroups of $\mathbb{Z}$ are on the form $\mathbb{Z}m$ where $m$ is a positive integer.

4. The subset of $GL(n, \mathbb{C})$ consisting of transformations that have unit determinant is a subgroup of $GL(n, \mathbb{C})$ because the inverse of a transformation with unit determinant also has unit determinant, and the product of two transformations with unit determinants has unit determinant. That is, if $\det(A) = 1$, then $\det(A^{-1}) = 1$ and $\det(AA^{-1}) = 1$. This subgroup is denoted $SL(n, \mathbb{C})$ where $SL$ stands for special linear group.

5. The set of unitary\(^3\) transformations of $\mathbb{C}^n$, denoted $U(n)$, is a subgroup of $GL(n, \mathbb{C})$ because the inverse of a unitary transformation is also unitary and the product of two unitary transformations is unitary.

**Definition 1.6.** The set of unitary transformations $U(n)$ is a subgroup of $GL(n, \mathbb{C})$ and is called a unitary group. Similarly, the set of orthogonal transformations of $\mathbb{R}^n$ is a subgroup of $GL(n, \mathbb{R})$. It is called the orthogonal group and is denoted $O(n)$.

In simpler terms, the unitary group consists of complex unitary $n \times n$ matrices such that $U^\dagger U = 1$. There are special cases of both the unitary and the orthogonal group called the special unitary group and the special orthogonal group. These are denoted $SU(n)$ and $SO(n)$, respectively. $SU(n)$ comes about from the unitary  

\(^3\)A matrix $U$ is unitary if $U^\dagger U = 1$, i.e., $U$ is a unitary transformation.
group if we demand that \( \det(U) = 1 \). The group \( SO(n) \) is commonly referred to as rigid rotations of \( \mathbb{R}^n \). These groups are more advanced than they appear, and will be properly introduced in the section about representations.

In physics, we like things that are invariant under some form of change. A special type of subgroup called a normal subgroup is important in this respect because it is invariant under some transformation \( gHg^{-1} = H \) where \( H \) is the normal subgroup, \( H \subset G \) and \( g \) is some element in \( G \). This notation implies that an element \( g \) of a group \( G \) acts on every element of it’s subgroup \( H \). In general we call the set of all such products a coset. If the group the element acts on happens to be a normal subgroup, i.e. the left cosets are the same as the right cosets, one can actually multiply the cosets which brings about another group called the quotient group \( G/H \) with the members \( \{ h \in H : hg = gh, \forall g \in G \} \).

**Definition 1.7.** Let \( G \) be a group and \( H \) a subgroup of \( G \). \( H \) is a normal subgroup if and only if \( \forall g \in G \)

\[
gHg^{-1} = H.\]

\( G \) is said to be **simple** if it only contains the trivial normal subgroup \( H = \{ e \} \). Furthermore, \( G \) is said to be **semi-simple** if it has no abelian subgroups.\[3\]

**Definition 1.8.** Let \( H \) be a subgroup of a group \( G \) and \( g \) be any element of \( G \). The left coset of \( gH \) is defined to be the set of group elements of the form \( gh \), as \( h \) ranges over the elements of \( H \)

\[
gH = \{ h \in H : gh \}.\]

Similarly, for the right coset

\[
Hg = \{ h \in H : hg \}.
\]

If \( H \) is a normal subgroup of \( G \), the set of left cosets of \( H \) in \( G \) is a group denoted by \( G/H \) under the operation

\[
(gH) \circ (g'H) = (g \circ g')H.
\]

It might be more satisfying to write the elements \( h \) of \( H \) instead of the whole set of elements as we have done. But when we know that one element \( g \in G \) acts on all elements of \( H \) it is sufficient to write \( H \).
1.3 Group Action

Earlier we introduced what we called transformation groups. In the language of abstract group theory, we can generalize these kinds of transformations.

**Definition 1.9.** Let $G$ be a group and $M$ a set. The left action of $G$ on $M$ is a map $\Phi : G \times M \to M$ such that

1. $\Phi(e, m) = m, \forall m \in M$;
2. $\Phi(g_1 g_2, m) = \Phi(g_1, \Phi(g_2, m))$.

One usually denotes $\Phi(g, m)$ by $g \cdot m$ or $gm$. A subset $N \subset M$ is called a left invariant if $gm \in N, \forall g \in G, \text{ whenever } m \in N$. The right action is defined similarly, but with $mg \in N$.

While this definition is fine, I find that the use of the group of permutations gives a much much more intuitive idea of what a group action is. The symmetric group $S_n$ permutes the elements of a group. Say $X = \{1, 2, 3\}$, then a permutation of the elements can be $\{2, 3, 1\}$ which constitutes the permutation group $S_3^4$. What is the operation? Well, 1 has been switched with 2, 2 with 3 and 3 with 1. This reminds us of a function where $f(1) = 2, f(2) = 3$ and $f(3) = 1$. More specifically, the symmetric group(group of permutations) is a bijection from the set $X$ to itself. That is,

$$f : X \to X.$$ 

Multiplication in this group is simply the composition of functions, so we multiply two permutations $f$ and $g$ by $f \circ g$. If $g$ permutes the elements from $\{1, 2, 3\}$ to $\{3, 1, 2\}$, the composition of $f$ and $g$, $f \circ g = f(g(x))$ where $x \in X$, yields the identity element.$^5$

**Definition 1.10.** Let $G$ be a group and $X$ a set. Let $S(X)$ be the set of all permutations of $X$. An action of $G$ on $X$ is a homomorphism $\phi : G \to S(X)$.

So, a group action is really a group of transformations on some space $X$. In particular a group action of a group $G$ on a set $X$ permutes the elements of $X$.

**Definition 1.11.** Let $G$ act on $M$ and let $m_0 \in M$. The orbit of $m_0$, denoted by $Gm_0$, is

$$Gm_0 = \{m \in M | gm_0 = m \text{ for some } g \in G\}$$

$^4$On a set of 3 elements.

$^5$Check that this is true!
The action is called transitive if \( Gm_0 = M \). The stabilizer of \( m_0 \) is
\[ G_{m_0} = \{ g \in G \mid gm_0 = m_0 \} \]. The group action is called effective if \( gm = m \) for all \( m \in M \) implies that \( g = e \).

Looking at definition 1.11 doesn’t necessarily do anything for you, so to get a better feel for what an orbit is, it is necessary to look at a couple of examples.

**Example 1.4.**
1. Let \( M = \mathbb{R}^2 \) and \( G = \text{SO}(2) \), the group of rotations in the plane. The action is rotation of a point in the plane about the origin by an angle \( \theta \) which makes orbits as circles centered at the origin. The action is effective but not transitive. The stabilizer of every point in the plane is \( \{ e \} \), except at the origin, for which the whole group is the stabilizer.

2. Let \( M = S^1 \), the unit circle, and \( G = \text{SO}(2) \), the group of rotations in two dimensions. The action is the displacement of a point on the circle with one unique orbit, i.e. the entire circle. The action is effective and transitive and the stabilizer of every point on the circle is \( \{ e \} \).

### 1.4 Useful definitions

Two representations can be equivalent, and it is important to know how, so this definition states what it means for two representations to be *equivalent*.

**Definition 1.12.** Let \( V \) and \( W \) be two vector spaces over the field \( F \). Then, the two representations \( \rho : G \to GL(V) \) and \( \pi : G \to GL(W) \) are said to be **equivalent** if there exists an isomorphic map \( \alpha : V \to W \) such that \( \forall g \in G, \)
\[ \alpha \rho(g) \alpha^{-1} = \pi(g). \]

In addition to the direct product, which given a set \( A \) and \( B \), defines the product \( A \times B \) of \( A \) and \( B \) that produces the set of ordered pairs \( (x, y) \) where \( x \in A \) and \( y \in B \), we need another type of product called the semi-direct product.

