# Elementary Particles in Curved Spacetime

L.C.W. Jeronimus
CTN, RuG

*28 August, 2012*

**Abstract**

Elementary particles can be identified with the unitary irreducible representations (UIR's) of the isometry group of a given spacetime. These UIR's are labeled by the eigenvalues of the Casimir operators of the isometry group and hence they represent invariant physical properties of the elementary particles. These properties therefore depend entirely on the spacetime background of the particle. To compare these labels for different spacetime backgrounds, one can use the method of contraction.

# Contents

**1 Introduction** **5**

**2 Isometry groups and GR** **6**
2.1 Symmetry transformations . . . . . 6
2.2 Spacetime, Manifolds and Coordinates . . . . . 6
2.2.1 S.R. and G.R. . . . . 9
2.3 Induced transformations . . . . . 10
2.3.1 Induced transformations on scalar fields . . . . . 10
2.3.2 Induced transformations on co- and contravariant vector fields . 11
2.3.3 Induced transformations on tensorfields . . . . . 11
2.4 Inertial observers . . . . . 12
2.5 General Relativity . . . . . 15
2.5.1 Covariant derivatie, Lie derivative and the Killing equations . . 17
2.6 Symmetry operations form a group . . . . . 21

**3 Discrete groups** **23**
3.1 Different Substructures of a Group . . . . . 24
3.2 Derived Groups . . . . . 26
3.3 Representations . . . . . 28
3.4 Representations acting on Vector spaces . . . . . 31
3.5 Equivalent Representations and Character . . . . . 33
3.6 Reducibility; Irreps and unitarity . . . . . 34
3.7 Great Orthogonality Theorem; Decomposition . . . . . 37
3.7.1 Great Orthogonality Theorem (GOT) . . . . . 37
3.7.2 Decomposition . . . . . 39
3.8 Tensor product representation; Clebsch Gordan decomposition . . . . . 40

**4 Lie groups** **42**
4.1 General properties Lie groups . . . . . 42
4.2 Representations of Lie Groups . . . . . 43
4.2.1 Non-Compact Lie groups and unitary representations . . . . . 44
4.3 Lie Algebra; $L$ . . . . . 44
4.4 Lie product; Structure Constants . . . . . 46
4.5 Additional structures of the Lie Algebra . . . . . 47
4.6 Casimir Operators, irrep labels and CSCO . . . . . 48
4.7 Representations of Lie Algebras . . . . . 49
4.7.1 Adjoint Representation of Lie Algebra . . . . . 49
4.8 What Discrete Group properties carry over to Lie groups? . . . . . 50

4.9 $SO(3)$ and $SU(2)$; Important compact groups . . . . . 50
4.10 $SU(2)$ is the universal covering group of $SO(3)$ . . . . . 53
4.11 Irreps of $SO(3)$ . . . . . 54

**5 The Lorentz and Poincaré Group** **58**
5.1 Finite dimensional non-unitary irreps of Lorentz group . . . . . 61
5.2 Infinite dimensional unitary irreps of the Lorentz group . . . . . 63
5.3 Poincaré Algebra and the Casimir operators . . . . . 65
5.4 Unitary irreps of the Poincaré group . . . . . 66
5.4.1 Induced representation method in general . . . . . 67
5.4.2 Induced representation method for Poincaré . . . . . 68

**6 Conclusion** **70**

**A (In)finite dimensional Vector Spaces and the concept of length** **71**

# Acknowledgements

As this thesis wraps up a long time of studying, I would like to make use of the opportunity to thank some people that were important to me during these years. I started my master in theoretical physics on the risk of and admittedly a bit in fear of not-succeeding. But I'm glad that I decided to give in to my intrinsic motivation: curiosity for the ideas and concepts that I think govern the fundamental laws of this universe. I'v learned a lot, satisfied my curiosity and discovered my boundaries.

First of all I would like to thank Diederik Roest for his mentoring. Thank you for the interesting discussions, the uplifting speeches and personal talks in times of doubt. They all where very helpful. Your convincing positivity lifted my moral many times. Then of course I want to thank Giuseppe Dibitetto and Roel Andringa, for sharing your knowledge and time with me. Talking about the subject really appeared *the* way for me to get a grip and you guys provided in that. Thank you for letting me ask the same questions over and over again..without getting annoyed!

On a more personal level I would like to thank my parents very much for creating the possibility for me to fulfill my ambitions and discover my boundaries for so many years. I think this is the most valuable thing you could possibly give to me. Thank you for your faith and persistence and mom's room, for finishing the job.

Then I would like to thank Rike for her love and patience. Your reflective understanding and positivity fuels me and really is a blessing to me. You encouraged me to finish and volunteered to listen to incomprehensible stuff, for the sake of listening. Thank you.

And although chances are small that they will ever read this, I still would also like to thank my friends and Anna, my sister, for being wonderful friends. Thank you for the quality time diversions, for playing afternoon filling games, for the laughter and the reflexive talks. For showing me how to graduate: I've had many great examples, but it took some time for me to find my way.

Thank you all very much!

# 1 Introduction

What is an elementary particle? Or what characterizes an elementary particle? By trying to answer these questions we would like to find the properties of a particle that never change. Thus we want the properties to be invariant when regarded by two different observers. But *any* two observers? Can they accelerate for example with respect to eachother? Would the particles still have the same properties?

Classically we can think of physical variables like the position or the energy of the particle and we know that they obviously depend on the movement of observers. So these are obviously not suited to describe the elementary particle unambiguously. Einstein showed with his special relativity theory that for any two observers to find the same laws of physics in a flat Minkowski space, these observers had to be in inertial frames of reference, related by Poincaré transformations. Relativistically we know that the rest mass for example is equal for all inertial observers, so this could be a candidate.

Not only the properties that characterize elementary particles but also the laws of physics themselves should not depend on the observer. This is called the *principle of equivalence*. Special relativity is the theory where this principle is applied to Minkowsky space. But this principle is valid for any kind of spacetime in which our physics takes place or the invariant properties of the elementary particle are measured, for that matter. So what happens if we go to curved space? In general relativity theory we see that because of the observed equality of gravitational and inertial mass, the effects of gravity can equivalently be described by the curvature of spacetime. Free falling observers, which in terms of general relativity follow geodesics in the curved space, should observe the same physics as an observer would in the absence of gravity. Intuitively it might therefore be clear that what we define as inertial observers first of all depends on the spacetime in which these observers move. It might therefore be very well thinkable that the invariant properties of the elementary particle change if we move into curved spacetimes.

Our space on a universal scale *is* actually curved. This we know because of the measurements done on distant supernovea that turned out to accelerate away from us, indicating a positive cosmological constant $\Lambda$ [1], [2]. However on a local scale, here on earth this curvature due to the cosmological constant is really much to small to detect in any experiment done on the particles. Furthermore the curvature of spacetime due to the mass of the earth, i.e. the earths gravity, would "overrule" the curvature caused by $\Lambda$. The questions I'd therefore like investigate are first of all how one can define the invariant properties that characterize the elementary particles and furthermore how they depend on the spacetime and its curvature the particles are in. To answer these questions we'll have to define what we mean by an inertial observer in any given spacetime, curved or not. General relativity is the theory to answer this and therefore we'll review some of the ideas from this theory, important to us, in the next section.

# 2 Isometry groups and GR

We would like to compare two observers in a given spacetime and see what we can say about how the measurements they make are related to each other. For this we need some geometry and in the next few paragraphs I will review some of the basics of geometry and we will work gradually towards defining the important symmetry operations in a given spacetime.

## 2.1 Symmetry transformations

Transformations are operations that change something in general. In the context of physics they can be grouped into two categories, transformations that affect the "external spacetime" in which the physics takes place or describes the physics itself, as with general relativity, and those transformations that affect the "internal space" of particle fields. This internal space is a mathematical construct to describe the internal quantum numbers that define different types of particles. Actually "external spacetime" is as much a mathematical construct as is internal, it's just that internal is only accurately described since the emergence of quantum mechanics this century and external spacetime as a concept is as old as man could "translate" from one place to another and thought about it. For now we are focussed on transformations acting on spacetime only.

In physics we are mostly interested in those transformations that act on spacetime, but leave the physics untouched or invariant. Those special operations are called *symmetry transformations* and tell us a lot about the physics that can take place in those spacetimes. In classical field theories invariant physics can be guaranteed by an invariant action $\delta\mathcal{S} = 0$, where $\mathcal{S} = \int d\mathcal{L}$ and $\mathcal{L}$ is the Lagrangian density. In quantum mechanics an equivalent demand for an operation to be a symmetry transformation would be that the unitary representation $U(T)$ of the symmetry transformation $T$ commutes with the Hamiltonian $H$ that describes the time evolution of the system, i.e. $[U(T), H] = 0$.

Now it's almost time to look at what makes a transformation a symmetry transformation of spacetime. But first some technical definitions to understand the transformations a little better:

## 2.2 Spacetime, Manifolds and Coordinates

Spacetime in geometric term is a *manifold* (see [5] and [6]) and manifolds have particular good properties, such as differentiability. A differentiable manifold of dimension $n$ is any topological space $M$ that is locally isomorphic to the Euclidean space $\mathbf{R}^n$. This means that there exists a mapping called a *coordinate vector* on the manifold $M$, which is an isomorphic mapping

$$\mathbf{x}\colon U \to \mathbf{R}^n, \tag{2.1}$$

from a subset $U \subset M$ to an open subset of the Euclidean vector space $\mathbf{R}^n$. Any finite, say $n$-dimensional vector space has some set of $n$ vectors $\{e_{(\mu)}\}$, $\mu = 1, \ldots, n$ called a *basis*, which span the vector space. This means that any vector $\mathbf{x}$ in this space $\mathbf{R}^n$ can

be expressed as some linear combination of all vectors in this set, i.e. $\mathbf{x} = x^\mu e_{(\mu)}$, where the coefficients $x^\mu$ of this linear combination are called the *components* of the vector $\mathbf{x}$ in the basis $\{e_{(\mu)}\}$. The actual *coordinates* $x^\mu$ are therefore the $n$ components of an $n$-dimensional position vector $\mathbf{x} = x^\mu e_{(\mu)}$, which indicates the position on the manifold in a basis $\{e_{(\mu)}\}$. The dimension $n$ of the manifold indicates the number of variables one has to specify to fix the position of some point on the manifold. Since the mapping $\mathbf{x}$ is isomorphic this number should coincide with the dimension of the Euclidean target space $\mathbf{R}^n$ of the coordinate. As a result of imposing a coordinate onto the manifold, differentiation and integration, which only makes sense in Euclidean space, can be performed on the manifold.

This vector $\mathbf{x}$ is not to be interpreted as some "arrow" pointing from the origin to that point $p$ on the manifold, represented by the coordinates $x^\mu(p)$. Any vector on a manifold $M$ is to be interpreted as a geometrical object living in the so called *tangent space* $T_p$ on $p \in M$. This tangent space $T_p$ is simply the set of all possible vectors at $p$ and has the same dimension $n$ as $M$, i.e. it is a Euclidean space $\mathbf{R}^n$ at every point $p$ on the manifold. The position vector $\mathbf{x}$ for the point $p$ on the manifold therefore lives in $T_p$ at $p$; it can be regarded as $n$ numbers attached to $p$, describing the position of $p$ relative to its surroundings $U$.

A common basis is for any Euclidean coordinate space $\mathbf{R}^n$ is the *coordinate basis* $e_{(\mu)} = \frac{\partial}{\partial x^\mu} \equiv \partial_\mu$. This is introduced by identifying the tangent space $T_p$ to a point $p$ on the manifold $M$, with "the space of directional derivative operators along curves through $p$". A vector therefore becomes an operator as we will see. A curve $\gamma(\lambda)$ through $p$ is simply any mapping $\gamma\colon \mathbf{R} \to M$, such that there is a $\lambda \in \mathbf{R}$ for which $\gamma(\lambda) = p \in M$. For the directional derivative we need also the space $\mathcal{F} \ni f$ of all smooth, i.e. $C^\infty$, functions on $M$, $f\colon M \to \mathbf{R}$. Since then for some smooth function $f$ any curve $\gamma(\lambda)$ through $p$ defines an operator $\frac{d}{d\lambda}$, the directional derivative, on $f$, i.e. $\frac{d}{d\lambda}\colon f \to \frac{d(f\circ\gamma)}{d\lambda} \equiv \frac{df}{d\lambda}$. Because directional derivatives are linear $a, b \in \mathbf{R}, f, g \in \mathcal{F}$

$$\frac{\partial}{\partial\lambda}(af + bg) = a\frac{\partial f}{\partial\lambda} + b\frac{\partial g}{\partial\lambda} \tag{2.2}$$

and obey the Leibniz rule

$$\left(a\frac{\partial}{\partial\lambda} + b\frac{\partial}{\partial\eta}\right)(fg) = \left(a\frac{\partial f}{\partial\lambda} + b\frac{\partial f}{\partial\eta}\right)g + \left(a\frac{\partial g}{\partial\lambda} + b\frac{\partial g}{\partial\eta}\right)f \tag{2.3}$$

it follows that the set of directional derivatives itself forms a vector space of the same dimension $n$ as $M$ and below we derive that any directional derivative can be decomposed into $\{\partial_\mu\}$, the set of partial derivatives associated with a chosen coordinate set $x^\mu$.

$$\begin{aligned}\frac{df}{d\lambda} &\equiv \frac{d(f\circ\gamma)}{d\lambda}\\ &= \frac{d[(f\circ(x^\mu)^{-1})\circ(x^\mu\circ\gamma)]}{d\lambda}\\ &= \frac{d(x\circ\gamma)^\mu}{d\lambda}\frac{\partial(f\circ(x^\mu)^{-1})}{\partial x^\mu}\\ &\equiv \frac{dx^\mu}{d\lambda}\frac{\partial f}{\partial x^\mu}\\ &= \frac{dx^\mu}{d\lambda}\partial_\mu f.\end{aligned} \tag{2.4}$$

Since this holds for any smooth function $f$, we have that the directional derivative operator $\frac{d}{d\lambda}$ can de decomposed as

$$\frac{d}{d\lambda} = \frac{dx^\mu}{d\lambda}\partial_\mu. \tag{2.5}$$

The set $\{\partial_\mu\}$ consequently forms the coordinate basis for $T_p$ in which the vectors $\frac{d}{d\lambda}$ live. Given any manifold $M$ with coordinates $x^\mu$, the set $\{\partial_\mu\}$ forms a basis, the coordinate basis, for all vectors on this manifold. There are infinitely many bases for any vector space and to go from one $\{\partial_\mu\}$ to a primed basis $\{\partial_{\mu'}\}$, one performs a *basis transformation*. One advantage of this coordinate basis $\{\partial_\mu\}$ is that the basis transformation is naturally given by the chain rule

$$\partial_{\nu'} = \frac{\partial x^\mu}{\partial x^{\nu'}}\partial_\mu. \tag{2.6}$$

In defining equation (2.6) I used the term basis transformation, but more common is *coordinate transformation*. The matrix $\frac{\partial x^\mu}{\partial x^{\nu'}}$ is for example called a *general coordinate transformation* (gct). Basis transformations and coordinate transformations are two sides of the same coin really. If you do a coordinate transformation you choose another mapping on your manifold, say $x^\mu \to x^{\mu'}$. This automatically induces a basis transformation $\partial_\mu \to \partial_{\mu'}$. The reverse order would be to change the basis $\partial_\mu \to \partial_{\mu'}$ and accordingly adapt the coordinates $x^\mu \to x^{\mu'}$. Most of the time we let the transformations act on the coordinate mappings and we therefore talk of coordinate transformations.