**Definition 1.13.** Let \( H \) and be a group and \( G \) a map \( G : H \to H \). The component-wise operation on the ordered pair \( (g, h) \), where \( g \in G \) and \( h \in H \), \( (g, h) \circ (g_j, h_j) = (g_j \circ g, g_j(h_j)) \) denoted \( G \times H \) is called the **semi-direct product** between \( G \) and \( H \).
1.5 Representation theory

Let $G$ be a finite group and $\mathcal{V}$ be a finite dimensional vector space over some field $\mathbb{F}$ with dimension $\text{dim} \mathcal{V}_\mathbb{F} = n$. That is, $\mathcal{V} \cong \mathbb{F}^n$. It is also important to recall that in definition 1.4 we introduced the general linear group of a vector space $\mathcal{V}$, which we denote $GL(\mathcal{V})$. This is the group of linear invertible mappings from the vector space $\mathcal{V}$ onto itself, i.e. $\mathcal{V} \rightarrow \mathcal{V}$.

**Definition 1.14.** The pair containing the map $\pi$ and the vector space $\mathcal{V}$, $(\pi, \mathcal{V})$, is a representation of the finite group $G$ if $\pi : G \rightarrow GL(\mathcal{V}) \cong GL(n, \mathbb{F})$ is homomorphic such that

$$\forall g_i, g_j \in G, \quad \pi(g_i g_j) = \pi(g_i) \pi(g_j).$$

Now we need to take a step back and examine this a little bit. For this to be a representation, $\pi$ must be a homomorphism which means that it must be a map such that the algebraic structure between two groups is preserved somehow. For us this means that $\pi$ will simply transport the group structure of $G$ to the group structure of $GL(\mathcal{V})$, like for example $GL(n, \mathbb{C})$. We can also look at this with the familiar linear transformation. That is, a representation can be viewed as the group action on a vector space. This is explicitly done by the linear transformation given by

$$\pi(g_1, g_2)v = [\pi(g_1)\pi(g_2)]v = \pi(g_1)[\pi(g_2)v]$$

Now, why would we do this? Why would we put ourselves through this and try to understand this abstract concept? There are a lot of advantages of studying representations. The one that is most important for physicists is that if $\pi$ is one-to-one, the group $G$ is identified as a subgroup of $GL(n, \mathbb{C})$. This means that we’re simply looking at a group of $n \times n$ invertible matrices due to the linear transformation\(^6\)! We have, in a way, condensed group theory down to linear algebra which is easier and more practical to work with. In fact, we’re still allowed to use concepts from linear algebra even though $\pi$ is not one-to-one. Let’s look at some examples.

**Example 1.5.** 1. Let $G$ be any group and $\mathcal{V} = \mathbb{C}$. This means that $\pi$ will carry our group to $\mathbb{C}^*$. 

$$\pi : G \rightarrow GL(1, \mathbb{C}) \cong \mathbb{C}^*$$

where $\mathbb{C}^*$ is $\mathbb{C} \setminus \{0\}$. This is the trivial representation, so $\pi$ will carry every group element to $1$. In other words, $\pi(g) = 1$. Thus,

$$\pi(g_1 g_2) = \pi(g_1)\pi(g_2) = 1 \cdot 1 = 1.$$  

\(^6\)As we know, any linear transformation can be represented by a matrix.
So this means that any group action, i.e. any element of $G$ that acts on any element (complex number) in $GL(1, \mathbb{C})$ will take a complex number and send it back to itself. This is the trivial group action.

Another useful example includes the symmetric group of permutations on the labels $\{1, 2, 3\}$.

**Example 1.6.** Let $G = S_3$, the symmetric group of permutations on the labels $\{1, 2, 3\}$ and $V = \mathbb{C}^3$. The identity element $e$ is represented by the matrix

$$123 \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

So 1 maps to the first column vector, 2 maps to the second column and 3 maps to the third column. Now we can write down the six permutations of the group.

$$123 \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad 132 \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad 231 \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$213 \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad 312 \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad 321 \mapsto \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

The representation is the column vectors 1, 2 and 3 while the group elements of $S_3$ are the matrices. This is a special case where $n = 3$, the more general case of the symmetric group is when $S_n$ is not specified.

Although we have used the Lie groups $SU(n)$ and $SO(n)$ quite sloppily in the former sections we will not do it here as an example of a representation. They are, but we’ll introduce them properly soon. Another important representation is the *adjoint representation* that will make more sense after we’ve introduced Lie groups and Lie algebras.

**Example 1.7.**

### 1.5.1 Irreducible representations

Up until now we’ve used $V$ for our vector space, but we’ll continue with $V$ instead because it’s simpler to write, so don’t be alarmed! The representations we will be working with are ones that cannot be broken down into smaller pieces. In other words, a non-zero finite representation $(\pi, V)$ that has no proper subrepresentations
(\pi, W), where W ⊂ V, is called irreducible. In order to understand the last sentence, we must first investigate what a subrepresentation is, and this involves invariant subspaces. Also, it is important to understand irreducible representations because of their physical significance!

**Definition 1.15.** Let G be a group. Let V be a vector space and let V' be a subspace of V denoted V' ⊆ V. Then, V' is called invariant if \( \pi(g)V' \subseteq V' \), \( \forall g \in G \).

I.e., if \( \pi(g) \) carries \( V' \) back to itself, then we’ll way that \((\pi, V')\) is a subrepresentation of \((\pi, V)\). Now we have a representation inside a larger representation. No matter what representation we have, we will always have two invariant subrepresentations, namely \( 0 \) and \( V \) itself. If there are no other invariant subspaces we say that the representation \((\pi, V)\) is irreducible.

**Definition 1.16.** Let \( \pi \) be a homomorphism \( \pi : G \rightarrow GL(V) \) such that \((\pi, V)\) is a representation. \((\pi, V)\) is said to be irreducible if there exists no other invariant subspaces(subrepresentations) than \( 0 \) and \( V \). If such invariant subspaces exists, \((\pi, V)\) is said to be reducible.

Now, I said that this is important, but I didn’t say why. Elementary particles are irreducible in the same sense that representations are irreducible, so when physicists talk about irreducible representations of the symmetry group of the theory, they talk about particles. Well, we classify particles, so how would we classify the corresponding irreps? We do that through **Casimir operators** for spin and mass, which we will get back to later.

### 1.6 Lie Groups and Lie Algebras

In order to explain what Lie groups and algebras are, we cannot strictly stay in the domain of group theory. We have to include elementary concepts from differential geometry as well. Specifically we have to include the concept of a manifold. This really takes us back to set theory, because a manifold is a collections of points. These points form special kinds of sets and so it is natural to step into the realm of topology as well. Usually we have to define a metric on a set in order to be able to talk about distances between points for example, but in a topological space we discard the distance function while keeping the open set of the metric space\(^8\).

---

\(^7\)Irreps is just short for irreducible representations.

\(^8\)A metric space consists of an open set \( X \) (does not contain it’s boundary points) on which a distance function \( d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \) is defined. I.e. \((X, d)\) is the metric space.
Definition 1.17. Let $X$ be non-empty set and $\mathcal{J}$ a collection of subsets of $X$ such that:

1. $X \in \mathcal{J}$.
2. $\emptyset \in \mathcal{J}$.
3. If $O_1, O_2, \ldots, O_n \in \mathcal{J}$, then $O_1 \cap O_2 \cap \ldots \cap O_n \in \mathcal{J}$.
4. If for each index $\alpha \in I$, $O_\alpha \in \mathcal{J}$, then $\bigcup_{\alpha \in I} O_\alpha \in \mathcal{J}$.

The pair $(X, \mathcal{J})$ is called a topological space. The set $X$ is called the underlying set and $\mathcal{J}$ is called the topology on $X$. The members of $\mathcal{J}$ are open sets.

It is often said that the metric space gives rise to the topological space and likewise that the topological space is associated with the metric space of which it has risen[4]. This might seem abstract and unnecessary, but it is important to understand topological spaces in order to understand manifolds, which is what we will use in our definition of Lie groups!

Example 1.8. 1. The metric space $(\mathbb{R}^n, d)$ gives rise to a topological space which we will call the Euclidean $n$-space with the usual topology.

2. Let $X$ be an arbitrary set. Let $\mathcal{J}$ be the set consisting of the empty set and $X$, i.e., $\mathcal{J} = \{\emptyset, X\}$. Then, $(X, \mathcal{J})$ is a topological space.

3. Let $\mathbb{Z}$ be the set of positive integers. For each positive integer $n$, let $O_n = \{n, n+1, n+2, \ldots\}$. Let $\mathcal{J} = \{\emptyset, O_1, O_2, \ldots, O_n, \ldots\}$. Then $(\mathbb{Z}, \mathcal{J})$ is a topological space.