Since any general vector $\mathbf{x}$ does not depend on the choice of basis or equivalently on the choice of coordinates, i.e. $\mathbf{x} = x^\mu\partial_\mu = x^{\mu'}\partial_{\mu'}$, we derive using equation (2.6) that the components $x^\mu$ must transform as

$$x^{\mu'} = \frac{\partial x^{\nu'}}{\partial x^\mu}x^\nu, \tag{2.7}$$

where $\frac{\partial x^{\nu'}}{\partial x^\mu}$ is of course the inverse of $\frac{\partial x^\mu}{\partial x^{\nu'}}$. Any vector $V$ with components $V^\mu$ that transform like the components $x^\mu$ of the position vector in equation(2.7) is correspondingly called a *contravariant* vector, because it transforms in the opposite way as the basis does and it lives in $T_p$. A common example of a contravariant vector is the tangent vector on a parametrized curve $x^\mu(\lambda)$, i.e. $\frac{dx^\mu}{d\lambda}$.

Any vector $\omega$ with components $\omega_\mu$ that transform analogously to equation (2.6), transforms thus in the same manner as the basis does and is likewise called a *covariant* vector. Covariant vectors $\omega$ in $p \in M$ live in a space dual to the tangent space $T_p$, called the *dual space* $T_p^*$. Given a (complex) vector space $T_p$ the dual space $T_p^*$ is formally defined to be the space of linear functionals on $T_p$, i.e. for any $\omega \in T_p^*$ we have $\omega \colon T_p \to \mathbf{C}$ also written as $\omega(V) \in \mathbf{C}$. The general requirement for any basis $\theta^{(\mu)}$ of a dual space $T_p^*$, dual to some vector space $T_p$ with basis $e_{(\nu)}$, is that it will have to obey $\theta^{(\mu)}(e_{(\nu)}) = \delta^\mu_\nu$. For the coordinate basis this follows evidently, since $\partial_\nu dx^\mu = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu$. A common example of a covariant vector is the gradient of a scalar function $\phi(\mathbf{x})$, i.e. $\frac{\partial \phi}{\partial x^\mu} = \partial_\mu \phi$. Note that both $\partial_\mu \phi$ and $\frac{dx^\mu}{d\lambda}$ are only the components of the two vector types $\omega$ and $V$. The entire vectors look like $\omega = \partial_\mu \phi dx^\mu$ and $V = \frac{dx^\mu}{d\lambda}\partial_\mu$. So since any tangent space $T_p$ is vector space at the point $p$ on the manifold and every vector space implies a dual space of functionals $T_p^*$ on this same point $p$ on the manifold, any vector $V$ can be expanded in the ordinary $\partial_\mu$ basis for $T_p$ or the $dx^\mu$ basis for $T_p^*$. The vector itself is therefore neither co- nor contravariant, but its components are.

Given a metric $g_{\mu\nu}$ and its inverse $g^{\mu\nu}$ on a manifold, defining the notion of distance on a manifold (see Appendix A), this relates the co- and contravariant components of a vector directly by $V_\mu = g_{\mu\nu}V^\mu$ and $V^\mu = g^{\mu\nu}V_\mu$.

Why are coordinate transformations so interesting? Because they could be *symmetry* transformations. Like mentioned, any transformation leaving the classical action $\mathcal{S}$ invariant is a symmetry transformation of this theory. To distinguish all possible symmetry transformations of a particular theory, gives us a lot of information about the physics. The symmetry transformations therefore form some subset of all possible coordinate transformations.

### 2.2.1 S.R. and G.R.

Take special and general relativity for example. In general relativity one demands that gct's given by equations(2.6) and (2.7), are a symmetry transformation. But for special relativity this is to general to be a symmetry. In special relativity one demands that the Minkowsky metric is invariant under a transformation, which results in *linear* transformations of the form of equations(2.8) and (2.10) below. In section 2.5 well go into detail on these sets of operations called the Lorentz and Poincaré group.

In the special relativistic case, we will see that the two bases $\{\partial_\mu\}$ and $\{\partial_{\mu'}\}$ are related by a constant matrix $T^{-1}$ as follows,

$$\partial_{\mu'} = g_{\mu\rho}T^\rho{}_\tau g^{\tau\nu}\partial_\nu = (T^{-1})^\mu{}_\nu \partial_\nu \equiv T_\mu{}^\nu \partial_\nu, \tag{2.8}$$

then we say that the transformation is linear, since in this case, unlike the general case with equation (2.6), it *is* true that for some $a, b \in \mathbf{R}$

$$a\partial_{\mu'} + b\partial_{\rho'} = a(T^{-1})^\mu{}_\nu \partial_\nu + b(T^{-1})^\rho{}_\sigma \partial_\sigma = (a\partial_\mu + b\partial_\rho)'. \tag{2.9}$$

The vector components transform like

$$x^{\mu'} = T^\mu{}_\nu x^\nu. \tag{2.10}$$

Inverting this equation goes like

$$\begin{aligned}\underbrace{\left(g_{\mu\rho}T^{\rho}{}_{\tau}g^{\tau\sigma}\right)}_{(T^{-1})^{\sigma}{}_{\mu}} x^{\mu'} &= \left(g_{\mu\rho}T^{\rho}{}_{\tau}g^{\tau\sigma}\right) T^{\mu}{}_{\nu}x^{\nu} \\ &= \delta^{\sigma}_{\mu}x^{\mu} \\ &= x^{\sigma}\end{aligned} \tag{2.11}$$

Hence

$$x^{\sigma} = \left(g_{\mu\rho}T^{\rho}{}_{\tau}g^{\tau\sigma}\right) x^{\mu'} \equiv T_{\mu}{}^{\sigma}x^{\mu'}. \tag{2.12}$$

In the next section concerning induced transformations, we will regard the linear transformations only, but keep in mind that they are a special subset of the general transformations.

## 2.3 Induced transformations

Transformations on spacetime in general induce transformations on fields, basically because of their dependence on the spacetime coordinate. How an object transforms under spacetime transformations defines the type of field one is dealing with. One can distinguish scalar fields an vector fields for example, which are all example of the more general *tensor* fields.

### 2.3.1 Induced transformations on scalar fields

A scalar $\phi$ is a mathematical object that has a magnitude but no direction. Physical examples include mass and energy. A scalar field $\phi(x^{\mu})$ then is some object which has a magnitude, but no direction, in every point on a manifold $M$ with coordinate $x^{\mu}$. However the scalar field $\phi$ itself is independent of whatever basis $\partial_{\mu}$, or equivalently whatever coordinates $x^{\mu}$, we like to use to express its argument $x^{\mu}$ in. We therefore will demand that two scalar fields $\phi(x^{\mu})$ and $\phi'(x^{\mu'})$, in two bases $\partial_{\mu}$ and $\partial_{\mu'}$ respectively related by equation(2.8), to be the same:

$$\phi'(x^{\mu'}) = \phi(x^{\mu}) \tag{2.13}$$

This equation defines any general scalar field. Rewriting equation (2.13), using equation(2.11), we get

$$\phi'(x^{\mu'}) = \phi((T^{-1})^{\mu}{}_{\nu}x^{\nu}), \tag{2.14}$$

which is true for the whole infinite space $\mathbf{R}'^{n} \ni x^{\mu'}$, so we may as well drop the accents on the coordinates and get

$$\phi'(x^{\mu}) = \phi((T^{-1})^{\mu}{}_{\nu}x^{\nu}). \tag{2.15}$$

This is what is called an *induced transformation*. Until now we didn't mention if the coordinate transformation $T$ is actually a symmetry transformations and with good reason, since equation (2.15) holds whether or not $T^{-1}$, and thus $T$ also, is a symmetry transformation. It merely defines how a scalar field itself transforms $\phi \to \phi'$ after a coordinate transformation $x \to x' = Tx$.

### 2.3.2 Induced transformations on co- and contravariant vector fields

We saw the distinction between co- and contravariant vectors. We now discuss the linear transformation rules for contravariant vector fields. A contravariant vector field $V(x^\mu)$ is, analogously to the scalar field, a vectorial entity given at all points $p$ of some manifold $M$ with coordinate $x^\mu(p)$. Momentum and the electromagnetic vector potential are famous examples. Like the scalar field, the vector field itself is in general independent on the choice of coordinates, but unlike the scalar field a vector field has components that do depend on the coordinates, as is explicitly stated in equation(2.10) for a contravariant vector. The only consequence for the vector field being a field would be that any coordinate transformation induces a transformation on the vector field, analogous to the scalar field.

By demanding the vector field $V(x^\mu)$, like the scalar field, to be invariant under the transformation $x^\sigma \to x^{\sigma'} = T^{\sigma'}{}_{\sigma} x^\sigma$, i.e.

$$V^{\mu'}(x^{\sigma'})\partial_{\mu'} = V^\mu(x^\sigma)\partial_\mu, \tag{2.16}$$

and using equation (2.8) one derives how the vector field components transforms under a coordinate transformation $x^\sigma \to x^{\sigma'}$, i.e.

$$V^{\mu'}(x^{\sigma'}) = T^\mu{}_\nu V^\nu(x^\sigma) = T^\mu{}_\nu V^\nu((T^{-1})^\sigma{}_\nu x^{\sigma'}) \tag{2.17}$$

again holding for all $x^{\sigma'}$, therefore dropping the accents on the coordinate, we get

$$V^{\mu'}(x^\sigma) = T^\mu{}_\nu V^\nu((T^{-1})^\sigma{}_\nu x^\sigma). \tag{2.18}$$

The total transformation therefore can be regarded to have a coordinate part, letting you calculate the vector in the old coordinates in the point $(T^{-1})^\sigma{}_\nu x^\sigma$, followed by a component part $T^\mu{}_\nu V^\nu$, which in total gives you the components of the vector in the new coordinates in the point $x^\sigma$. The transformation rules for covariant vectors can be similarly derived. The are given by

$$\omega_{\mu'}(x^\sigma) = (T^{-1})^\mu{}_\nu \omega_\nu((T^{-1})^\sigma{}_\nu x^\sigma). \tag{2.19}$$

### 2.3.3 Induced transformations on tensorfields

The generalization of scalar fields, covariant and contravariant vector fields is the *tensor* field. To define a tensor we first need to define the operation of a *tensor product* $\otimes$. A tensor product $V \otimes \omega$ between a contravariant vector $V$ and a covariant vector $\omega$ is defined as

$$V \otimes \omega = (V^\nu \partial_\nu) \otimes (\omega_\mu dx^\mu) \equiv V^\nu \omega_\mu (\partial_\nu \otimes dx^\mu) = A^\nu{}_\mu (\partial_\nu \otimes dx^\mu), \tag{2.20}$$

where $A^\nu{}_\mu$ are the components of a so called rank $(1,1)$ tensor, since it has one upper and one lower index. This in contrast to the *inner product* $V \cdot \omega$, which is defined by

$$V \cdot \omega = (V^\nu \partial_\nu) \cdot (\omega_\mu dx^\mu) \equiv V^\nu \omega_\mu (\partial_\nu \cdot dx^\mu) = V^\nu \omega_\mu \delta^\mu_\nu = V^\mu \omega_\mu \in \mathbf{R}, \tag{2.21}$$

and is a (0,0) tensor or scalar as will be clear in a moment. A general tensor field of rank $(k, l)$ is a mathematical object $A(x^\sigma)$ that is defined as being a tensor product of $k$ contravariant and $l$ covariant vector fields at spacetime point $x^\sigma$

$$A(x^\sigma) = A^{\nu_1 \dots \nu_k}{}_{\mu_1 \dots \mu_l}(x^\sigma)(\partial_{\nu_1} \otimes \dots \otimes \partial_{\nu_l} \otimes dx^{\mu_1} \otimes \dots \otimes dx^{\mu_k}), \tag{2.22}$$

having components $A^{\nu_1 \dots \nu_k}{}_{\mu_1 \dots \mu_l}(x^\sigma)$, which transform as a direct product of $k$ contra and $l$ covariant vectors

$$A^{\nu'_1 \dots \nu'_k}{}_{\mu'_1 \dots \mu'_l}(x^\sigma) = T^{\nu'_1}{}_{\nu_1} \dots T^{\nu'_k}{}_{\nu_k} (T^{-1})^{\mu'_1}{}_{\mu_1} \dots (T^{-1})^{\mu'_l}{}_{\mu_l} A^{\nu_1 \dots \nu_k}{}_{\mu_1 \dots \mu_l}((T^{-1})^\sigma{}_\nu x^\sigma) \tag{2.23}$$

The set of all $(k, l)$ tensors forms a vector space, with a *tensor product* of $k$ contra- and $l$ covariant bases as a basis.

A scalar field $\phi(x^\sigma)$ therefore is a $(0, 0)$ tensor, since it transforms as one and covariant tensor fields $\omega(x^\sigma)$ and contravariant tensor fields $V(x^\sigma)$, are $(0, 1)$ and $(1, 0)$ tensor fields respectively.

As will be explained in great detail in the next sections, all symmetry transformations of a given system in general form a *group*. It is possible that some coordinate transformations $x^\sigma \to x^{\sigma'} = \tilde{T}^{\sigma'}{}_\sigma x^\sigma$ leave the tensor field *invariant*, i.e.

$$A^{\nu'_1 \dots \nu'_k}{}_{\mu'_1 \dots \mu'_l}(x^\sigma) = A^{\nu_1 \dots \nu_k}{}_{\mu_1 \dots \mu_l}(x^\sigma). \tag{2.24}$$

As a side note, to be used later on, transformations $\tilde{T}$ leaving the tensor field invariant form a special group called the *little group* of that particular tensor field.