Topological spaces are really sets where we can define (among other things) neighbourhoods and limits. We haven’t defined continuity etc. in the domain of topology, so it would be beneficial for the reader to look in to those things⁹. So now that we have looked more closely at what a topological space is, we can begin to talk about the concept of a manifold. A manifold is really a set of points, with some additional structure. In fact, it is a topological space in which you can label the points in the set by coordinates, i.e. real numbers $\{x_1(p), x_2(p), \ldots, x_n(p)\} \in \mathbb{R}$ at each point $p \in M$ to give us the location of that point¹⁰. This is the property that distinguishes a manifold from a normal topological space.

⁹Introduction to Topology by Bert Mendelson does this wonderfully.

¹⁰The points of a manifold in General Relativity also takes into account the time, so it not only tells you where you are at $p$, but also when. That is, the points are now events.
1.6.1 Pre-discussion of manifolds including physics

When we want to find the motion of an object through space-time, we want to find a parametric description of a curve in 3 spatial dimensions and one for \( t \), thus a curve in \( \mathbb{R}^4 \) given that time is a coordinate like \( x(t), y(t) \) and \( z(t) \). However, the laws of motion in space-time is not the same for all observers unless they are what we call *inertial frames*. In Newtonian mechanics we use the Galileo transformation to connect the frames, but in special relativity we use the Lorentz transformation to connect them. The space-time manifold is a flat affine manifold \( \mathbb{R}^4 \), where affine simply means that it is a vector space as well. As we said earlier, we can cover the points of \( \mathbb{R}^n \) by a coordinate frame \( \{ x^i \} \), these frames are related to one another by the elements of the general linear group. The elements are general linear transformations \( L^i_j \in GL(n, \mathbb{R}) \), thus

\[
x'^i = L^i_j x^j.
\]

But what about when the space-time manifold is not flat? A typical example of a non-flat manifold is a the hypersphere \( S^n \). We cannot do as we did in the case of the flat space-time, i.e. we cannot represent all the points of a curved manifold with a single coordinate frame \( \{ x^i \} \), that we will name a *chart*. But what about many charts? Much in the same way Democritus reasoned that since the world is dynamic, it must be comprised of smaller parts called atoms, we can make a dynamic(meaning curved in our case) surface by the means of many charts put together. We call this collection of charts an *atlas*. Each chart maps one open region of the curved surface, and so the collection of these charts, i.e. the union, makes up the entire surface(manifold). Like in the flat case, we must have a way to map from one chart to the next, and this includes the regions where the charts overlap. This enables us to get a complete coordinate description of our surface through our beloved atlas.

We can, however, run into a problem. Namely the fact that a manifold can be described by many atlases. Since this is a fact of life that we cannot do anything about, we must make it so that the *universal* laws of physics also accommodates curved space-time and that they are the same in *every* atlas. This is where the vector space structure comes in handy, because the inertial frames correspond to a vector at each position of the manifold. In this way, all observers experience the same laws of physics in all reference frames. This principle is called *the principle of general covariance*. 
1.6.2 Differentiable Manifolds

One of the important things to keep in mind from the section above is the fact that these higher dimensional 'blobs' we call manifolds are locally Euclidean. We need to introduce a new concept in order to be able to give a good description of a manifold, namely that of a homeomorphism.

**Definition 1.18.** Topological spaces \((X, \mathcal{J})\) and \((Y, \mathcal{J}')\) are called homeomorphic if there exist inverse functions \(f : X \to Y\) and \(g : Y \to X\) such that \(f\) and \(g\) are continuous. In this event the functions \(f\) and \(g\) are said to be homeomorphisms and we say that \(f\) and \(g\) define a homeomorphism between \((X, \mathcal{J})\) and \((Y, \mathcal{J}')\).

The crucial thing about homeomorphisms is that they preserve all the topological properties of the sets. These topological spaces we call manifolds are also Hausdorff, so we’ll have to define what we mean by that.

**Definition 1.19.** Let \((X, \mathcal{J})\) be a topological space. \((X, \mathcal{J})\) is said to be Hausdorff if \(\forall x, y \in X, \exists\) open subsets \(A, B\) where \(A \cap B = \emptyset\) such that \(x \in A\) and \(y \in B\).

This might seem complicated, but it’s not! Given two points in \(X\), we can separate those two points by two open subsets \(A\) and \(B\) that does not intersect. So \(x\) is only in \(A\) and \(y\) is only in \(B\). Examples of such spaces are metric spaces. Our manifolds will also have this property. We talked about charts and atlases, so let’s formalise them.

**Definition 1.20.** An \(m\)-dimensional chart for \(M\) is a bijective homeomorphic map \(\phi : U \to V \subset \mathbb{R}^m\) where \(U \subset M\). A chart is denoted by the pair \((U, \phi)\).

With this in place, we can define an atlas.

**Definition 1.21.** An \(m\)-dimensional atlas for \(M\) is a collection of charts \(\phi_\alpha : U_\alpha \to V_\alpha\), denoted \(\{(U_\alpha, \phi_\alpha)\}\), such that the images cover the whole of \(M\):

\[
M = \bigcup_\alpha U_\alpha.
\]

Note that \(V_\alpha = \phi_\alpha(U_\alpha)\). When reading the definitions of the atlas and chart, it is perhaps useful to look at figure 1. There is a one-to-one correspondence between arrays \((x^1, \ldots, x^m) \in V \subset \mathbb{R}^m\) and points \(x \in U \subset M\) given by the maps \(\phi\) and \(\phi^{-1}\). What this means is that we can call a chart a local coordinate system on \(M\). The
The notion of locality is important here, because it only makes sense for a small region of $M$, i.e. a subset $U$ of $M$.

To get a complete description, we need to address the intersection of $U_\alpha$ and $U_\beta$ shown by the shaded areas in Figure 1. Let $U_\alpha \cap U_\beta \neq \emptyset$, the sets $V_\alpha$ and $V_\beta$ contains subsets $\phi_\alpha(U_\alpha \cap U_\beta) \subset V_\alpha$ and $\phi_\beta(U_\alpha \cap U_\beta) \subset V_\beta$. The important thing is that any point $x \in U_\alpha \cap U_\beta$ has two coordinate descriptions $\phi_\alpha(x)$ and $\phi_\beta(x)$, and we can therefore change coordinates between the two charts. To make life less complicated we will define $U_{\alpha\beta} \equiv U_\alpha \cap U_\beta$. This means that there is an invertible mapping

$$\Phi_{\alpha\beta} \equiv \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_{\alpha\beta}) \subset \mathbb{R}^m \to \phi_\beta(U_{\alpha\beta}) \subset \mathbb{R}^m,$$

$$(x^1_\alpha, \ldots, x^m_\alpha) \mapsto (x^1_\beta, \ldots, x^m_\beta). \tag{2}$$

Not only must they be Hausdorff, but they must also be differentiable. What do we mean by that? The differentiable structure is really given by the atlas, as we will now define.

**Definition 1.22.** An atlas $\mathcal{A} = \{(U_\alpha, \phi_\alpha)\}$ is said to be differentiable or smooth if it consists of pairwise compatible charts, which means that all sets $\phi_\alpha^{-1}(U_{\alpha\beta})$ are open and the changes of coordinates $\Phi_{\alpha\beta} \equiv \phi_\alpha^{-1} \circ \phi_\beta$ are given by differentiable functions. We denote this a $C^\infty$-atlas.

The differentiable functions are denoted $C^k$. That is, they are continuously differentiable of all orders less than or equal to $k$. We will be concerned with $C^\infty$ which we call smooth. Two equivalent atlases $\mathcal{A}$ and $\mathcal{A}'$ have the property that $\mathcal{A} \cup \mathcal{A}'$ is again a $C^\infty$-atlas. This defines a equivalence relation on $C^\infty$-atlases. A differentiable structure $D$ on $M$ is said to be a equivalence class of the equivalence relation and the collection of atlases related to $D$, $\mathcal{A}_D$, is said to be the maximal atlas for $D$. And now for the grand reveal.

**Definition 1.23.** A $C^\infty$-differentiable manifold of dimension $m$ is a topological Hausdorff space together with a maximal atlas $A_D = \{(U_\alpha, \phi_\alpha)\}$, such that

1. $\bigcup \alpha U_\alpha = M$.
2. $\phi_\alpha : U_\alpha \subset M \to V_\alpha \subset \mathbb{R}^m$ is a homeomorphism.
3. The map $\Phi_{\alpha\beta}$ from $\phi_\alpha(U_{\alpha\beta})$ to $\phi_\beta(U_{\alpha\beta})$ is $C^\infty$ (infinitely differentiable)$^{11}$.

$^{11}$The shaded regions of $V_{\alpha,\beta}$. 

17
When we say manifold from here on what we mean is a $C^\infty$-differentiable manifold, or a smooth manifold. To get a better feel for what a manifold is in practice, we can take a look at some examples.

**Example 1.9.**  
1. The trivial example of a manifold is the real line itself, $\mathbb{R}$. The more general trivial example is $\mathbb{R}^m$ because this does in fact look like $\mathbb{R}^m$ both globally and locally.

2. The $n$-sphere is another example of a manifold of dimension $n$. It is defined as $S^n = \{ x = (x_1, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} | x_1^2, \ldots, x_{n+1}^2 = 1 \}$. The circle, namely $S^1$ and the surface on a sphere $S^2$ are quite useful examples.

3. The 2-torus, denoted $T_2$, is a special case of the more general $n$-torus. It is formed by bringing together opposite sides of a square.

4. This is a highly non-trivial example. The general linear groups for $\mathbb{R}^n$ and $\mathbb{C}^n$ are also examples of manifolds. They are, in fact, Lie groups! For any two positive integers $n$ and $m$ let $\mathbb{R}^{m \times n}$ be the vector space of all $m \times n$ matrices. Due to the fact that $\mathbb{R}^{mn}$ is isomorphic to $\mathbb{R}^{m \times n}$, it is given the topology of $\mathbb{R}^{mn}$. By definition, the general linear group $GL(n, \mathbb{R})$ takes the form

$$GL(n, \mathbb{R}) \equiv \{ A \in \mathbb{R}^{n \times n} | \det(A) \neq 0 \} = \det^{-1}(\mathbb{R} \setminus \{0\})$$  \hspace{1cm} (3)

Since

$$\det : \mathbb{R}^{n \times n} \to \mathbb{R}$$  \hspace{1cm} (4)
is continuous, we know that $\mathbb{R}^{m \times n} \simeq \mathbb{R}^{n^2}$ contains an open subset $GL(n, \mathbb{R})$. Ultimately this means that $GL(n, \mathbb{R})$ is a manifold. In the case of the general linear group $GL(n, \mathbb{C})$ we have a set of complex $n \times n$ matrices. This time $GL(n, \mathbb{C})$ is an open subset of the vector space of $n \times n$ complex matrices $\mathbb{C}^{n \times n} \simeq \mathbb{R}^{2n^2}$. If we follow the same arguments as we did for $GL(n, \mathbb{R})$ we can see that $GL(n, \mathbb{C})$ is a manifold of dimension $2n^2$ [5].

After what we’ve seen in this text, it seems as though Lie groups are nothing but the general linear groups and other matrix-groups, and we can just pack our bags and move on. That is not the case[6].

1.6.3 Why should we study Lie groups and Lie algebras?

Most of the things we know in physics follow from symmetry. The basic symmetries that we see in nature are the Poincaré transformations consisting of the Lorentz transformations + translations. These transformations are so powerful that they can tell us what particles exists in nature, and we’ve seen a lot of them! How do we classify them? We use their mass. For particles of mass $m \neq 0$, we can use another number that we call spin. For particles of mass $m = 0$, we use a different number called helicity in order to classify them. In fact, most particles at high energies in the standard model are massless. In the sector of massive particles we have leptons and quarks, while in the massless sector we have what we call Gauge bosons. The Gauge bosons with helicity $\pm 1$ we call the photon and the gluon. When we scatter particles at the LHC we assume unitarity, and by also assuming the Poincaré symmetry we can conclude that the coupling constants of the interaction between gluons are only consistent if they satisfy what is known as the Jacobi identity which we will get back to when we talk about Lie algebras. The coupling constants, also called the structure constants, thereby becomes the elements of the Lie algebra. What kind of particles that can exist and the interactions between them are all controlled by symmetry which makes Lie groups and algebras highly important.
1.6.4 Lie Groups

Definition 1.24. A Lie group is a non-empty set $G$ satisfying the following conditions:

1. $G$ is a group under the operation $G \times G \to G$, $(g_1, g_2) \mapsto g_1 \circ g_2$ and $G \to G$, $g \mapsto g^{-1}$, where the maps are smooth.

2. $G$ is a smooth manifold.

We usually just write the product $g_1 \circ g_2$ as $g_1 g_2$.

Example 1.10. 1. The unit circle in the complex plane is denoted by $S^1 = \{e^{i\theta} \mid \theta \in \mathbb{R} \text{ (mod } 2\pi)\}$. If we multiply and take the inverse of the group elements we get $e^{i\theta} e^{i\phi} = e^{i(\theta + \phi)}$ and $(e^{i\theta})^{-1} = e^{-i\theta}$, both of which are differentiable. Thus $S^1$ is a Lie group that we denote $U(1)$.

2. The matrix groups are of great importance in modern theoretical physics, and we’ve already talked a lot about them. Because of their importance, we’ll go over them briefly again. As we know, the product of the elements are simply ordinary matrix multiplication. We’ve also said that the matrices are invertible, so the inverse is just that. There are $n^2$ coordinates in $GL(n, \mathbb{R})$ given by $M = \{x_{ij}\}$. Some of the subgroups of $GL(n, \mathbb{R})$ are

(a) Orthogonal group $O(n) = \{M \in GL(n, \mathbb{R}) \mid MM^T = M^T M = I_n\}$

(b) Special linear group $SL(n, \mathbb{R}) = \{M \in GL(n, \mathbb{R}) \mid \det M = 1\}$

(c) Special orthogonal group $SO(n) = O(n) \cap SL(n, \mathbb{R})$

The Lorentz group is also a Lie group. That is,

$O(3, 1) = \{M \in GL(4, \mathbb{R}) \mid M\eta M^T = \eta\}$.

where $\eta = \text{diag}(-1, 1, 1, 1)$ is the Minkowski metric. We also have subgroups of $GL(n, \mathbb{C})$:

(a) Unitary group $U(n) = \{M \in GL(n, \mathbb{C}) \mid MM^\dagger = M^\dagger M = I\}$

(b) $SL(n, \mathbb{C}) = \text{exactly as before but with } M \in GL(n, \mathbb{C})$.

(c) Special unitary group $SU(n) = U(n) \cap SL(n, \mathbb{C})$.

While Lie groups are important, we will spend a lot of time looking at its corresponding algebra.
1.6.5 Why bother with Lie Algebras?

Professor Robert Gilmore of Drexel University has a wonderful discussion of why one should in fact bother with Lie algebras[7], from which the following is inspired. The composition law of Lie groups can be extremely difficult to work with due to its non-linear nature. If one could somehow linearize it, while still clinging to its original composition information we would benefit greatly. It turns out that this is possible through Lie algebras, which linearizes Lie groups. This process is simplified because of the homogeneity of Lie groups, i.e. every point on the manifold is locally equivalent to every other point, in fact, we only need to choose one point because they are locally the same. The linearization of a Lie group goes on in the neighbourhood of this point, and consists of Taylor expanding the group composition function around the points that defines the group operation. However, there is one special point, namely the identity. It is of no geometrical difference to the other points, only algebraic. If we linearize the Lie group about the identity point, we get a new set of operators in the process which are the objects that form a Lie algebra. Because the Lie algebra is a linear vector space, vector addition is what constitutes the composition of two group operations around the identity element. We can even construct more sophisticated group products, like the commutator, which do in fact hold information from the original group composition law.