## 2.4 Inertial observers

To understand what we mean by 'observing the same physical laws' in more concrete mathematical terms, we first must talk about *inertial frames of reference*. How are they defined, from Newton's definitions to Einstein's modifications resulting in the special and general theory of relativity. In the end we would like to find out how to describe the motion of an inertial observer in curved space and how two such observers are related.

An inertial observer first of all, is an observer that for simplicity is fixed to the origin of some inertial frame of reference, which is a set of coordinates $\mathbf{x} = x^k \partial_k$ on the manifold we call spacetime, with particular important properties. Newton's definition [3], [4] of an inertial frame is any in frame $M$ in which the trajectory of a free particle, i.e. no interactions, is described by a straight line. Hence we can define a Newtonian inertial frame as follows:

**Definition 2.1.** *Newtonian inertial frame*

A *Newtonian inertial frame* $M$ is any frame such that a trajectory $\mathbf{x}(t)$ for a free particle, $\sum \mathbf{F} = 0$, is given by

$$\mathbf{x}(t) = \dot{\mathbf{x}}t + \mathbf{x}(0), \tag{2.25}$$

with $\dot{\mathbf{x}} = d\mathbf{x}/dt \equiv \mathbf{v}$ the *constant velocity* and $\mathbf{x}(0)$ the position at $t = 0$ of the object.

Another way to put it is that an inertial frame is any frame for which Newton's second law $\sum \mathbf{F} = \dot{\mathbf{p}} = m\ddot{\mathbf{x}}$ holds. Clearly equation (2.25) solves Newton's second law for free particles. Newton's first law, stating that $\ddot{\mathbf{x}} = 0$ must hold for a free particle, then states that such an inertial frame, the one in which $\mathbf{v}$ is measured, exists.

When we talk about the 'same physical laws' of nature, we actually mean the *equations of motion* (e.o.m.) that describe the evolution of the system under investigation. The Newtonian way of deriving the e.o.m. comes down to gathering all forces $\mathbf{F}$, resulting from all physically interactions acting on the particle and then writing down

$$\sum \mathbf{F} = m\ddot{\mathbf{x}}. \tag{2.26}$$

These are the *Newtonian* e.o.m. . In general the e.o.m. fix our physics.

It is called inertial because only in such a frame the *inertia* $m$, known as the mass, can be measured. This can be demonstrated by going to an explicit *non*-inertial frame. Let's accelerate a frame $\tilde{M}$ with respect to an inertial frame $M$ in the $+\hat{x}$ direction by an amount of $a$ and let's focus on this direction only forgetting about the $\hat{y}$ and $\hat{z}$ directions. Then $\tilde{x}$ of $\tilde{M}$ is related to the coordinate $x$ of $M$ by

$$\tilde{x} = x - \frac{1}{2}at^2 \tag{2.27}$$

Hence the trajectory $\tilde{x}(t)$ of a free particle, for which in the inertial frame $\sum \mathbf{F} = 0$ holds, would in the accelerated frame result in $\tilde{x}(t) = -\frac{1}{2}at^2 + \dot{x}t + x(0)$, indicating the presence of a *fictitious* force $\tilde{\mathbf{F}} = \tilde{m}a$.

These frames of reference clearly do not describe the same physics, since in the latter the particle cannot be described as 'free' any longer. In this frame of reference a free particle is described by $\tilde{F} \neq 0$ An observer in the inertial frame $M$ could try to measure the mass $m$ of the particle by applying a *real* force $\mathbf{F} = F\hat{x}$ to the particle, measuring the acceleration $a_0$ in his inertial frame of reference and deriving that the mass $m$ of the particle hence must be equal to

$$m = \frac{F}{a_0}. \tag{2.28}$$

An observer in $\tilde{M}$ however, unaware of the fact that his frame of reference is accelerated could apply the same force $\mathbf{F} = F\hat{x}$ to the particle, but would measure an acceleration $\tilde{a} = a + a_0$ leading to the false conclusion that the mass $\tilde{m}$ should be equal to

$$\tilde{m} = \frac{F}{\tilde{a}} = \frac{F}{a + a_0}, \tag{2.29}$$

since he doesn't know that he is the one that is being accelerated.

Now for two observers with frames of reference $M$ and $M'$ to find the same physics, i.e. the same e.o.m in their respective coordinates, they must be related by special transformations that take you from the observer in $M$ to the observer in $M'$. By inspection one can derive the transformations $\{T\}$ that would relate two Newtonian inertial frames of reference: Let's say that $M$ is an inertial frame of reference and an *event* in $M$ has coordinates $(\mathbf{x}, t)$. We're adopting the so called *passive perspective* on coordinate transformations by describing the *same* event in $M'$ by new coordinates

$(\mathbf{x}', t')$. Let's look at what happens when these sets of coordinates are related by a set of special transformations $\{T_k\}$, with

**Definition 2.2.** *Galilean Transformations* $\{T_k\}$

$$(\mathbf{x}, t) \to T_1[(\mathbf{x}, t)] = (\mathbf{x}', t') = (\mathbf{R}\mathbf{x}, t) \tag{2.30}$$
$$(\mathbf{x}, t) \to T_2[(\mathbf{x}, t)] = (\mathbf{x}', t') = (\mathbf{x} + \mathbf{u}t, t) \tag{2.31}$$
$$(\mathbf{x}, t) \to T_3[(\mathbf{x}, t)] = (\mathbf{x}', t') = (\mathbf{x} + u_0, t + t_0) \tag{2.32}$$

where $\mathbf{R}$ is some constant orthogonal $3 \times 3$ matrix, representing a *fixed* rotation of $M'$ with respect to $M$ ($M'$ is explicitly not rotat*ing*), $\mathbf{u}$ represents the constant relative velocity of $M'$ with respect to $M$ and $\mathbf{u}_0$ and $t_0$ are respectively a fixed spatial and time translations from $M$ to $M'$. Provided that $\mathbf{F}$ transforms as a 3-vector under rotations, the equation $\sum \mathbf{F} = m\ddot{\mathbf{x}}$ transforms into $\sum \mathbf{F}' = m\ddot{\mathbf{x}}'$ and one says that they transform *covariantly*: they have exactly the same form under this particular set of transformations. It implies that both $M$ and $M'$ are Newtonian inertial frames of reference, exactly what we hoped for.

Einstein noticed that not all physical laws were covariant under the Galilean transformations. Especially electromagnetic waves behaved strangely: they had the same speed in *any* Newtonian inertial frame of reference, which in particular broke the covariance under the Galilean "boost" (2.31). Instead of claiming that the laws describing electromagnetism were false, Einstein postulated that in fact our notion of time and space had to be redefined. This resulted in the *special theory of relativity*. The Galilean transformations are the non-relativistic equivalent of the relativistic *Poincaré* transformations. These define new inertial frames in which the speed of light always has the same value $c$, which will be treated more extensively in chapter 5.

But the "method of inspection" used here to make the galilean transformations plausible is of course far from satisfactory. We would like to derive a general recipe for relating two inertial observers to each other, given a certain spacetime. Inspection for example does not answer *why* exactly the Galilean transformations relate two inertial observers, it just states that they do. I will now try to motivate that to find this recipe, the metric $g_{\mu\nu}$ is the magic wand. To do this we must use more modern approaches using Lagrangian mechanics. Both relativistic and non relativistic e.o.m. can be derived from an *action* using the *variational principle*. I will now skip a few, or better, a lot of steps, but any good book on Classical dynamics like [4] will treat it and these lecture notes [5] and the closely related [6] treat it extensively.

The basic idea is that we start from an action $\mathcal{S} = \int d^n x \mathcal{L}$ in an $n$-dimensional space, where $\mathcal{L}$ is the Lagrangian density. We then derive the equations of motion, using the variational principle accumulating into the Euler-Lagrange equations, given here for the dynamics of a free complex scalar field $\phi$:

$$\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{2.33}$$

The idea is now that $\mathcal{L}$ in general depends on the metric $g_{\mu\nu}$ , which function it is to define an *inner product* $g_{\mu\nu} x^\mu x^\nu$ on the spacetime of interest, from which physically

relevant concepts like length and angles can be derived. Causal structures also depend on the metric and for this to remain invariant under a transformation, we have to make sure that $g_{\mu\nu}$ is invariant. In Appendix A one can find some more information on the definition of the metric in general. It suffice to say that given any transformation leaving the metric $g_{\mu\nu}$ invariant should be considered as a symmetry of the theory, meaning that the equations of motion transform covariantly under these transformations, like we saw in the case of the Galilean transformations and Newton's equations of motion in classical dynamics.

The free complex scalar field for example the Lagrangian looks like

$$\mathcal{L} = \frac{1}{2} g^{\mu\nu} \partial_\nu \phi^* \partial_\mu \phi - \frac{1}{2} m^2 \phi^2. \tag{2.34}$$

and is clearly dependent on the metric. But also in classical mechanics we use metric all the time. A classical Lagrangian $L = \int \mathcal{L}\, d^3x$ a massive particle would in general be of the form $L \sim \mathbf{v}^2 = v_i v^i$, but this is merely the Euclidean inner product $\delta_{ij} v^i v^j$, where $\delta_{ij}$ is the Kronecker delta and is regarded as the metric of classical Euclidean space.

As was mentioned there are in general two types of transformations that one can perform on a given $\mathcal{L}$ : *spacetime* transformations involving the spacetime coordinates $(x, t)$ and *internal* transformations acting on the "inner space" for example by replacing $\phi$ by its complex conjugate $\phi^*$ for complex scalar fields. We will focus solely on the spacetime transformations. Given a spacetime with a metric $g_{\mu\nu}$ the question how two observers should be related if they were to measure the same physical laws, can be recast into terms involving the metric $g_{\mu\nu}$ , since the *invariance* of this metric under these transformations guarantees the covariance of the equations of motion in the two frames of reference. The question thus becomes: Given a spacetime with metric $g_{\mu\nu}$ , what transformations leave $g_{\mu\nu}$ invariant? To answer this quite general question we need some more tools. They are provided to us by General Relativity, the theory that relates spacetimes and their metrics to energy and matter. Besides the many great books written on this subject, for me the lecture notes [5] and [6] were particularly useful.

## 2.5 General Relativity

General Relativity is the theory that relates the metric of a space time to the energy density present. We need to borrow some concepts from it and therefore this quick fly through is al but self containing. In general relativity the metric tensor is regarded as a tensor field $g_{\mu\nu} = g_{\mu\nu}(x)$. It describes the spacetime at the point $x$ and from it one can derive the curvature of the spacetime, if present. The central equations in GR are the *Einstein equations* and they are given by

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G_N T_{\mu\nu} \tag{2.35}$$

$R_{\mu\nu}$ is called the Ricci tensor and is a contraction of the most important tensor of GR, the *Riemann-curvature tensor* $R^\lambda{}_{\sigma\mu\nu}$ and $R$ is the contraction of $R_{\mu\nu}$ called the Ricci scalar. The Riemann-curvature tensor $R^\lambda{}_{\sigma\mu\nu}$, which as the name suggests gives us a way to define the curvature of a given space and is in its full glory given by

$$R^{\lambda}{}_{\sigma\mu\nu} = \partial_\mu \Gamma^{\lambda}_{\sigma\nu} - \partial_\nu \Gamma^{\lambda}_{\sigma\mu} + \Gamma^{\lambda}_{\mu\rho}\Gamma^{\rho}_{\nu\sigma} - \Gamma^{\lambda}_{\nu\rho}\Gamma^{\rho}_{\mu\sigma} \tag{2.36}$$

and $\Gamma^{\mu}_{\nu\lambda}$ is the *Christoffel connection* and, when derived from the metric, which we will use here it is given by

$$\Gamma^{\mu}_{\nu\lambda} = \frac{1}{2} g^{\mu\rho} \left( \partial_\lambda g_{\rho\nu} + \partial_\nu g_{\rho\lambda} - \partial_\rho g_{\nu\lambda} \right). \tag{2.37}$$

We can define $\Gamma^{\mu}_{\nu\lambda}$ to be torsion free, i.e. it is symmetric in its lower indices $\Gamma^{\mu}_{\nu\lambda} = \Gamma^{\mu}_{(\nu\lambda)}$. So putting it all together we see from the left hand side of equation (2.35) that the Einstein equations are basically second order differential equations of the metric $g_{\mu\nu}$ . On the right hand side of the Einstein equations $T_{\mu\nu}$ represents the energy-momentum density of the spacetime of interest. We are however interested in the vacuum solutions to the Einstein equation, since we just want to study the effects of 'empty' space on the particles. So we put $T_{\mu\nu}$ to 0. However, one can easily show [5], [6] that by deriving the Einstein equations from an action principle (so called Einstein-Hilbert action), an additional scalar can be added to the Lagrangian density, resulting in the modified vacuum Einstein equations:

$$\boxed{R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = g_{\mu\nu} \Lambda} \tag{2.38}$$

where $\Lambda$ is the *cosmological constant* which can be interpreted as the energy density of the vacuum and gives rise to a curvature of the universe. We know from experiments [2], [1] that $\Lambda$ actually is very small, and positive making the universe a *de Sitter* space.

Back to the observers: we want to define a family of frames of reference in which one is sure to measure the same physical laws. Einstein shows us that because of the empirical fact that inertial and gravitational mass are equal an experiment cannot distinguish between an accelerated laboratory of a fixed laboratory in the presence of a gravitational field. Doing experiments in a free falling laboratory in the presence of a gravitational field is therefore equivalent to doing them in a lab in the absence of gravity, albeit *local* such that possible non-uniformity of the gravitational field has no influence. This is called *Einstein's principle of equivalence* and can be put in the following way.

**Definition 2.3.** *Einstein principle of Equivalence*
Any experiment done locally in a free falling frame of reference should obtain the same results as the same experiment done in an inertial frame in the absence of gravity.

To make sure, the term inertial here is the relativistic variant of Newton's inertial frame. The relativistic inertial frames are related by Poincaré transformations as mentioned, but we will try to derive this fact from symmetry considerations. The Einstein principle of equivalence can also be translated into the statement that any for any curved spacetime, it's local tangent space is always a flat spacetime with Minkowski metric.