The simple answer to the question in bold above is that we can study Lie groups through the more simple linear vector space called the Lie algebra while still retaining the information from the Lie group. It is important also to state that the global properties of the Lie group are destroyed in the invertible linearization, so we only conserve the local properties about the identity element. There is, however, a way to get back to the Lie group, sort of, from the algebra which we will talk about shortly.
1.6.6 Lie Algebras

**Definition 1.25.** Let $F$ be a field and $g$ a $F$-vector space. Then $g$ is a **Lie algebra** if there exists a mapping $[\cdot, \cdot] : g \times g \to g$, named the Lie bracket, such that

1. It is anti-symmetric:
   \[ [X, Y] = -[Y, X], \quad \forall X, Y \in g. \tag{5} \]

2. It is bilinear:
   \[
   [aX + bY, Z] = a[X, Z] + b[Y, Z] \\
   \text{and} \\
   [Z, aX + bY] = a[Z, X] + b[Z, Y], \quad \forall X, Y, Z \in g. \tag{6}
   \]

3. The **Jacobi Identity** is satisfied:
   \[ [X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0, \quad \forall X, Y, X \in g. \tag{7} \]

So it is really a vector space over some field $F$, for example $\mathbb{R}$ or $\mathbb{C}$, with a binary operation we call the Lie bracket, which has to satisfy some specific conditions. In most cases relevant to physics the Lie bracket will simply be our beloved commutator $[X, Y] = XY - YX$. Some texts refer to the anti-symmetric property as **skew-symmetric**. Although we’ll we working with the commutator, it can be enlightening to see other examples of Lie algebras as well as the ‘usual’ ones.

**Example 1.11.** 1. $\mathbb{R}^3$ endowed with the Lie bracket $[u, v] = u \times v$, commonly known as the cross product.

2. The Heisenberg Lie algebra $\mathcal{H}$ over some field $F$ is generated by the matrices
   \[
   X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.
   \]
   where $[Y, X] = Z$ and $[Y, Z] = [X, Z] = 0$. One can also check that they satisfy the conditions given in the definition.

Although I have not shown explicitly how to linearize some Lie group, it is important to show how to get back to the Lie group even though the global information was lost in the initial linearization. That is, how to invert the linearization map. We do this by the use of the **exponential function**. Lie theory is about symmetries that
are continuous. So what’s the difference between a continuous and a discontinuous symmetry? A continuous symmetry implies that there are group elements arbitrarily close to the identity transformation of the group, and these transformations change nothing, hence the symmetry. The set of symmetry transformations of a square cannot contain the rotation of an angle close to 0 degrees. It is simply not part of that set. However, a circle can rotate by an extremely small circle, and it would still be a symmetry transformation, and there are are arbitrarily many rotations for the circle. We can write an element $g$ close to the identity $I$ in the Lie group by the small real number $\epsilon$ and some operator $X$, which is is also called a generator\(^{12}\) because in some sense it generates a finite transformation, in the Lie algebra as

$$g(\epsilon) = I + \epsilon X. \quad \text{(8)}$$

This transformation doesn’t do much to the group element, however, if we apply it $k$ times we get a finite transformation.

$$g(\epsilon)^k = (I + \epsilon X)(I + \epsilon X) \ldots (I + \epsilon X)^k. \quad \text{(9)}$$

Let $k$ be a really big number, $\theta$ some finite transformation and $\epsilon = \theta/k$. Then, because we want to consider the smallest possible transformations, i.e. infinitesimal ones, we have to apply the transformation an infinite amount of times in order to get a resulting finite transformation as $k \to \infty$. We write this as

$$h(\theta) = \lim_{k \to \infty} g(\theta)^k = \lim_{k \to \infty} (I + \frac{\theta}{k} X)^k = \sum_{n=0}^{\infty} \frac{(\theta X)^n}{n!} = e^{\theta X}. \quad \text{(10)}$$

Lets now consider a matrix group of continuous transformations. If we Taylor expand an element about the identity we get

$$h(\theta) = I + \frac{dh}{d\theta}\bigg|_{\theta=0} \theta + \frac{d^2h}{d\theta^2}\bigg|_{\theta=0} \theta^2 + \ldots = \sum_{n} \frac{d^n h}{d\theta^n}\bigg|_{\theta=0} \theta^n \quad \text{(11)}$$

$$= e^{\frac{dh}{d\theta}|_{\theta=0} \theta} = \sum_{n} \frac{d^n h}{d\theta^n}\bigg|_{\theta=0} \theta^n, \quad \text{(12)}$$

where we can see that $X = \frac{dh}{d\theta}|_{\theta=0}$. The lesson here is that the teachings of a group can be brought forth by studying its generators. For matrix groups, then, we can actually define a Lie algebra as the $n \times n$ matrices $X$ such that $e^{tX}$ is an element of the Lie group $G$, where $t \in \mathbb{R}$. Although the combination rule, $\odot$, for matrix groups\(^{12}\)For example the linear combinations of the matrices that span the Lie algebra of $SL(n, \mathbb{R})$.\)
is matrix multiplication, one must not fall into the trap of thinking that this also applies for the elements of its algebra. The operation given above tells you that for matrix algebras, the difference between the products \( XY \) and \( YX \) is indeed a part of the Lie algebra, but \( XY \) and \( YX \) themselves need not be. In general the combination of two elements of a matrix Lie group will take the form

\[
e^X \circ e^Y = e^{X+Y+\frac{1}{2}[X,Y]+\frac{1}{12}[X,[X,Y]]-\frac{1}{12}[Y,[X,Y]]+...} \in G
\]

which includes the Lie bracket we saw in definition 1.23. The exponential maps the Lie algebra to the manifold of the Lie group, however, the mapping back to the Lie group is not complete in the sense that it does not map to the entire manifold[7].

1.6.7 Structure Constants

The Lie algebra of a Lie group is nothing but a linear vector space with a mapping \([,]\) called the Lie bracket. This means that we can define the usual quantities of a vector space such as the inner product and its basis vectors. Earlier we said that the dimension of a vector space is determined, for example, by the number of its basis vectors and so if we have a Lie group i.e. a manifold of dimension \( n \), we can say that we have an \( n \)-dimensional Lie algebra with \( n \) basis vectors, or basis operators denoted \( X_1, X_2, \ldots, X_n \), previously called generators. We know that we can express any vector in a vector space as a linear combination of its basis vectors, and since the commutator of two operators is in the Lie algebra, we can write

\[
[X_i, X_j] = C^k_{ij} X_k.
\]

The structure of a Lie algebra is a well defined thing, and it is determined solely by the structure constants \( C^k_{ij} \), which are anti-symmetric due to the commutator being anti-symmetric. That is, \( C^k_{ij} + C^k_{ji} = 0 \). But what is a structure constant? Well, it’s a constant, but not an ordinary constant. In short, the entire group multiplication law is practically summarized by the structure constants[8]. In other words, the structure constants specify the multiplication of two generators in the Lie algebra. The basis vectors of the Lie algebra are generators, and the multiplication law is given by the Lie bracket. Because the vector space \( g \), i.e. the Lie algebra, is linear we can extend this to all elements in \( g \), meaning that the Lie brackets are determined by the structure constants, \( \forall X \in g \).

The adjoint representation mentioned earlier is generated by the structure constants. This representation consists of a set of matrices \( T_i \) defined as

\[
(T_i)^k_j = -C^k_{ij},
\]
and since the structure constants satisfy the Jacobi identity,

\[ [T_i, T_j] = C^k_{ij} T_k. \] (17)

This means that the structure constants themselves provide the representation of the algebra. We can summarize by Lie’s Theorems.

**Theorem 1.1.** 1.

2. \([X_i, X_j] = C^k_{ij} X_k\) are satisfied by the generators \(X_i\) where \(C^k_{ij}\) are the structure constants.

3. The identity \(C^k_{ij} C^m_{kl} + C^k_{jl} C^m_{ki} + C^k_{li} C^m_{kj} = 0\) follows from the Jacobi identity. The structure constants are anti-symmetric: \(C^k_{ij} = -C^k_{ji}\).

### 1.6.8 The Dynkin Index

The German mathematician Wilhelm Karl Joseph Killing made contributions of great value to Lie theory. One of them is the Killing form. In order to properly understand Dynkin indices, we need to understand Killing forms.

### 1.7 The Poincaré Group and its Algebra

#### 1.7.1 The Lorentz Group and its Algebra

#### 1.7.2 The Poincaré Group and its Algebra

#### 1.7.3 Casimir Operators

### 2 Tensors

A lot of books that I know of in physics starts explaining what a tensor is by simply stating how it transforms under some change of basis. While this is not false, or an insufficient explanation for physicists, one cannot help but ask the question: But what is a tensor really? What does it do? ‘Well, they are generalisations of vectors that transforms as we told you’ physicists might say, and they’re not wrong. However, it is my intention to try to present a more satisfactory explanation of this concept and peer down, if not to the core, at least far enough down so that we’ll have fewer questions about the nature of tensors and get a good feeling when the topic comes up in a conversation! Because the reader and her/his way
2.1.1

of learning is unknown to me, I will not choose between an informal ‘physical’ explanation or a more rigorous mathematical explanation, I’ll do both, starting with the mathematical explanation\textsuperscript{13}.

2.1 Semi-Rigorous Mathematical Definition

Tensors can be characterised as functions that takes in vectors and dual vectors and spews out scalars, i.e. ordinary numbers on the real line. A question arises, what is a dual vector?

2.1.1 Dual Vectors

Vector spaces include a set of defined coordinate axes, one might even be given the freedom to choose them. In any case, the basis of the vector space is defined by the vectors parallel to the coordinate axes. Usually these basis vectors are of unit length and orthogonal to each other. A linear transformation, i.e. a function $f$ on some vector space $V$ must satisfy

$$
f(v + w) = f(v) + f(w)$$
$$f(\alpha v) = \alpha f(v)$$

(18)

for any vectors $v$, $w$ and scalar $\alpha$. Interestingly enough, any linear vector space $V$ has a similar partner $V^\ast$. I said similar, so what’s the difference? $V^\ast$ is simply the space of linear maps from $V$ to $\mathbb{R}$, i.e. it is a homomorphism from $V \rightarrow \mathbb{R}$ denoted $\text{Hom}(V, \mathbb{R})$. While $V$ consists of vectors, $V^\ast$ consists of linear functionals which are sometimes called covectors, dual vectors and sometimes one-forms\textsuperscript{14}.

That’s all well and good, but what do they do? Dual vectors are linear functionals, so they take in a vector and gives you something back. That something is a real number. For example, let $\beta \in V^\ast$ and $v \in V$, then $\beta(v) \in \mathbb{R}$. It is important to remember that $\beta$ is a linear function, not an arbitrary one. So dual vectors act on vectors.

The dual basis (the basis in $V^\ast$) has the same dimension as the basis in $V$, which

\textsuperscript{13}That’s not to say that the two ways of explaining tensors are fundamentally different in the sense that they explain different things. They do in fact explain the same object, but with some remarkably big differences of approach.

\textsuperscript{14}It is not only in relativity we find dual vectors. The bra-vector $\langle \psi |$ is also a dual vector, and so when we take the dual vector and act on a vector $| \psi \rangle$ we get the usual inner product from quantum mechanics $\langle \psi | \psi \rangle$. 

26
means that it has the same number of basis vectors. To elaborate on this we can pick a basis $e_1, \ldots, e_n \in V$ and $e^1, \ldots, e^n \in V^*$. The basis vectors with upper indices are linear functionals which act on vectors such that
\[
e^i(e_j) = \delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}
\] (19)

A linear functional or a covariant vector are vectors with lower indices given by $w = w_i e^i$ and if we switch to more familiar indices: $w = w_\mu e^\mu$. Vectors are called contravariant vectors with upper indices $V = V^\nu e_\nu$, and so if we let $w$ act on $V$ we get
\[
w(V) = w_\mu V^\nu e^\mu(e_\nu) \\
= w_\mu V^\nu \delta^\mu_\nu \\
= w_\mu V^\mu,
\] (20)

which is in fact the ordinary dot product. The dual space of a dual space is the original vector space itself in the sense that $V^{**} = V$, thus
\[
V(w) = w(V) = w_\mu V^\mu \equiv \langle w, V \rangle = \langle V, w \rangle.
\] (21)

In this way we can look at vectors as linear mappings from single dual vectors to real numbers[9]. This is the heterogeneous scalar product between a vector and a dual vector. For basis vectors in braket notation we get
\[
\langle e^\mu, e_\nu \rangle = \delta^\mu_\nu.
\] (22)

### 2.1.2 Dual Spaces In Terms of Row and Column Vectors

A vector $v \in V$ can be written as a column vector,
\[
\begin{pmatrix}
  v_1 \\
v_2 \\
\vdots \\
v_n
\end{pmatrix}.
\] (23)

The neat thing is that $\beta \in V^*$ can be written as a row vector,
\[
(v_1, v_2, \ldots, v_n).
\] (24)

\[\text{It's actually quite sloppy to say that } V^{**} = V. \text{ Rather, } V^{**} \text{ is isomorphic to } V \text{ which means that they are similar but fundamentally separate things and one usually denotes an isomorphism by using } \simeq \text{ instead of } =. \text{ So } V^{**} \simeq V. \text{ However, it suffices to think of it as 'the same as'.}\]
2.1.3

So in this way we can apply the row vector to the column vector and get a scalar! In order to do this we need to pick a basis and an inner product. We will see that the use of this in physics consists of choosing the inner product to be the ordinary dot product which of course produces a scalar.

Now we’re equipped with the knowledge of what a dual vector in a dual space is and can return to the concept in question, namely tensors.

2.1.3 Tensors

In a way, the tensor is quite similar to the dual vector. I said similar again, what do I mean by that? The dual vector is a linear map from a vector space to $\mathbb{R}$. A tensor is a multilinear map from a set of vectors and dual vectors to $\mathbb{R}$. So, like the dual vector, a tensor takes in vectors (and dual vectors) and returns a real number. The fact that it is multilinear means that it is linear for several variables.

Definition 2.1. Let $V$ be a finite-dimensional vector space and let $V^*$ be the corresponding dual space. A tensor of rank $(r, s)$ is a multilinear map

$$T : V^* \times \ldots \times V^* \times V \times \ldots \times V \rightarrow \mathbb{R}. \quad (25)$$

So the tensor takes in $r$ dual vectors and $s$ vectors and produces a number. An even more general statement would be to say that the multilinear map goes to the field\textsuperscript{16} $K$ of which $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{Q}$ are examples. Earlier we said that it is multilinear, but didn’t state exactly what meant by that. When we fix all but one of the dual vectors or vectors it is linear for the remaining variable[10]. Tensors are quite general, they can be scalars, vectors, dual vectors and objects of higher rank. In fact, a scalar is a rank $(0, 0)$ tensor and the components of a (contravariant)vector is a rank $(1, 0)$ tensor, while a dual vector or a covariant vector is a rank $(0, 1)$ tensor.

The set of general rank $(k, l)$ tensors form a vector space, and in this space we are able to multiply the tensors by real numbers and add them together. As we know, we can define bases in vector spaces, and in order to do so in the case of tensors we need to consult with a somewhat difficult concept, namely the tensor product.

\textsuperscript{16}Roughly speaking, a field is a set of elements where you can add, subtract, multiply and divide which are commutative operations. Because subtraction is really glorified adding, and division is really glorified multiplication, we only need addition and multiplication. Then, a field is a set of elements with two operations: addition and multiplication. It is a commulative group under both operations if we omit zero. A field is also distributive.
Definition 2.2. Let $T$ be a rank $(r,s)$ tensor and $E$ be a rank $(k,l)$ tensor. The tensor product between $T$ and $M$ denoted $T \otimes E$ constructs a new tensor of rank $(r+k, s+l)$ by the rule

$$
\begin{align*}
(T \otimes E)(v^1, \ldots, v^r, \ldots v^{r+k}; w_1, \ldots, w_s, \ldots, w_{s+l}) & = T(v^1, \ldots, v^r; w_1, \ldots w_s) \times E(v^{r+1}, \ldots, v^{r+k}; w_{s+1}, \ldots w_{s+l}).
\end{align*}
$$

An important aspect of tensor products is that they do not commute, i.e. $T \otimes E \neq E \otimes T$. The procedure in (8) is quite straightforward: $T$ acts on the first set of dual vectors and vectors, and then $E$ acts on the dual vectors and vectors that are left, and the answers are multiplied. It is also important to note that the entries of $T$ and $E$ are separate dual vectors and vectors. That is, they are not components[11].

2.1.4 A Brief Discussion on Tensor Products

A ton of questions might arise from the definition of the tensor product given above, so let's try and demystify it a little bit.

Example 2.1. Let $f : V \times V \to \mathbb{R}$ be a bilinear map with a basis $\beta = \{v_1, \ldots, v_n\}$ for $V$. Then, we can write the vectors $x$ and $y$ as $x = \sum_i x_i v_i$ and $y = \sum_j y_j v_j$, respectively. If we then take in the ordered pair $(x, y)$ in $f$ we get $f(x, y) = f\left(\sum_i x_i v_i, \sum_j y_j v_j\right)$.