The Einstein definition of an inertial observer is given by a special trajectory in curved space, a *geodesic*. It is the exact curved space equivalent of a straight line and is derived by demanding that the second derivative of $x$ with respect to time, transform

covariantly under general coordinate transformations. This guarantees its validity in all inertial frames of reference. The geodesic equation is again a differential equation, which solutions $x^\mu$ are the geodesics. This leads to the following definition of an inertial observer in GR.

**Definition 2.4.** *Einstein inertial observer*
An *Einstein inertial observer* is an observer that follows a geodesic in spacetime. A geodesic is a path $x^\mu$ which solves the geodesic equation

$$\ddot{x}^\mu + \Gamma^\mu_{\nu\lambda}\dot{x}^\nu\dot{x}^\lambda = 0. \tag{2.39}$$

In fact in flat space $\Gamma^\mu_{\nu\lambda} \to 0$ and the geodesic equations return to our familiar Newton equations for an inertial observer: $\ddot{x}^\mu = 0$. We conclude that in general curved spacetimes what observers should agree upon depends on the background. To answer the question which transformations relate two inertial observers in the presence of gravity, i.e. in curved space, we need a way to compare the metric tensor $g_{\mu\nu}$ from point to point. For this we need the *covariant derivative* and the *Lie derivative*.

### 2.5.1 Covariant derivatie, Lie derivative and the Killing equations

**Definition 2.5.** *Covariant Derivative*
The covariant derivative is defined as the differential operator $\nabla_\mu$ which transforms covariantly under general coordinate transformations. Given a metric $g_{\mu\nu}$ and derived Christoffel symbol $\Gamma^\mu_{\nu\lambda}$, the explicit form of the covariant derivative of a (1,1) tensor $T^\mu_\nu$ is given by

$$\nabla_\mu T^\nu_\rho = \partial_\mu T^\nu_\rho + \Gamma^\nu_{\mu\lambda}T^\lambda_\rho - \Gamma^\lambda_{\mu\rho}T^\nu_\lambda \tag{2.40}$$

and can be generalized easily to an arbitrary (p,q) tensor. It can be shown [5],[6] that in general $\nabla_\mu$ or actually $\Gamma^\mu_{\nu\lambda}$ can chosen to be *metric compatible*, i.e. $\Gamma$ is chosen such that $\nabla_\mu$ commutes with the metric $\nabla_\mu g_{\rho\sigma} = 0$.

The definition of the Lie derivative requires that we compare the metric field $g_{\mu\nu}(x)$ at two infinitesimally separated places. The metric $g_{\mu\nu}(x)$ being an (0,2) tensor transforms under general coordinate transformations $x^\nu \to y^\mu(x^\nu)$ as

$$g'_{\mu\nu}(y(x)) = \frac{\partial x^\rho}{\partial y^\mu}\frac{\partial x^\lambda}{\partial y^\nu}g_{\rho\lambda}(x) \tag{2.41}$$

where $g'_{\mu\nu}(y(x))$ is a *new* metric tensor, namely the one at $y(x)$. Now regard the difference in line elements $ds^2$ between the spacetime points $y(x)$ and $x$

$$g'_{\mu\nu}(y)dy^\mu dy^\nu - g_{\mu\nu}(x)dx^\mu dx^\nu. \tag{2.42}$$

We know that the line element $ds^2$ is a scalar under general coordinate transformations. This means that under a passive coordinate transformation $x \to y(x)$, meaning that we regard the *same* metric, yet where using a different set of coordinates $y$, the line element transforms like

$$g_{\mu\nu}(x)dx^\mu dx^\nu \to g_{\mu\nu}(y)dy^\mu dy^\nu = g_{\mu\nu}(x)dx^\mu dx^\nu \tag{2.43}$$

as simple as that. Hence equation (2.42) can be written as

$$g'_{\mu\nu}(y)dy^\mu dy^\nu - g_{\mu\nu}(y)dy^\mu dy^\nu = \left(g'_{\mu\nu}(y) - g_{\mu\nu}(y)\right) dy^\mu dy^\nu \tag{2.44}$$

Hence now we're comparing two metrics at the same point $y$: $g'_{\mu\nu}(y)$ is the metric acquired after *actively* taking the metric from $x \to y = y(x)$, while $g_{\mu\nu}(y)$ is the metric at $y$ acquired by the *passive* relabeling $x \to y = y(x)$. Now if $g'_{\mu\nu}(y) = g_{\mu\nu}(y)$ this difference of course would identically vanish, meaning that the particular coordinate transformation $x \to y(x)$ left the metric *invariant*, hence $x \to y(x)$ is a *symmetry* of the metric.

**Definition 2.6.** *Isometry*

Any transformation $x \to y(x)$ that leaves the metric invariant is called an *isometry*. Since the metric $g_{\mu\nu}(x)$ determines all properties of the spacetime at the given point $x$, these isometries in general leave the spacetime invariant, guaranteeing covariance of the equations of motion, i.e. they imply a symmetry of any physical theory.

We now would like to find out how to *derive* these isometries. Given that these isometries are continuous coordinate transformations $x \to y(x)$, we also could take the infinitesimal transformation to first order:

$$y^\mu = x^\mu + \delta x^\mu \equiv x^\mu + \epsilon V^\mu(x). \tag{2.45}$$

where $V^\mu(x)$ can be defined, and therefore is, as a vector field. This is so because any infinitesimal coordinate transformation $\delta x^\mu$ transforms as a vector

$$\delta y^\mu = \delta y^\mu(x) \to \delta y'^\mu = \frac{\partial y^\mu}{\partial x^\nu}\delta x^\nu \tag{2.46}$$

For a vector field $W^\mu$ the Lie derivative is defined as

**Definition 2.7.** *Lie derivative $L_V$ of a vector field $W^\mu$*

$$L_V W^\mu = \lim_{\epsilon\to 0}\frac{W^\mu(y(x)) - W'^\mu(y(x))}{\epsilon} \tag{2.47}$$

with $y(x)$ of course given by the infinitesimal transformation (2.45). $W'^\mu(y(x))$, being a transformed vector is given by

$$\begin{aligned} W'^\mu(y(x)) &= \frac{\partial y^\mu}{\partial x^\nu}W^\nu(x) \\ &= (\delta^\mu_\nu + \epsilon\partial_\nu V^\mu(x))W^\nu(x) \\ &= W^\mu(x) + \epsilon W^\nu(x)\partial_\nu V^\mu(x) \end{aligned} \tag{2.48}$$

$W^\mu(y(x))$, on the the other hand, is merely the same vector field described in another coordinate and is therefore given by the Taylor expansion around $x$

$$W^\mu(y(x)) = W^\mu(x) + \epsilon V^\nu\partial_\nu W^\mu(x) + \mathcal{O}(\epsilon^2). \tag{2.49}$$

Substituting these expressions into equation (2.47) results in

$$L_V W^\mu = V^\nu\partial_\nu W^\mu - W^\nu\partial_\nu V^\mu. \tag{2.50}$$

Replacing $\partial_\mu$ by $\nabla_\mu$ we make the already covariant equation, explicitly covariant.

Solving the similar problem for covariant vectors $W_\mu$, we can generalize and finally give the the Lie derivative $L_V$ of a tensorfield $T(y(x))$

**Definition 2.8.** *Lie derivative $L_V$ of a vector field $T^{\mu_1\dots\mu_l}{}_{\nu_1\dots\nu_k}$*

$$\begin{aligned} L_V\, T^{\mu_1\dots\mu_l}{}_{\nu_1\dots\nu_k} &= V^\rho \nabla_\rho T^{\mu_1\dots\mu_l}{}_{\nu_1\dots\nu_k} \\ &- T^{\rho\dots\mu_l}{}_{\nu_1\dots\nu_k} \nabla_\rho V^{\mu_1} - \dots - T^{\mu_1\dots\rho}{}_{\nu_1\dots\nu_k} \nabla_\rho V^{\mu_l} \\ &+ T^{\mu 1\dots\mu_l}{}_{\rho\dots\nu_k} \nabla_{\nu_1} V^\rho + \dots + T^{\mu 1\dots\mu_l}{}_{\nu_1\dots\rho} \nabla_{\nu_l} V^\rho \end{aligned} \tag{2.51}$$

For the metric $g_{\mu\nu}$ , which is (0,2) tensor this general result (2.51) gives us

$$\begin{aligned} L_V\, g_{\nu\lambda} &= V^\rho \underbrace{(\nabla_\rho g_{\nu\lambda})}_{0} + g_{\rho\lambda}\,(\nabla_\nu V^\rho) + g_{\nu\rho}\,(\nabla_\lambda V^\rho) \\ &= \nabla_\nu V_\lambda - \nabla_\lambda V_\nu \end{aligned} \tag{2.52}$$

Demanding that the transformation (2.45) is an isometry, means that the Lie derivative must vanish. We conclude that this is equivalent to stating that the vector field $V^\mu$ must solve the so called *Killing equations*.

$$\boxed{L_V\, g_{\nu\lambda} = 0 \Leftrightarrow \nabla_\nu\, V_\lambda + \nabla_\lambda\, V_\rho = 0} \tag{2.53}$$

Any vector $V^\mu$ that solves equation (2.53) is correspondingly called a *Killing vector*. There is a theorem [6] that relates the maximum number of independent Killing vector fields to the dimensionality of the geometry $D$.

**Theorem 2.9.** The maximum number $n$ of independent Killing vector fields, which span the isometry group of a spacetime with metric $g_{\mu\nu}$ of dimensionality $D$ is given by

$$n = \frac{D(D+1)}{2} \tag{2.54}$$

In chapter 5 we will see that for the Minkowski spacetime $D = 4$ of course, and therefore $n = 10$, which coincides with the number of generators of the Poincaré group. Therefore the Poincaré group is called a *maximally symmetric spacetime* in $D = 4$. The other maximal symmetric spacetimes are the de Sitter and anti de Sitter spacetime. All three are solutions to equations (2.38) for $\Lambda = 0$, $\Lambda > 0$ and $\Lambda < 0$ respectively.

Looking back at that the expression $L_V\, W^\mu$ (2.50) is antisymmetric in the indices $\mu$ and $\nu$, hence one can associate it with a commutator

$$V^\nu \partial_\nu W^\mu - W^\nu \partial_\nu V^\mu = [V, W]^\mu = L_V\, W^\mu = -L_W\, V^\mu \tag{2.55}$$

This commutator is a Lie bracket since it can be shown to satisfy the Jacobi identity, which we will encounter in chapter 4.3 were it is a requirement for the operators to generate a so called *Lie algebra*. The Lie algebra generated by Lie derivatives is generates the group of general coordinate transformations, also known as the group of *diffeomorphisms*.

Furthermore one can show that the commutator of two Lie derivatives themselves is again a Lie derivative. The next identity expresses this fact

$$[L_V, L_W] = L_{[V,W]} \tag{2.56}$$

This implies that if $V$ and $W$ are Killing vectors, $[V, W]$ is a Killing vector too, ensuring closure of the Lie algebra (see also section 4.3). All in all, this means that the Killing vectors are the generators of a symmetry group. For Killing vectors this is per definition the *Isometry group* of the particular spacetime.

So we conclude that we have found the recipe to find the isometry group of a given spacetime. These transformations generated by th Killing vectors are the fundamental spacetime symmetries we were looking for: They relate two inertial observers, defined by the spacetime. The recipe is:

- Solve the modified Einstein vacuum equations (2.38) for $g_{\mu\nu}(x)$ for certain value for $\Lambda$.
- Then, given the metric, solve the Killing equations (2.53) to determine the Killing vectors $V^\mu$
- $V^\mu$ form a *Lie algebra* of the isometry group $\Rightarrow$ $V^\mu$ they are the *generators* of the isometry group

As an example we can solve the Killing equations for the Minkowski metric $\eta_{\mu\nu}$, which is a flat solution to the vacuum Einstein equation ($\Lambda = 0$). Because for a flat space the Christoffel symbols identically vanish and we therefore have

$$\nabla_\nu V_\lambda + \nabla_\lambda V_\rho = 0 \Rightarrow \partial_\mu V_\nu + \partial_\nu V_\mu = 0. \tag{2.57}$$

This has the following solution

$$V^\mu = \delta x^\mu = (\omega)^\mu{}_\nu x^\nu + \epsilon^\mu \tag{2.58}$$

which we will recognize in chapter 5 as the Poincaré transformations.

But why does a symmetry of the spacetime dictate what representations of the particles are possible? In quantum mechanics the wavefunctions $|\psi\rangle$ represent all possible states of some system. The states of a free particle for example, form a linear in general infinite dimensional vector space, endowed with an inner product $\langle\phi|\psi\rangle$, i.e. a Hilbert space (see Appendix A). Two inertial observers looking at the same particle in general will describe it by two different wavefunctions $|\psi\rangle$ and $|\psi'\rangle$. They however must observe the same physical laws, since the particle doesn't care in what set of coordinates it is being described in. This means that

$$\langle\phi|\psi\rangle = \langle\phi'|\psi'\rangle \tag{2.59}$$

must hold, which is equivalent to stating that $|\psi\rangle$ and $|\psi'\rangle$ in Hilbert space must be related by a *unitary* transformation $U(T)$ depending on the isometry transformation $\{T\}$ that relates two inertial observers in the particular spacetime. We say that $U(T)$ must form a *unitary representation* of that isometry group. Hence to find out what the observers can suspect to measure in experiments involving free particles, we need to take a closer look at group theory and representation theory.

## 2.6 Symmetry operations form a group

Symmetry operations are transformations that leave a certain object invariant. This object of interest could for example be a geometrical figure like a square or a vector. What these operations are that leave the object invariant, of course totally depends on what object we're looking at. Therefore the set of symmetry operations is associated with that object and called the symmetry set or group of that object. One can imagine that a square has a certain set of symmetry operations associated with it and so does a circle, albeit a totally different set. In the above discussion for example we investigated the symmetry operations of a given spacetime, called isometries. We will now discuss why any set of symmetry operations, including the isometry transformations, form in general a special set called a *group* and what group theory could tell us about elementary particles.

If a first symmetry operation leaves the object invariant, a second symmetry operation, applied after the first one, leaves the object invariant too. These two operations therefore can be combined into a third operation, the composed operation: 'do operation two after operation one', inevitably leaving the object again invariant and thus being a symmetry operation itself. When we consider the set of all operations leaving an object invariant, it must be that this composed operation is in fact a member of this set, which thereby is called 'closed' under this particular composition law, or more general 'multiplication law'. Also each symmetry operation is one that can be undone, by the inverse operation. If an operation is composed with its inverse operation, the net result would be 'to do nothing' which clearly is a symmetry operation, if there ever was any, since it per definition leaves the object invariant. So if an operation is a symmetry operation, its inverse must be so too. This set of symmetry operations together with the rules they must obey forms what is called a *group* and can be nicely summarized in the next set of axioms. There are lots of books written about group theory, reaching from more mathematical[7], [8] to more physical [9], [10] approaches. This text is just a quick reminder and does not serve as a complete overview. I will only highlight the parts important to my research, but some general definitions are necessary to follow the arguments.

**Definition 2.10.** *Group;* $G \equiv (G, \circ)$

A set of elements $G = \{g_1, g_2, \dots\}$ together with a composition law $\circ$ is called a *group* $(G, \circ)$, which we from now on will write in short as $G$, if it obeys the following rules:

1. (*Closure*) If $g_i, g_j \in G$ then $g_i \circ g_j = g_k \in G$
2. (*Associativity*) $(g_i \circ g_j) \circ g_k = g_i \circ (g_j \circ g_k)$
3. (*Existence of identity*) $\exists e \in G$ such that $\forall g \in G, g \circ e = e \circ g = g$
4. (*Existence of inverse*) $\forall g_i \in G, \exists g_i^{-1} \in G$ such that $g_i \circ g_i^{-1} = g_i^{-1} \circ g_i = e$

Another important fundamental property of groups is that they can be *abelian* or *non-abelian*. The demand for a group to be abelian is that its elements *commute*, i.e. $g_i \circ g_j = g_j \circ g_i, \forall g_i, g_j \in G$. This mathematical definition physically means that the order of applying the symmetry operations to the object doesn't matter. Commuting operations are actually more an exception then a rule, resulting in more non-abelian

then abelian groups. The groups that are abelian have in general a much simpler structure as we will see.

# 3 Discrete groups

One can make a first distinction between two types of groups, the *continuous* and *discrete* groups. A discrete group has elements that are *countable*. It does not have to be a finite group though. For example the group $Z_+$, the set of integers with '+' as the multiplication law, is countable but infinite, whereas the group $D_4$, which is the group of symmetry operations on a square, is finite and countable. So both groups are discrete, but only $D_4$ is finite.

In chapter 4 we will describe the more important continuous or *Lie groups* in more detail, but for now let me just summarize some general properties and theorems of discrete groups, which are needed (although in some slightly modified way) to later classify the Lie groups and their properties more clearly. These theorems will be stated mostly without proof, and can be looked up in any standard group theory textbook:

**Definition 3.1.** *Order $|G|$ of the finite group $G$*

The *order* $|G|$ of a finite group $G$ is simply the number of distinct elements belonging to that group:

$$\text{If } \; G = \{g_1, g_2, \ldots, g_n\}, \text{then } \; |G| = n \tag{3.1}$$

Hence the order of an infinite group is $\infty$.

**Definition 3.2.** *Order of an element $g \in G$*

In contrast to the order of a group, the order of an *element* $g \in G$, where $G$ is a group of order $n$ is the number $k \leqslant n$, such that $g$ composed $k$-times with itself equals identity:

$$\underbrace{g \circ g \circ \ldots \circ g}_{k-times} \equiv g^k = e \tag{3.2}$$

From now one I will leave the abstract composition $\circ$ out by writing $g_i \circ g_j$ simply as $g_i g_j$, suggesting ordinary multiplication. Keep in mind that the composition, although often simply *being* multiplication, in principle could be any operation defining a group.

If one multiplies the whole group $G$ with some element $g \in G$, the resulting set would be the group $G'$, composed of the same distinct elements as $G$, but possibly with its elements in a different order. However since the group does not care about the order of its elements, this new group $G'$ just equals the old group $G$, resulting in the next theorem:

**Theorem 3.3.** *Rearrangement Theorem*

$$gG \equiv G' = G \tag{3.3}$$

A direct consequence of this Theorem is that each row or column of the multiplication table[1] of a group $G$ contains each element of $G$ exactly once. So this multiplication table can be considered as a square $n \times n$ matrix, with each entry giving the composition or multiplication of two of the elements of $G$.

[1] A multiplication table is simply the square matrix of all possible multiplications within the group.

## 3.1 Different Substructures of a Group

A group in general can be divided into parts or substructures in numerous ways. A first obvious substructure would be a subgroup. But other partitions of a general finite group can also be defined, containing each different information. These partitions help us to classify and categorize the different operations of the group into partitions of operations that are 'alike'. This resemblance between the operations depends on the way one partitions the group. An elementary tool for this categorizing process is by defining equivalence relations between the elements of the group. The first substructure we consider is the subgroup:

**Definition 3.4.** *$H$ is a subgroup of $G$; $H \subset G$*
A *subgroup* $H$ of group $G$ is a subset $\{h_i\}$ of elements of order $|H| \leqslant |G|$ that, under the multiplication law of $G$, form a group by themselves, i.e. the set $\{h_i\}$ obey the group axioms of $G$.

**Definition 3.5.** *Proper Subgroup*
The subgroup is called *proper* if $H$ is neither equal to $G$ (since this is by definition a trivial subgroup of $G$), nor equal to $\{e\}$ (being the identity of $G$ and also being a trivial subgroup). Therefore for a proper subgroup $|H| < |G|$.

**Definition 3.6.** *Equivalence Relation; Equivalence Class*
An *equivalence relation* is an abstract mathematical relation $\sim$ between elements of a set $S$, which obeys specific rules:

1. (*reflexivity*) $a \sim a$
2. (*symmetry*) $a \sim b \quad \Leftrightarrow \quad b \sim a$
3. (*transitivity*) $a \sim b$ and $b \sim c \quad \Leftrightarrow \quad a \sim c$

Now any equivalence relation $\sim$ partitions a general set $S$, which could be a group, into *disjoint equivalence classes* $C_j$, each class containing elements equivalent to each other. They are disjoint since if two disjoint classes $C_1$ and $C_2$ would have at least one element in common, this would imply (because of the transitivity of the equivalence relation) that all elements of the two classes would be equivalent. Thus implying that the two classes are in fact one and the same class, contrary to hypothesis. So if we find a mathematical relation that turns out to be an equivalence relation, it automatically partitions the group in equivalence classes associated with that particular relation.
An important example of an equivalence relation is *conjugacy*:

**Definition 3.7.** *Conjugacy Relation; Conjugacy Classes $C_j$*
Two elements $g_i$ and $g_j$ are *conjugate* to one another if there exists a third element $g \in G$, such that $gg_ig^{-1} = g_j$.

One can check that this conjugacy relation does satisfy the criteria of being an equivalence relation, thereby partitioning the group into important disjoint equivalence classes, called *conjugacy classes*: $C_j$, $j = 1, 2, \ldots, k$, with $k$ being the number of distinct classes.

The number of elements in the $k$ disjoint classes of one group, is not the same for every class $C_j$. For example, the identity element is always in a class $\{e\}$ by itself, since it commutes with every other element. Moreover, for abelian groups, one can prove easily that each element forms a conjugacy class by itself, since the elements all commute.

It is easily seen that only the class that holds the identity $e \in G$, i.e. $\{\mathrm{e}\}$, forms a subgroup of $G$, albeit a trivial one. The rest of the classes, because of being distinct, don't contain the identity and therefore never form subgroups of $G$.

Besides conjugacy classes, there is another important substructure of a group $G$ which has to do with the subgroups $H \subset G$. It is called a *coset* of $G$:

**Definition 3.8.** *Cosets of a subgroup $H \subset G$; Invariant subgroup*

When multiplying a subgroup $H \subset G$ from the left by one element $g \in G$, we create a *leftcoset* of $H$: $gH = \{gh_1, gh_2, \dots, gh_{|H|}\}$. The set of all cosets of $H$, $\{g_iH\} = \{g_1H, g_2H, \dots\}$, is acquired by multiplying $H$ with all $g \in G$. The label $i$ now doesn't range over the whole group $G$, but merely labels the distinct cosets and thus in general $i \leqslant |G|$. It is called the index of $H$ in $G$, which will be become clear with Lagrange's Theorem. Analogously, by multiplying from the right, we create the *rightcosets*: $Hg_i$.

In contrast to the classes, which have varying number of elements, the cosets all have the same number of elements being of course the order $|H|$ of the subgroup. One can prove that this coset also defines an equivalence class, i.e. the elements in one coset are equivalent to one another. As with conjugacy, this equivalence relation partitions the group into disjoint equivalence classes, the cosets $g_iH$ themselves. This partition however is a completely different one then the partition acquired when conjugacy is considered as the equivalence relation, i.e. conjugacy classes and cosets both partition the group, but in different types of equivalence classes. Furthermore also some sort of rearrangement theorem for cosets holds: The coset $g_iH$ in this case is labeled by $g_i$. We could also label it by any other member of the coset $g_iH = \{g_ih_1, g_ih_2, \dots, g_ih_{|H|}\}$, since multiplying $H$ with any member of this coset would result in the same coset, because $H$ is a group and the rearrangement theorem holds for $H$: $g_iH = \{g_ih_1, g_ih_2, \dots, g_ih_{|H|}\}$, $H$ is a group, so $h_jH = H$, therefore $g_ih_jH = g_iH, \forall h_j \in H$.

Therefore we conclude that a coset can be labeled by any of its elements.

In general the left and right coset of a subgroup are not equal. If they are equal, the subgroup is called *normal* or *invariant*.

**Definition 3.9.** *Invariant Subgroup $H$*

A subgroup $H$ is called *invariant* or *normal* if the set of distinct left cosets equals the set of distinct right cosets or equivalently if the subgroup is invariant under a similarity transformation:

$$\forall g_i \in G, g_iH = Hg_i \Leftrightarrow g_iHg_i^{-1} = H \tag{3.4}$$

We see that this similarity transformation is just an example of a conjugacy operation and in fact therefore is an equivalence relation, so we conclude that $H$ being invariant or normal is the same as being *self-conjugate*. One can prove that an invariant subgroup always consists of a set of complete conjugacy classes. Two important theorems have to do with the cosets of a subgroup:

**Theorem 3.10.** *Coset Theorem*
Two cosets $g_iH$ and $g_jH$ of a subgroup $H$ either contain the same elements, i.e. are equal, or have no elements in common:

$$g_iH \cap g_jH = \begin{cases} \varnothing \\ g_iH = g_jH \end{cases} \tag{3.5}$$

This is of course a consequence of the fact that the cosets define an equivalence relation and therefore the cosets are disjoint. So, together with the Rearrangement Theorem, this states that every element $g \in G$ is in exactly one coset. It also follows that every distinct coset $g_iH$ of a subgroup $H$ has an equal amount of elements: $|H|$. Since every element is in exactly one coset, and the cosets all have an equal amount of elements, $|H|$, it must be so that $|H|$ is contained in $|G|$ an integer number of times. This is summarized in Lagrange's theorem:

**Theorem 3.11.** *Lagrange's Theorem*
The number of distinct cosets of a subgroup $H$ is a devisor $i$ of $|G|$, called the *index* of $H$ in $G$:

$$i = \frac{|G|}{|H|} \tag{3.6}$$

The index of $H$ in $G$ is equal to the number of distinct cosets of $H$. Later we will see that the so called *quotient group* $G/H$, were the cosets $g_iH$ themselves are considered as the 'elements' of the group, can be considered as a group with $i$ elements. Importantly, Lagrange's Theorem only gives the *possible* orders of $H$, given a group $G$. It does not tell if the possible devisor of $|G|$, which could represent an order of $H$, can be associated with a subgroup $H$ at all. It merely states that *if* $H$ is a subgroup of $G$, then its order must be a devisor of $|G|$.

## 3.2 Derived Groups

Starting with a general, finite group $G$, one can try to define derived structures, in a particular way associated with $G$, that by themselves obey the group axioms. Let's call them *derived groups*. In general when considering a group, it is impossible and even meaningless to classify that group as being derived from a more general group. But in this context, it is useful to see how other groups can be constructed starting from a general group. Two examples of derived groups are the *quotient group* and the *direct product group*.

**Definition 3.12.** *Quotient Group: $G/H$*
The set of cosets of an *invariant* subgroup $H \subset G$, $G$ finite, together form a group, called the *quotient group* $G/H = \{Hg_1, Hg_2, \ldots, Hg_i\}$ under the following multiplication law $*$ :

$$Hg_i * Hg_j \equiv H(g_i \circ g_j) = Hg_k \in G/H \tag{3.7}$$

Since $H$ is an invariant subgroup, it doesn't matter whether we take the left or right cosets as the elements, since they are equal. The order of the quotient group $|G/H|$

is equal to the index of $H$ in $G$, as mentioned earlier in Lagrange's theorem. So $|G/H| = |G|/|H| = i$

Looking at how the order of the quotient group is derived from the orders of $G$ and $H$, one might be tempted to write the group $G$ as some product of $G/H$ and $H$: $G \stackrel{?}{=} G/H \times H$. In some cases it is correct to write a group as the product of two other groups. However first we need to define what we mean by this product $\times$. A group that *can* be written as a product of two of its subgroups is called a *direct product group*. Since the direct product of groups is an essential ingredient of elementary particle physics, some properties of these groups and their irreps are shortly summarized here, though not derived.

**Definition 3.13.** *Direct Product:*$\times$

Let $H$ and $H'$ be two groups. Now define the set of ordered pairs $G = \{(h, h')|h \in H, h' \in H'\}$. Then the set $G$ is called the *direct product* of $H$ and $H'$, written as $G = H \times H'$.

**Definition 3.14.** *Direct product group:* $G = H \times H'$

The set $G$ forms a *direct product group* when the multiplication law $*$ of $G$ is defined to be pairwise: Let $g_1, g_2 \in G$, with the ordered pairs $g_i = (h_i, h'_i)$, then

$$g_1 * g_2 = (h_1, h'_1) * (h_2, h'_2) \equiv (\underset{\in H}{h_1 \circ h_2}, \underset{\in H'}{h'_1 \circ h'_2}) = (h_3, h'_3) = g_3 \tag{3.8}$$

From this definition of the multiplication law defining the direct product group, it follows immediately that $(e, e')$ is the identity of $G$ and every element $g = (h, h') \in G$ has inverse $g^{-1} = (h, h')^{-1} = (h^{-1}, h'^{-1})$.

The definition of the multiplication of ordered pairs also implies that the sets $\tilde{H} = \{(h, e')|h \in H\}$ and $\tilde{H}' = \{(e, h')|h' \in H'\}$ are invariant subgroups of $G = H \times H'$. $\tilde{H}$ and $\tilde{H}'$ are of course isomorphic to $H$ and $H'$ respectively.