Because it's bilinear we can pull out the sums to get

$$
\begin{align*}
f(x, y) & = \sum_{i,j} x_i y_j f(v_j, v_j) \\
& = \left(\sum_{i,j} f(v_i, v_j) v^i \otimes v^j\right)(x, y)
\end{align*}
$$

This last step needs some elaboration. If $\alpha, \beta \in V^*$, then $\alpha \otimes \beta$ defines a new mapping. So when $\alpha \otimes \beta$ acts on a pair $(x, y)$ we get $(\alpha \otimes \beta)(x, y) = \alpha(x)\beta(y)$.

But what does this have to do with the last step in (9)? $x^i y^j$ is the same as the dual basis acting on $x$ times the dual basis acting on $y$, i.e., $x^i y^j = v^i(x) v^j(y)$ where $v^i, v^j \in V^*$. Thus, given the definition with $\alpha$ and $\beta$ we get

$$
\begin{align*}
x^i y^j & = v^i(x) v^j(y) \\
& = (v^i \otimes v^j)(x, y).
\end{align*}
$$

29
So we can build a bilinear map $b$, and even a multilinear map $T$ using a linear combination of the tensor products!\[12\]

There are often two different ways of explaining tensor products. The concrete explanation involves what we have just done by building multilinear map which is really a use of the tensor product of dual vectors and vectors. The other explanation is harder to grasp and more abstract because the tensor products are taken between vector spaces. If this still is confusing i will close by saying this: Given three vector spaces $V, W$ and $Z$, the tensor product is a multilinear mapping from $f : V \times W \rightarrow Z$. So it reduces multilinear algebra to linear algebra. Instead of focusing on $n$ vector spaces, lets focus on 2 which is simpler to handle. The bilinear mapping takes elements from $V \times W$ to the mysterious space we denote by $\otimes$, which can be illustrated by a chart

$$
\begin{array}{ccc}
V \times W & \xrightarrow{\text{bilinear } f} & Z \\
\downarrow \text{bilinear } \otimes & & \\
V \otimes W & \xrightarrow{\text{A unique linear map } \tilde{f}} & \\
\end{array}
$$

Figure 2: Chart of mappings that illustrates what a tensor product does when applied to two vector spaces $V$ and $W$\[13\].

As I said, in this way the tensor product is a thing that reduces multilinear algebra to linear algebra. If this is confusing, it suffices to think about tensor products as they are presented in definition 2.2.

So now that we have defined the tensor product we can do what we set out to do, namely define a basis for the space consisting of all $(r,s)$ tensors. The basis will by definition be a tensor, and it will consist of basis dual vectors and basis vectors given by

$$
e_{\mu_1} \otimes \ldots \otimes e_{\mu_r} \otimes e^{\nu_1} \otimes \ldots \otimes e^{\nu_s},$$

(29)

30
where we have used the same notation for the basis vectors as in (3), but with * to highlight the dual basis. There are \( n^{r+s} \) tensors\(^\text{17}\) as given in (11), and so all tensors \( T \) of type \((r, s)\) can be written on the form

\[
T = \sum_{\mu_1, \ldots, \nu_1}^n T^{\mu_1 \cdots \mu_r \nu_1 \cdots \nu_s} e_{\mu_1} \otimes \cdots \otimes e_{\mu_r} \otimes e^{\nu_1} \otimes \cdots \otimes e^{\nu_s}.
\]

(30)

Using the Einstein summing convention we can omit the symbol for the sum. The expansion coefficients \( T^{\mu_1 \cdots \mu_r \nu_1 \cdots \nu_s} \) are the components of the tensor, and these components are what we will deal with. They can be defined as the action of the tensor \( T \) on the basis dual vectors and vectors by

\[
T^{\mu_1 \cdots \mu_r \nu_1 \cdots \nu_s} = T(e^{\mu_1}, \ldots, e^{\mu_r}, e_{\nu_1}, \ldots, e_{\nu_s}).
\]

(31)

The above examples of tensors might be easy to grasp. Upon asking the question of what a rank \((1, 1)\) is one realises that these things are hard to comprehend. In fact, one can look at tensors of higher rank in many different, but equivalent ways\([10]\). Following the definition, it is a map \( V^* \times V \to \mathbb{R} \). This tensor converts a contravariant vector field to the same contravariant vector field, i.e., it does nothing to it. Another question arises: how does that relate to the map \( V^* \times V \to \mathbb{R} \)? The dual space of the dual space of \( V \) is \( V \) itself, that is \( V^{**} = V \). If we fix the vector \( v \in V \), the tensor \( T(\tau, v) \) is a linear function of \( \tau \in V^* \) that we will denote \( A_v \in V^{**} = V \). Thus, we have

\[
\langle A_v, \tau \rangle = A(\tau, v).
\]

(32)

where \( \langle , , \rangle : V^* \times V \to \mathbb{R} \) is the bilinear function called the natural pairing of \( V^* \) and \( V \) onto \( \mathbb{R} \).

### 2.1.5 Tensor Fields on Manifolds

When physicists talk about tensors, they often mean tensor fields on manifolds. In the same way that a tensor is a generalisation of a vector, a tensor field is a generalisation of a vector field in which a vector is assigned at each point. In the case of tensor fields, it is then the case that a tensor is assigned at each point, and this point lives on a manifold.

Given some manifold \( M \), we define a tangent vector at some point \( p \in M \) to be

---

\(^{17}\)In four-dimensional spacetime \( n = 4 \) and so there are \( 4^{r+s} \) tensors.
an infinitesimal displacement on \( M \). The set of all these tangent vectors on \( M \) is denoted \( T_p \) and is called the **tangent space** at \( p \). Likewise, \( T^*_p \) is the **cotangent space** and elements of \( T^*_p \) are called **cotangent vectors**. The vector spaces \( V \) and \( V^* \) are replaced by \( T_p \) and \( T^*_p \). The same notions about contravariant and covariant vectors apply in the tangent spaces. The basis we construct on \( T_p \) takes the form \( \{ \partial/\partial x^1, \ldots, \partial/\partial x^n \} \in T_p \), and the basis for the cotangent space is denoted \( \{ dx^1, \ldots, dx^n \} \in T^*_p \). The reason we distinguish between symbols of the bases on the tangent and cotangent spaces is merely a because \( dx^\mu \) is the symbol for a linear map on \( \partial/\partial x^\nu \) given by

\[
dx^\mu (\partial/\partial x^\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu. \tag{33}\]

Say we have some vector \( v \) with a basis \( \partial/\partial x^\mu \). If the coordinate system is changed, the new components \( v'^\mu \) are related to the old components \( v^\mu \) by

\[
v'^\mu = \frac{\partial x'^\mu}{\partial x^\mu} v^\mu. \tag{34}\]

Similarly, given a basis \( \{dx^\mu\} \) of the components \( w_\mu \) of a dual vector \( w \) we define a transformation of the coordinates as

\[
w'_\mu = \frac{\partial x^\mu}{\partial x'^\mu} w_\mu. \tag{35}\]

In fact, this transformation cannot be anything else given (?) and (?). We can now generalise this in terms of transformations of type \((r, s)\) tensors:

\[
T_{\mu'1\ldots\mu'_{r}}^{\nu'1\ldots\nu'_{s}} = \frac{\partial x^{\mu'_{1}}}{\partial x^\mu} \cdots \frac{\partial x^{\mu'_{r}}}{\partial x^\mu} \frac{\partial x^{\nu'_{1}}}{\partial x^\nu} \cdots \frac{\partial x^{\nu'_{s}}}{\partial x^\nu} T_{\mu1\ldots\mu_{r}}^{\nu1\ldots\nu_{s}}. \tag{36}\]

### 2.2 Usual Definition as Seen in Physics Texts

I have to be a bit careful when i state what’s in the title of this section. Some of the things i stated in the mathematical definition was, in fact, taken from physics texts! However, there are physics texts where the focus on the mathematical aspects are neglected.