That $\tilde{H}$ and $\tilde{H}'$ are truly invariant can easily be checked by regarding $g * \tilde{h} * g^{-1}$, with $\tilde{h} \in \tilde{H}$

$$\begin{aligned} g * \tilde{h}_i * g^{-1} &= (h, h') * (h_i, e') * (h, h')^{-1} = \\ &= (h, h') * (h_i, e') * (h^{-1}, h'^{-1}) = \\ &= (hh_ih^{-1}, h'e'h'^{-1}) = \\ &= (hh_ih^{-1}, e') \in \tilde{H} \subset G, \end{aligned}$$

since $H$ is a group.

So $\tilde{H}$ is an invariant subgroup and the same holds for $\tilde{H}'$ of course. The following important properties hold for $\tilde{H}$ and $\tilde{H}'$:

$$\tilde{H} \cup \tilde{H}' = G \tag{3.9}$$

$$\tilde{H} \cap \tilde{H}' = (e, e') \tag{3.10}$$

Of course we can try to turn the argument around and consider a group $(G, *)$ with two invariant subgroups $H \subset G$ and $H' \subset G$, which satisfy

$$H \cup H' = G \tag{3.11}$$

$$H \cap H' = e \tag{3.12}$$

Then $G$ is said to be decomposable into $H \times H'$, i.e. $G \simeq H \times H'$, meaning $G$ is isomorphic to a direct product group. The fact that $H$ and $H'$ are invariant and (3.12) implies that $\forall h \in H, h' \in H'$, $h * h' = h' * h$, i.e. $H$ and $H'$ commute, which in itself implies that every $g \in G$ can be uniquely written as some combination $h * h'$.

Notice the subtle difference in reasoning: It makes no sense to talk about $h_i * h'_j$ in the context of two different groups $H$ and $H'$, simply because the multiplication rule $*$ between elements of two different groups is not defined. However defining the groups $\tilde{H}$ and $\tilde{H}'$ of ordered pairs as the groups isomorphic to $H$ and $H'$ as defined above, allows us to define a product between two elements $\tilde{h}$ and $\tilde{h}'$ by the direct product. Taking a direct product of two groups $H$ and $H'$ using the ordered pairs thus creates a direct product group by definition. These groups can be the same or totally different groups, as long as the direct product of ordered pairs is used as the multiplication rule, it is guaranteed to create a direct product group. On the other hand starting from a given group $G$ with invariant subgroups $H$ and $H'$, the product between elements $h * h'$ of course *is* defined. However this product $*$ is just the regular multiplication law of $G$. If the subgroups furthermore obey (3.11) and (3.12), $G$ is isomorphic to a direct product group $H \times H'$.

Because $H$ and $H'$ are necessarily invariant subgroups of the group $G$ if $G$ is isomorphic to their direct product group $H \times H'$, for finite subgroups $H$ and $H'$ it becomes interesting to look at the quotient groups $G/H$ and $G/H'$. $G/H$ consists of all distinct cosets of $H$: $G/H = \{h'_1 H, h'_2 H, \ldots, h'_{|H'|} H\}$. The fact that the label of the distinct cosets of $H$ runs from 1 to $|H'|$ can be seen by realizing that since $G = H \times H'$, the index of $H$ in $G$ must be $|H'|$. For if we regard $H' = \{h'_1, h'_2, \ldots, h'_{|H'|}\}$, the cosets $h'_i H$ are all distinct. If they weren't, two of them would be equal: say $h'_i H = h'_j H$. This would mean that there are two elements $h_p, h_k \in H$, such that $h'_i h_p = h'_j h_k$. But because we previously concluded that every element $g \in G$ can uniquely be written as a product $h'h$, we see that this implies that $h'_i = h'_j$. But this gives a contradiction, since elements of a group $H'$ are supposed to be distinct. Therefore $h'_i H \neq h'_j H$, saying they are distinct. We conclude that $G/H$ follow the multiplication table of $H'$ in a one-to-one manner, i.e. $G/H$ and $H'$ are *isomorphic*, with notation $G/H \cong H'$. The same argument would lead to the conclusion that $G/H' \cong H$. So the quotient group $G/H$ of a direct product group $G = H \times H'$, is always isomorphic to the 'other' part of the direct product, $H'$.

If groups are isomorphic, meaning 'changing in the same way', we consider them to be equal. For isomorphic groups, the group structure, expressed in the multiplication table, is exactly equal, so the two groups are merely different representatives of the same group structure and can therefore be considered equal. This isomorphism and closely related homomorphism theory is essential when the all important *representations* of groups are to be considered, so a bit more on those ideas in the following section on representations.

## 3.3 Representations

Until now, we only considered abstract groups as sets of elements obeying certain rules and having a certain (sub)structure, fixed in the multiplication table. The elements

of the groups are in our case symmetry operations, acting on the objects of interest. The way these abstract group elements act depends on the objects they are acting on. We say that a particular group has different *representations*, depending on the object on which the symmetry operations are supposed to act. We demand that the structure of representations resembles the group structure of the symmetry group under consideration. To this extend we introduce the concepts of *homo-* and *isomorphisms*:

**Definition 3.15.** *Homomorphism;* $\phi\colon G \to G'$

A group *homomorphism* $\phi\colon G \to G'$ is a mapping from one group $(G, \circ)$ to another group $(G', \cdot)$, which preserves the group multiplications :

$$\phi(g_i \circ g_j) = \phi(g_i) \cdot \phi(g_j), \tag{3.13}$$

where $(\ \circ\ )$ is the multiplication rule for the group elements $g \in G$ and $(\ \cdot\ )$ is the multiplication rule for the homomorphisms $\phi(g) \in G'$.

**Definition 3.16.** *Image of a homomorphism;* $Im\ \phi$

The image of a homomorphism $\phi\colon G \to G'$ is the subset $Im\ \phi \subset G'$ that is reached by $\phi$ acting on all of $G$, i.e.:

$$Im\ \phi = \{g' \in G' | \phi(g) = g', g \in G\} \tag{3.14}$$

**Definition 3.17.** *Kernel of a homomorphism;* $Ker\ \phi$

The kernel of a homomorphism $\phi\colon G \to G'$ is the subset $Ker\ \phi \subset G$ which is taken to the identity element $e' \in G'$, i.e.:

$$Ker\ \phi = \{g \in G | \phi(g) = e'\} \tag{3.15}$$

**Definition 3.18.** *Isomorphisms*

A group *isomorphism* $\phi\colon G \to G'$ is simply a homomorphism from one group $(G, \circ)$ to another group $(G', \cdot)$, that is *bijective*, meaning one-to-one and onto. The onto requirement is simply guaranteed by restricting $G'$ to the image of $\phi$. For the homomorphism to be one-to-one, we use the well known result that the kernel of $\phi$ must consist of $e \in G$ only, i.e. $\phi\colon G \to G'$ is an isomorphism if the following requirements are met:

1. $\phi\colon G \to S$ is a homomorphism, with $G' \subset S$.
2. $Im\ \phi = G' \subset S$
3. $Ker\ \phi = \{e\} \subset G$

Now let us consider for example the symmetry group of the square, called $D_4$. An element $g \in D_4$ can be represented by a $2 \times 2$ matrix if we consider the square as a 2D object, i.e. an object spanned by 2D vectors $\vec{x} = (x_i, x_j)$. Of course one can also regard a square as a 2D object *embedded* in a 3D space, but now the third component of the vector doesn't play a role so we can put it to 0: $\vec{x} = (x_i, x_j, 0)$. In the 3D case, one dimension just doesn't 'participate' or better is left invariant under any $D_4$ operation. However, since acting on a 3D vector, $g \in D_4$ must therefore be represented by a $3 \times 3$ matrix. Both the $2 \times 2$ and the $3 \times 3$ matrices actually obey the group structure

rules of $D_4$. We say that they form *representations* of respectively 2D and 3D of $D_4$. Whether or not a given mathematical set of entities together with a multiplication law forms an actual representation of a group, depends on whether or not it *preserves the multiplication rule* of that group, i.e. whether the set is the image of a group homomorphism or not:

**Definition 3.19.** *Representation* $(D, V)$ *of Group* $G$*;* $D\colon G \to G'(V)$
A *representation* $(D, V)$ of a group $G$ is a homomorphic mapping $D\colon G \to G'(V)$, together with $V$ the *representation space*, preserving the group multiplications:

$$D(g_i \circ g_j)\mathbf{v} = [D(g_i) \cdot D(g_j)]\mathbf{v}, \tag{3.16}$$

with $\mathbf{v} \in V$ and $(\circ)$ the group multiplication rule and $(\cdot)$ the representation multiplication rule. Or shortly, leaving the representation space out

$$D(g_i \circ g_j) = D(g_i) \cdot D(g_j). \tag{3.17}$$

We conclude with the remark that for the representation group $G'$, we will exclusively use either the groups $GL(n, \mathbb{C})$, or $GL(V)$. $GL(n, \mathbb{C})$ is the so called *general linear matrix group* of all non-singular $n \times n$ matrices with complex entries, the multiplication rule in this group $(\cdot)$ is just matrix multiplication. On the other hand $GL(V)$ is the *general linear operator group* of linear operators acting on a vector space $V$ and in this case the multiplication rule $(\cdot)$ means 'composition'. In fact $GL(n, \mathbb{C})$ is just one realization of a $GL(V)$ group, because as we will see one can write a linear operator acting on a finite vector space explicitly as a matrix, by making an explicit choice of base. To take $GL(V)$ as the representation group avoids this arbitrary choice and takes the space of all base-*in*dependent linear operators as the space in which the representations live.

Before we look at representations acting on representation spaces as linear operators acting on vector spaces, let's get back to the quotient groups one more time and consider the quotient group consisting of the cosets of an important invariant subgroup of any finite group, the *kernel* of a representation (same as the kernel of a homomorphic mapping ): $Ker\ D$, with $D\colon G \to G'$.

**Definition 3.20.** *Kernel of Representation:* $Ker\ D \subset G$
The *kernel* of a representation $D\colon G \to G'$ is the subset of $G$ that maps onto the identity of $G'$: $e'$

$$Ker\ D \equiv \{g \in G | D(g) = e' \in G'\} \tag{3.18}$$

It's easy to prove that the kernel of a homomorphism indeed always forms an invariant subgroup. Since, for $g \in Ker\ D$ and $g_i$ any element of $G$, $D(g_i g g_i^{-1}) = D(g_i)D(g)D(g_i^{-1}) = D(g_i)e'D(g_i)^{-1} = e'$. Hence $g_i g g_i^{-1} \in Ker\ D$ and $Ker\ D$ is therefore an invariant subgroup of $G$. The fact that it *always* is a subgroup of $G$ restricts for finite groups the possible homomorphisms representing the group $G$, but we won't go into detail here. Two important possible kernels of a given representation $D(g)$ are:

1. $Ker\ D = G$ itself, then all elements of $G$ are mapped onto $e' \in G'$; this is the kernel of what is called the *trivial representation* and is by definition many-to-one, or *unfaithful*.

2. $Ker\ D = e \in G$. Unlike with the kernel of the unfaithful representation in case 1, this homomorphism is actually *bijective*, and therefore the mapping is an *isomorphism*. Thus the representation $D$ with this particular kernel is said to be *faithful*.

## 3.4 Representations acting on Vector spaces

Now we will discuss the mapping $D\colon G \to GL(V)$, which is a representation of $G$ onto the space of linear operators acting on a vector space $V$[2].

The functions $f \in L^2$ obey also the vector space axioms, but now the basis in which a function can be expressed has infinitely many 'vectors'. Consider for example the Fourier basis; if a function obeys certain periodicity conditions it can be expanded in the following basis: $f(x) = \sum_{n=1}^{\infty} f_n e^{2\pi i n x}$. Now the exponents $e^{2\pi i n x}$ can be regarded as the basis 'vectors' and there are infinitely many of them. The 'components' of the vector $f(x)$ are thus $f_n$ in this basis, called the Fourier components.

The demand that $GL(V)$ is *linear* is translated into the behavior of its components $T \in GL(V)$: $D\colon V \to V$ is linear iff

$$T(a\mathbf{v} + b\mathbf{w}) = aT(\mathbf{v}) + bT(\mathbf{w}) \tag{3.19}$$

with $a$ and $b$ two complex scalars and $v, w \in V$. Now a linear transformation $T \in GL(V)$, representing a symmetry operation $g \in G$, acting on a vector $\mathbf{v} = v^i e_{(i)}$ in a given basis $\{e_{(i)}\}$ can be expressed explicitly in matrix formulation using the transformation of each individual basis vector $e_{(i)}$, which reads

$$e'_{(j)} = Te_{(j)} = [D(g)]_{ij} e_{(i)} \equiv D_{ij} e_{(i)} \tag{3.20}$$

This is explicitly *not* an ordinary vector multiplication, instead it states that any transformed vector $e'_{(j)}$ can be written as a linear combination of basis vectors $e_{(i)}$. Also note that since $i$ is merely a dummy index, the seemingly distinct bases $\{e_{(i)}\}$ and $\{e_{(j)}\}$ are the same. By making use of the linearity of $T$, the transformation of a general vector $\mathbf{v}$ is then given by:

$$T\mathbf{v} = T(v^j e_{(j)}) = v^j T e_{(j)} = v^j D_{ij} e_{(i)} = [D_{ij} v^j] e_{(i)} \tag{3.21}$$

We think of this transformation $T$ as acting on the vector $\mathbf{v}$, therefore in principle leaving the basis unchanged, however the only information we have is how $T$ acts on the basis vectors $e_{(j)}$ expressed in terms of $e_{(i)}$ by equation (3.20). We therefore use this equation and consider the transformed vector $T\mathbf{v}$ as another vector $\mathbf{u} = u^i e_{(i)}$. We can write its components $u^i$ in the basis $\{e_{(i)}\}$ as:

$$u^i = D_{ij} v^j \tag{3.22}$$

[2]See Appendix A

In this case it *is* an ordinary vector multiplication, when regarding the coordinates $v^j$ as a column vector. So the vector components transform with $D_{ij}(g)$, depending on the chosen basis $\{e_{(i)}\}$ and on the transformation $g \in G$. Note that we did not make a basis transformation, we merely transformed the vector $\mathbf{v}$ into vector $\mathbf{u}$ and then expressed the components $u^i$ in the *same* basis $\{e_{(i)}\}$.

If we do decide to make a basis transformation, say $\{e_{(i)}\} \overset{S^{-1}}{\to} \{f_{(k)}\}$, with the linear superposition

$$f_{(k)} = S^{-1}_{ik} e_{(i)} \tag{3.23}$$

$D$ would take a different form in this new basis $\{f_{(k)}\}$, call it $D'$. $S$ is required to be invertible. In this new basis, the *same* linear transformation $T$ acting on one of the basis vectors and representing some element $g \in G$, would take the following form:

$$T f_{(l)} = D'_{kl} f_{(k)}, \tag{3.24}$$

If we furthermore write our vector components in this new basis as $\mathbf{v} = v'^k f_{(k)}$ and $\mathbf{u} = u'^k f_{(k)}$ then the transformation $T$, taking the vector $\mathbf{v}$ to $\mathbf{u}$, will analogously to equation (3.22) result in component transformation

$$u'^k = D'_{kl} v'^l \tag{3.25}$$

If we now express the old basis $\{e_{(i)}\}$ as a linear combination of the new basis vectors $\{f_{(k)}\}$, i.e. we invert equation (3.23)

$$e_{(i)} = S_{ki} f_{(k)} \tag{3.26}$$

then we are able to express our vectos $\mathbf{v}$ and $\mathbf{u}$ directly in terms of this basis as

$$\mathbf{v} = v^i e_{(i)} = v^i S_{ki} f_{(k)} \tag{3.27}$$

thus giving $v'^k = v^i S_{ki}$ and similarly $u'^k = u^i S_{ki}$. The component $u'^k$ in the basis $\{f_{(k)}\}$ of the transformed vector $\mathbf{u} = T\mathbf{v}$ are then given by

$$u'^k = S_{ki} u^i = S_{ki} \left( D_{ij} v^j \right) = \left( S_{ki} D_{ij} S^{-1}_{jl} \right) v'^l \tag{3.28}$$

Comparing this equation with equation (3.25), we conclude that the components of $D'$ and $D$ are related by

$$D'_{kl} = S_{ki} D_{ij} S^{-1}_{jl} \tag{3.29}$$

i.e. in matrix notation

$$D' = S D S^{-1} \tag{3.30}$$

and therefore we can conclude that any two matrix representations in different bases, are in fact equivalent, or

$$D(g) \sim D'(g). \tag{3.31}$$

So the set of *all* linear transformations $T = \{T(g)\}$ representing all $g \in G$, corresponds to a whole *conjugacy class* of matrix representations $D'(g) = SD(g)S^{-1}$. Therefore it is more convenient to choose the basis independent target space $GL(V)$ instead of the basis dependent $GL(n, \mathbb{C})$. Now for this representation to be a group homomorphism, the linear operators must preserve the group multiplication rules:

$$T(g_i \circ g_j) = T(g_i) \cdot T(g_j). \tag{3.32}$$

A vector space $V$ for which this holds for the whole group $G$ is also known as a *$G$-module*.

## 3.5 Equivalent Representations and Character

Following the previous discussion about the difference between the matrix and linear operator representations of a group, it seems reasonable to regard the matrix representations which are related by a similarity transformation (3.24) as equivalent, since they merely represent the same transformation in a different basis. We therefore regard representations of a group as distinct *upto* conjugacy:

**Definition 3.21.** *Equivalent Matrix Representations*
Two matrix representations $D$ and $D'$ are *equivalent* if $\forall g \in G$ *all* matrices $D(g)$ and $D'(g)$ are related by the same similarity transformation:

$$D(g) = SD'(g)S^{-1}, \tag{3.33}$$

for some transformation matrix $S$. In terms of linear operator representations, this just means that both $D(g)$ and $D'(g)$ correspond to the same linear transformation $T(g)$, representing one and the same group element $g \in G$.

In order to distinguish distinct, i.e. non-equivalent, representations, we will introduce a mathematical construct called the *character* of a representation, which is a so called *class* function, depending only on the conjugacy classes of a representation, not on the particular matrix form itself:

**Definition 3.22.** *Character of a representation:* $\mathcal{X}$
The *character* $\mathcal{X}_j$ *of one conjugacy class* $C_j$ is defined as the trace of one representation matrix $D_j(g)$ in that class, i.e. $g \in C_j$:

$$\mathcal{X}_j = Tr\ D_j(g) \tag{3.34}$$

The character $\mathcal{X}$ of the *whole representation* $D$ is subsequently the set $\{\mathcal{X}_j(g)|g \in G\}$

It is easy to see that $\mathcal{X}_j$ defined in this way is a function of class $j$: Let's say $g_1$ and $g_2$ are in the same conjugation class $C_j$, i.e. there is an element $g \in G$ such that $g_2 = gg_1g^{-1}$. Any given representation must respect these multiplication rules. So for $D_j(g_1 \in C_j)$, another member of the matrix representation $D_j(g_2 \in C_j)$ from the same class $j$, must be related to $D_j(g_1)$ by a matrix $D(g)$ such that $D_j(g_2) = D(g)D_j(g_1)D^{-1}(g)$. Therefore the character $\mathcal{X}_j$ is equal for both matrices, and in fact the whole class $C_j$, and is given by

$$\mathcal{X}_j = Tr\ D_j(g_2) = Tr\ D(g)D_j(g_1)D(g)^{-1} = Tr\ D_j(g_1)D(g)^{-1}D(g) = Tr\ D_j(g \in C_j)\ , \tag{3.35}$$

because of the cyclic properties of the trace. So for the whole conjugacy class, we have the same character. One representation $D$ in general contains several conjugacy classes, $C_1, C_2, \ldots, C_k$, each with a distinct character $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_k$. We saw that two *elements* $D(g_1)$ and $D(g_2)$ of a matrix representation have the same character $\mathcal{X}_j$ if $g_1, g_2 \in C_j$, they are conjugated within that representation $D$. But the equivalence of two whole representations $D$ and $D'$ (See def.(3.21)) can be rephrased in terms of the character of the representations: Two *representations* $D$ and $D'$ are equivalent when the characters for *all corresponding classes* are equal $\{\mathcal{X}_j(g)\} = \{\mathcal{X}'_j(g)\}$.

There are two easily mixed up notions here: First of all, the cyclic property of the trace has as a consequence that the character of a whole representation is invariant under a change of basis by a similarity transformation S. The character of the whole rep $D$ is $\mathcal{X} = \{\mathcal{X}_j\}$, but this set and hence $\mathcal{X}$ is invariant since all its elements $\mathcal{X}_j$, i.e. every character of conjugacy class $C_j$, is itself invariant under a basis transformation. Secondly the cyclic property of the trace is also used to demonstrate the fact that the character for all representations $D_j(g \in C_j)$ of elements belonging to the same conjugacy class $C_j$ are identical. For example, the transformation matrix $S$ from (3.33), regarding the whole representation $D$, has the particular form of $D(g)$, the matrix representation of $g$, when regarding the character $\mathcal{X}_j$ of only one conjugacy class. This is because $g$ is exactly that element that conjugates $g_1$ and $g_2$.

We are in general only interested in the conjugacy classes $C_j$ and their characters $\mathcal{X}_j$ of a particular representation $D$, not the individual elements $D_j(g)$ in each class. Particularly the *number of classes* $k$ plays an important role in the search for the fundamental building blocks for the representations, called irreps, as we will see.

## 3.6 Reducibility; Irreps and unitarity

It is easy to see that once one has some matrix representation $D$ of a certain group $G$, one can construct a second representation by a so called direct sum $\oplus$ operation: $D' = D \oplus D \equiv \left(\begin{smallmatrix} D & \varnothing \\ \varnothing & D \end{smallmatrix}\right)$. This new matrix set $D'$ is indeed a representation if $D$ is one, since then $\forall g_1, g_2 \in G$

$$\begin{aligned} D'(g_1 g_2) &= \begin{pmatrix} D(g_1 g_2) & \varnothing \\ \varnothing & D(g_1 g_2) \end{pmatrix} \\ &= \begin{pmatrix} D(g_1)D(g_2) & \varnothing \\ \varnothing & D(g_1)D(g_2) \end{pmatrix} \\ &= \begin{pmatrix} D(g_1) & \varnothing \\ \varnothing & D(g_1) \end{pmatrix} \begin{pmatrix} D(g_2) & \varnothing \\ \varnothing & D(g_2) \end{pmatrix} \\ &= D'(g_1)D'(g_2) \end{aligned}$$

However this representation $D'$ is of course not a really 'new' one, since it is build from already known representations $D$. We can argue the other way around and say that if some representation $D$ can be written in a way that it is a direct sum of other, more fundamental, representations $D^{(\mu)}$, then we call $D$ *reducible*. If $D^{(\mu)}$ can't be written in this so called block-diagonal form, then it is are called an *irreducible* representation or *irrep* in short.

**Definition 3.23.** *Reducibilty (of matrix representations); Irreps ∼ non-decomposable matrix set*

If all matrices $D(g)$ of representation $D$ can be *simultaneously* diagonalized by the same basis transformation $S$, then the representation $D$ is called (completely) *reducible*, if not *irreducible*, i.e. if there exists a basis transformation $S$, such that $\forall g \in G$,

$$SD(g)S^{-1} = \begin{pmatrix} D^{(1)}(g) & \varnothing \\ \varnothing & D^{(2)}(g) \end{pmatrix}, \tag{3.36}$$

then $D$ is (completely) reducible. If $D^{(\mu)}$ cannot be reduced any further, they are called the *irreducible representations* or short *irreps* of $G$.

Just like a finite group has some finite number of classes $k$, as was mentioned in section 3.1, a finite group also has a finite number $m$ of irreps. As we will see in the section on the Great Orthogonality Theorem, these two numbers are related and help to identify all irreps of a certain finite group.

We have actually used *complete* reducibility (also called *decomposable*) here, which means that the upper right entry of the matrix contains only zeros. Strictly speaking, ordinary (non-complete) reducibility means that the upper right entries of the transformed matrices can be non-zero. But there is a theorem, called Maschke's theorem, which states that every reducible representation of a *finite* group is in fact completely reducible. And since we, up to now, were talking about finite groups, complete reducibilty will be assumed. Maschke's theorem for linear operator representations is elaborated in the following discussion.

Reducibility is, besides one of the central themes of general group theory, also *the* central theme of this thesis. This is so because the irreps form the building blocks of any representation of a group and we will later find out that they can be labeled in specific ways, making a fundamental characterization of a symmetry group possible. Furthermore since particles are to be associated with vectors in Hilbert space associated , on which the representations of the symmetry operations act, as was mentioned in section 3.4, the representations need to be unitary. They transform under some representation of a symmetry group of the theory, in the flat spacetime case this is the Poincaré group as we saw. In the chapter on Lie groups 4 we discuss the fact that the fundamental intrinsic properties of those particles, such as mass and spin, can be associated with the labels of the irreps of the Poincaré group. Once we understand this mechanism in flat spacetime, i.e. represented by the Poincaré symmetry group, we will apply the same mechanism in *curved* spacetime, were the symmetry operations form the de Sitter group. In the view of this it is more appropriate to consider the linear operator representations $GL(V)$, acting on a vector space $V$, as the natural representation of a group. In the language of linear operator representations, reducibility is intuitively translated into the existence of *subspaces*:

**Definition 3.24.** *Reducibility (of linear operator representations); Irreps ∼ invariant subspaces*

Say $T\colon G \to GL(V)$ is a representation of $G$. If there exists a subspace $U \subset V$, such that $\forall g \in G$ and $\forall \mathbf{u} \in U$

$$T(g)\mathbf{u} \in U, \tag{3.37}$$

meaning $U$ is an *invariant subspace* of $V$ under all operations of $G$, then $T$ is said to be *reducible*.

This definition of course is parallel to definition 3.23 for matrix representations, since these irrep matrices act only on subspaces that are closed under these matrix multiplications.

In order for $T$ to be *completely* reducible we rely on two theorems: The first theorem states that every representation of a finite group is equivalent to a unitary representation, the second theorem states that every unitary representation is decomposable. Together they are called Maschke's theorem for linear operator representations:

**Theorem 3.25.** *Maschke's Theorem*
Every linear operator representation of a *finite* group is *decomposable*.

The same can of course be proven for matrix representations. To prove Maschke's theorem, we need to introduce a few more important concepts: the *group invariant inner product*, *orthonormal basis* and *unitarity transformations*[3]:
Having defined an inner product we can now define what we mean by a *unitary transformation*. It has everything to do with the unitarity of matrices, as in the requirement that the inverse of a matrix $U$ is equal to its hermitian conjugate, the complex transposed: $U^{-1} = U^\dagger = (U^*)^T$ and therefore $UU^\dagger = U^\dagger U = I$. For linear operators $T$ the unitarity requirement is defined as in the following:

**Definition 3.26.** *Unitary Transformation*
Given a scalar product ( , ) defined on a normalized vector space $V$, a linear transformation $T$ is said to be *unitary* iff $\forall\ \mathbf{u}, \mathbf{v} \in V$:

$$(T\mathbf{u}, T\mathbf{v}) = (\mathbf{u}, \mathbf{v}) \tag{3.38}$$

With unitary transformations and the scalar product properly defined, one can proof that any reducible *unitary* linear operator $T$, which requires the existence of an invariant subspace $U \subset V$ by def. 3.24, is in fact *completely reducible*. If we now can choose a scalar product such that a linear operator always is unitary, we can state that every linear operator can be made unitary and therefore is completely reducible, proving Maschke's Theorem.

**Definition 3.27.** *Group invariant inner product;* { , }
The inner product { , } which per definition is group invariant, i.e. implies that $T$ is unitary, is given by:

$$\{\mathbf{v}, \mathbf{v}'\} \equiv \frac{1}{|G|} \sum_{g \in G} \big(T(g)\mathbf{v}, T(g)\mathbf{v}'\big) \tag{3.39}$$

Where ( , ) is the original inner product discussed in Appendix A. It is easy to prove that a a symmetry operations represented by a linear transformation $T$ acting on the vectors $\mathbf{v}, \mathbf{v}'$ leaves this inner product invariant, i.e.

$$\{T(g')\mathbf{v}, T(g')\mathbf{v}'\} = \frac{1}{|G|} \sum_{g \in G} \big(T(g')T(g)\mathbf{v}, T(g')T(g)\mathbf{v}'\big) = \frac{1}{|G|} \sum_{g'' \in G} \big(T(g'')\mathbf{v}, T(g'')\mathbf{v}'\big) = \{\mathbf{v}, \mathbf{v}'\}$$

So for a vector space with this inner product defined on it, every linear operator is unitary. The idea now is that by Gram-Schmidt orthogonalisation, we can always choose our finite dimensional basis to be orthonormal. If we furthermore have a reducible unitary operator $T$, this means that there exists a submodule $U$ that is closed under all symmetry operations $T(g)$. But to be *completely* reducible, this submodule must be the *orthogonal complement* of the rest of the vectorspace, $W = V \backslash U$, i.e. $W = \{\mathbf{w} \in V \mid (\mathbf{w}, \mathbf{u}) = 0\ \forall \mathbf{u} \in U\}$. This ensures that both $U$ and $W$ are closed under the symmetry operations of the group, and therefore that $T$ is *completely* reducible.