---

\(^{18}\)The set of all the tangent spaces on a manifold of dimension \( n \) we call the **tangent bundle** \( T(M) \) of dimension \( 2n \). In addition, the cotangent bundle is denoted by \( T^*(M) \).
2.2.1 Index Gymnastics and Important Tensors

3 Various Mathematical Tools

3.1 Fourier Transforms

3.2 The Dirac Delta Distribution

3.3 Green’s Functions

3.4 Complex Analysis

4 Classical field theory

4.1 The need for quantum fields

The Schrödinger equation was and is a great triumph in physics. However, at the time of its arrival special relativity had already been developed by Einstein in 1905, and since the Schrödinger equation was not relativistic the need for one was evident. That is, there is a symmetry between space and time in special relativity, but the Schrödinger equation

\[ i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + V\psi \]

has a first-order time derivative and a second order spatial derivative. So the apparent asymmetry in this equation doesn’t satisfy special relativity.

His attempt to add special relativity to this equation which we now call the Klein-Gordon equation was not successful because he failed to add the spin of the electron and so the fine structure constant turned out to be the wrong value. Paul Dirac saved the day with his equation including massive spin-\(1/2\) particles which is consistent with both quantum mechanics and special relativity. In the realm of special relativity, a multitude of particles live and interact through e.g. particle-antiparticle pairs, so it becomes evident that we need a theory in which these particles can live. That is where the fields come in. In later chapters we will talk about the Dirac field and the Klein-Gordon field. There’s a number of reasons for why a field theory of quantum mechanics is needed, and we could spend a lot of time going through them all, but
I think it’s sufficient to know some of the reasons. We will intersect with them later on after all.

4.2 Lagrangian and Hamiltonian Field Theory

In classical mechanics, the Lagrangian is the quantity that tells us something about the dynamics of a system with \( N \) degrees of freedom described a set of generalized coordinates \( q_i(t) \) for \( i, \ldots, N \), often denoted \( q \). The Lagrangian is a function of \( q \) and its time derivative \( \dot{q} \), that is \( L = L(q, \dot{q}) \). The first time one meets this function is takes the form where the potential term is subtracted from the kinetic term, \( L = \sum_i \frac{1}{2} m_i \dot{q}_i^2 - V(q) \). However, this is the simplest form of the Lagrangian. This brings us to the action which is the time integral of the Lagrangian. The famous principle of least action tells us that if some object’s initial position and final position is fixed by the coordinates at some initial time \( t_{\text{init}} \) and some final time \( t_f \) such that \( q(t_{\text{init}}) = q_{\text{init}} \) and \( q(t_f) = q_f \), the trajectory of that object is path for which the action \( S \) is an extremum. More precisely,

\[
\delta S = 0 \quad \text{(37)}
\]

\[
\delta \int_{t_{\text{init}}}^{t_f} dt L(q, \dot{q}) = 0. \quad \text{(38)}
\]

The variation of the Lagrangian is

\[
\delta L = \sum_i \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i, \quad \text{(39)}
\]

and this means that the variation in the action becomes

\[
\delta S = \int_{t_{\text{init}}}^{t_f} dt \sum_i \left[ \frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right] = 0 \quad \text{(40)}
\]

\[
= \int_{t_{\text{init}}}^{t_f} dt \sum_i \left[ \frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right] \delta q_i = 0 \quad \text{(41)}
\]

So now we obtain the essential equation in rudimentary classical mechanics, namely the Euler-Lagrange equation.

\[
\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \quad \text{(42)}
\]

The Hamiltonian formalism requires a slightly different coordinate in addition to \( q \) named the conjugate momenta. That is,

\[
p = \frac{\partial L}{\partial \dot{q}_i} \quad \text{(43)}
\]
which means that the Hamiltonian is a function of both \( p \) and \( q \). I.e.,

\[
H(p, q) = \sum_i p_i \dot{q}_i - L.
\]  

(44)

This brings us to a description using fields. By this we mean that we go from using functions of time to using functions of space-time, \( q_i(t) \to \phi_i(x) \), and the functions \( \phi_i(x) \) are simply denoted \( \phi \). The space-time point \( x \) is as we introduced it in the section about the Poincaré group where we talked about representations on fields. It is the local fields we’re interested in and in this case the Lagrangian takes the form of an integral over the Lagrangian density \( \mathcal{L} \) which is often referred to as the Lagrangian itself. Thus, the Lagrangian takes the form

\[
L = \int d^3x \mathcal{L}(\phi, \partial_\mu \phi).
\]  

(45)

This means that the action takes the form

\[
S = \int dt L = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi).
\]  

(46)

But what about the boundary conditions in this case? How do we transfer the boundary conditions of between the two values \( t_{\text{init}} \) and \( t_f \) to a field? The boundary conditions are not much unlike those we use when we define the domain of the probability density of a state in quantum mechanics. That is, the boundary conditions of a field requires the field to decrease sufficiently fast at infinity. In other words, it dies out at infinity in a way such that we can neglect the boundary conditions. We don’t have to think about them! The action is defined the same way as before, thus

\[
\delta S = \int d^4x \sum_i \left[ \frac{\partial \mathcal{L}}{\partial \phi_i} \delta \phi_i + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \delta (\partial_\mu \phi_i) \right]
\]  

(47)

\[
= \int d^4x \sum_i \left[ \frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right] \delta \phi_i = 0.
\]  

(48)

Then, the Euler-Lagrange equation takes the form

\[
\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)}.
\]  

(49)

If we now introduce a small change in the Lagrangian density in the form of a partial derivative \( \partial_\mu K_\mu(\phi) \) we can analyse what happens to the action. It is no big surprise that we want the Lagrangian \( \mathcal{L}' = \mathcal{L} + \partial_\mu K_\mu(\phi) \) to give rise to the stationary action

35
$S'$. The integral over some finite volume $V$ bounded by the surface $\Sigma$ can be treated using the famous theorem by Stokes, i.e.

$$\int_V d^4x \partial_\mu K^\mu = \int_\Sigma dA n^\mu K^\mu$$

where $n^\mu$ is the normal vector of and $dA$ is an infinitesimal surface element of $\Sigma$. Because this is a term on the boundary of the volume it will vanish when the field goes sufficiently fast to zero at infinity.[14] So this term is either zero or stationary due to the requirement that the fields are kept constant on $\Sigma$. In general, then, total derivatives do not change the Lagrangian and thus leaves the action stationary.

Earlier we defined what we meant by conjugate momenta by not using fields, and the definition is similar when we do. That is,

$$\Pi_i(x) = \frac{\partial L}{\partial (\partial_0 \phi_i(x))}.$$  

(51)

By extension, the Hamiltonian density is

$$\mathcal{H} = \sum_i \Pi_i(x) \partial_0 \phi_i(x) - L$$

(52)

which yields the total Hamiltonian

$$H = \int d^3x \mathcal{H}.$$  

(53)

4.3 Noether’s Theorem

It is of great interest in classical field theory to expose our fields to some change and see if they leave the action invariant. If so, we say that the transformation on the fields is a symmetry transformation. These symmetries give rise to important and fundamental quantities, namely those that are conserved. Let $\phi_i$ be fields with action $S$. We perform an infinitesimal transformation on the fields and the coordinates by a set of parameters $\epsilon^a$ where $a = 1, \ldots, N$ such that

$$x^\mu \rightarrow x'^\mu = x^\mu + \epsilon^a A^\mu_a(x)$$

$$\phi_i(x) \rightarrow \phi'_i(x') = \phi_i(x) \rightarrow \epsilon^a F_{i,a}(\phi, \partial \phi),$$

(54) (55)
5 The Klein-Gordon Field

6 The Dirac Field

When Schrödinger devised his equation to explain individual particles he kept track of each quantum state. It doesn’t take very long for one to wonder how this plays out when you have many states, and you’re right to think that it’s hard to keep track of them all! That is why the Schrödinger equation is terrible when we consider many particle states, the bookkeeping gets too complicated. That being said, it still was a triumph of physics, and is known as first quantization. Paul Dirac wanted to avoid keeping track of each changing particle state.

7 Quantization of Free Fields

7.1 Fock Space

8 Perturbation Theory and Feynman Diagrams

9 Elementary QED Processes & Functional Methods

10 Cross sections and the S-matrix; Feynman Rules for QED

11 Non-Abelian Gauge Invariance

12 Quantization of Non-Abelian Gauge Theories

13 Higgs Mechanism & Electroweak Theory
References