[3]For the definitions of inner product etc see Appendix A again.

Thus by choosing this orthonormal basis, we ensure that every vector space $V$, on which a unitary transformation $T$ acts, is completely reducible. We just have proven in (3.39) that such an inner product can always be constructed and hence that every unitary transformation is completely reducible. Together with the fact that every operator can be made unitary with the right scalar product, proves that every representation of a finite group is completely reducible.

Therefore we conclude that for finite discrete groups the vector space on which its linear operator representations act, can be divided into subspaces that transform among themselves. For infinite groups this does not hold in general. These subspace form the fundamental building blocks of any space on which linear operator representations act. Analogously to the matrix representations, these are also called the *irreps* of that particular group. We note that in this case we call the invariant subspace $U \subset V$ on which the 'actual' representation $T$ acts, an irrep of $G$. Whereas with the matrix representations, we called the representations $D^{(i)}$ themselves the irreps of the group.

## 3.7 Great Orthogonality Theorem; Decomposition

To find all irreps of finite groups we consider the two central lemma's in group theory: Schur's Lemma's. Without proof (see [9] or any other group theory book) we state them here in terms of linear operator representations $T$:

**Theorem 3.28.** *Schur 1*

Let $T(g)$ be irreducible, then if there exists an operator $\hat{Q}$, that commutes with every $T(g)$, $\forall g \in G$: $\hat{Q}T(g) = T(g)\hat{Q}$, then $\hat{Q} = \lambda \hat{I}$, where $\hat{I}$ is the identity operator.

This lemma is important for defining the *Casimir* operators of a group (see section 4.6), which are the tools to label any irreps and hence characterize the elementary particles associated with it.

**Theorem 3.29.** *Schur 2*

Let $T(g)$ and $T'(g)$ be two non-equivalent irreps acting on $U$ and $U'$ respectively. Let an operator $\hat{Q}$ map $U$ onto $U'$: $\hat{Q}: U \to U'$. If $\forall g \in G$, $\hat{Q}T(g) = T'(g)\hat{Q}$, then $\hat{Q} = \hat{O}$, with $\hat{O}$ the null-operator that maps $U$ onto $\mathbf{0'} \in U'$: $\hat{O}\mathbf{u} = \mathbf{0'}$

### 3.7.1 Great Orthogonality Theorem (GOT)

Using Schur 1 and 2, translated into matrix form, we state without proof the following relation for *inequivalent irreducibel* matrix representations $D^{(\mu)}, D^{(\nu)}$:

**Theorem 3.30.** *Great Orthogonality Theorem for matrix representations*

$$\sum_g D^{(\mu)}_{ir}(g) D^{(\nu)}_{sj}(g^{-1}) = \frac{|G|}{n_\mu} \delta^{\mu\nu} \delta_{ij} \delta_{rs}, \tag{3.40}$$

where $n_\mu$ is the dimensionality of $D^{(\mu)}$. We saw that we can take any representation of a finite group to be unitary, so then (3.40) becomes:

$$\sum_g D^{(\mu)}_{ir}(g) D^{(\nu)*}_{js}(g) = \frac{|G|}{n_\mu} \delta^{\mu\nu} \delta_{ij} \delta_{rs} \tag{3.41}$$

This, of course, is just an inner product of the two vectors

$$D^{(\mu)}_{ir} = \{D^{(\mu)}_{ir}(g_1), D^{(\mu)}_{ir}(g_2), \ldots, D^{(\mu)}_{ir}(g_{|G|})\}$$

and

$$D^{(\nu)}_{js} = \{D^{(\nu)}_{js}(g_1), D^{(\nu)}_{js}(g_2), \ldots, D^{(\nu)}_{js}(g_{|G|})\}.$$

$D^{(\mu)}_{ir}$ has $|G|$ components, i.e. is a $|G|$ dimensional vector. Furthermore the indices $i, r$ both run from 1 to $n_\mu$, giving $n^2_\mu$ possible linear independent vectors $D^{(\mu)}_{ir}$ for a given irrep $\mu$. The same of course holds for $D^{(\nu)}_{js}$. Equation (3.41) therefore expresses the orthogonality of two inequivalent irreps $D^{(\mu)}$ and $D^{(\nu)}$, hence GOT. However the total number of linear independent vectors in a $|G|$ dimensional vector space can never exceed $|G|$. Resulting in the inequality

$$\sum_\mu n^2_\mu \leqslant |G|, \tag{3.42}$$

limiting thereby $M$ the total number of irreps of a finite group. This inequality actually turns out to be an *equality*, as can be derived by expressing the GOT in terms of characters as we will do below and regarding the so called regular representation, which we will not treat, for its main purpose is to prove the actual equality of equation (3.42). We will just state this as a given fact:

$$\boxed{\sum_\mu n^2_\mu = |G|}. \tag{3.43}$$

Rewriting the GOT in terms of characters is of course done by taking the traces of (3.40):

$$\begin{aligned} Tr(\sum_g D^{(\mu)}_{ir}(g) D^{(\nu)*}_{js}(g)) &= \sum_g \mathcal{X}^{(\mu)}(g)\mathcal{X}^{(\nu)*}(g) \\ &= \frac{|G|}{n_\mu}\delta^{\mu\nu}\delta_{ij}\delta_{rs}\delta_{ir}\delta_{sj} \\ &= \frac{|G|}{n_\mu}\delta^{\mu\nu}\underbrace{\delta_{ij}\delta_{ij}}_{n_\mu} \\ &= |G|\delta^{\mu\nu} \end{aligned} \tag{3.44}$$

Rewriting we have:

$$\frac{1}{|G|}\sum_g \mathcal{X}^{(\mu)}(g)\mathcal{X}^{(\nu)*}(g) = \delta^{\mu\nu} \tag{3.45}$$

Which again can be read as a normalized scalar product of the character vectors: $\langle \mathcal{X}^{(\mu)}, \mathcal{X}^{(\nu)} \rangle = \delta^{\mu\nu}$ Since $\mathcal{X}$ is a class function, we can replace the sum over the group elements by a sum over all classes $C_i$, multiplied by the elements per class $\alpha_i$, where $i = 1, 2, \ldots, k$, with $k$ the total number of equivalence classes:

$$\boxed{\frac{1}{|G|}\sum_{i}^{k}\alpha_i\mathcal{X}_i^{(\mu)}(g)\mathcal{X}_i^{(\nu)*}(g)=\delta^{\mu\nu}} \tag{3.46}$$

With arguments analogous to those used in the derivation of equation (3.42) on the dimensionalities of the vector spaces involved, it can likewise be proven that the number of irreps $M$ is limited from above by the number of conjugacy classes $k$:

$$M \leqslant k \tag{3.47}$$

which again, turns out to be the all important equality

$$\boxed{M=k}, \tag{3.48}$$

stating that the number of irreps equals the number of conjugacy classes for finite groups. This equality holds because equation (3.46) can also be rewritten such that one sums over the irreps, not over the classes, resulting in $k \leqslant M$, hence implying equation (3.48). This equality is of fundamental importance, since it restricts the number of irreps $M$ to equal the number of classes $k$ for a finite group, thereby fixing the fundamental structure all possible representations of finite group. Furthermore we can use equation (3.48) to distinguish the separate building blocks, i.e. irreps, of any given reducible representation. This process is called *decomposition* and makes use of the great orthogonality theorem for characters (3.46).

### 3.7.2 Decomposition

The idea of decomposition is that any given reducible matrix representation $D$ can be written in block diagonal form, with its irreps $D^{(\mu)}$ on the diagonal. This block diagonal form is written as a linear direct sum combination of irreps $\forall g \in G$:

$$D(g)=\bigoplus_{\mu} a_\mu D^{(\mu)}(g), \tag{3.49}$$

with $a_\mu$ the frequency coefficient with which a certain irrep occurs in the linear combination. The key problem is to find these numbers $a_\mu$ for all $M=k$ irreps of a given group. The solution comes from taking the trace of equation (3.49) and using the orthogonality of the characters. In taking the trace we focus on a general group element $g$:

$$\mathcal{X}(g)=\sum_{\mu} a_\mu \mathcal{X}^{(\mu)}(g) \tag{3.50}$$

Multiplying equation (3.50) from the left by $\mathcal{X}^{(\nu)}$ and summing over $\nu$ and making use of equation (3.46) we can find the coefficients $a_\nu$:

$$\boxed{a_\nu=\frac{1}{|G|}\sum_{g}\mathcal{X}(g)\mathcal{X}^{*(\nu)}(g)=\langle\mathcal{X},\mathcal{X}^{(\nu)}\rangle} \tag{3.51}$$

To conclude we see that these coefficients tell us *exactly* how a general representation $D$ decomposes into irreps $D^{(\mu)}$ of the group by making use of equation (3.49). The

equations (3.43), (3.46), (3.48) and (3.51) together determine all decompositions of any given reducible representation into irreps of the group.

We are also interested in the vector spaces *on which a certain reducible representation acts irreducibly.* This means that we are looking for that basis on which $D(g)$ or $T(g)$ acts in such a way that the submodules are closed $\forall g \in G$. This is equivalent to finding the bases for the different irreps, when the decomposition of the particular representation is known.

## 3.8 Tensor product representation; Clebsch Gordan decomposition

Consider a system with symmetry group $G$, with irreducible representations $D^{\mu}(G)$ and $D^{\nu}(G)$, with dimensions $d_{\mu}$ and $d_{\nu}$ respectively. One can then construct a direct product representation $D^{\mu}(G) \times D^{\nu}(G)$ of $G$, with dimension $d_{\mu} \cdot d_{\nu}$ defined as follows:

**Definition 3.31.** *Direct Product Representation*

$$D^{(\mu)} \otimes D^{(\nu)} \equiv D^{(\mu \times \nu)} \tag{3.52}$$

In terms of matrix entries this direct product is written as:

$$[D^{(\mu)}]^{a}{}_{b}[D^{(\nu)}]^{c}{}_{d} \equiv [D^{(\mu \times \nu)}]^{ac}{}_{bd} = [D^{(\mu \times \nu)}]^{A}{}_{B}, \tag{3.53}$$

with $A = (ac)$ the first index and $B = (bd)$ the second index of the matrix $D^{(\mu \times \nu)}$. This implies that the direct product representation $D^{(\mu)} \otimes D^{(\nu)}$ has $\left[D^{(\mu \times \nu)}\right]^{ac}{}_{bd}$ as a matrix, for which the entries are given by (3.53). That the direct product of two irreps forms a rep itself can easily be seen as follows:

$$\begin{aligned} D^{(\mu \times \nu)}(g_1) D^{(\mu \times \nu)}(g_2) &= [D^{(\mu)}(g_1) \otimes D^{(\nu)}(g_1)][D^{(\mu)}(g_2) \otimes D^{(\nu)}(g_2)] = \\ &= D^{(\mu)}(g_1) D^{(\mu)}(g_2) \otimes D^{(\nu)}(g_1) D^{(\nu)}(g_2) = \\ &= D^{(\mu)}(g_1 g_2) \otimes D^{(\nu)}(g_1 g_2) = \\ &= D^{(\mu \times \nu)}(g_1 g_2) \quad \square \end{aligned} \tag{3.54}$$

where in the second line we made use of the pairwise character of the direct product discussed in section (3.2). Note that by convention a direct product of groups is denoted by $\times$ and a direct product of representations by $\otimes$. The meaning however is the same: The direct product of two sets $S_1$ and $S_2$, either groups or representations in our case, is just the set of ordered pairs $\{(s_1, s_2) | s_1 \in S_1, s_2 \in S_2\}$ together with a multiplication rule on how to multiply two pairs (See section 3.2).

If the system under consideration is in such a situation that the symmetry group of the system can be regarded as $G^{(1)} \times G^{(2)}$, with $G^{(1)} = G^{(2)}$, this means that the system is invariant under two independent symmetry operations $g_1 \in G^{(1)}$ and $g_2 \in G^{(2)}$. This direct product representation $D^{(\mu)} \otimes D^{(\nu)}$ is an irrep of $G \times G$ iff $D^{(\mu)}$ and $D^{(\nu)}$ are irreps of $G$. However sometimes (See [10]) $G \times G$ cannot be regarded as the symmetry group of the system, and can therefore be decomposed into the so called *Clebsch-Gordan series*:

$$D^{(\mu)} \otimes D^{(\nu)} = \bigotimes_{\rho} a_\rho D^{(\rho)} \tag{3.55}$$

Analogously to the previous discussion, the coefficients in this particular decomposition, called *Clebsch-Gordan coefficients*, are derived by taking traces and inner products. First of all the characters of a direct product representation have an important property:

$$\begin{aligned} \mathcal{X}^{(\mu\times\nu)}(g) &= [D^{(\mu\times\nu)}(g)]^{A}{}_{A} = \\ &= [D^{(\mu\times\nu)}(g)]^{ac}{}_{ac} = \\ &= [D^{(\mu)}(g)]^{a}{}_{a}[D^{(\nu)}(g)]^{c}{}_{c} = \\ &= \mathcal{X}^{(\mu)}(g)\mathcal{X}^{(\nu)}(g) \end{aligned} \tag{3.56}$$

With this property the Clebsch-Gordan coefficients can be easily calculated:

$$a_\rho = \langle \mathcal{X}^{(\rho)}, \mathcal{X}^{(\mu\times\nu)} \rangle = \langle \mathcal{X}^{(\rho)}, \mathcal{X}^{(\mu)}\mathcal{X}^{(\nu)} \rangle \tag{3.57}$$

# 4 Lie groups

Thus far we have only considered groups with a countable number of elements, i.e. discrete groups. These sets may contain a finite or infinite number of elements, the fact that they are countable makes the group discrete. However sometimes a system is invariant under symmetry operations that can be infinitely close to one another. Take for example a system which is rotationally invariant, but this time for *any* rotation. The group of rotations in 3D that leave the system invariant, called $SO(3)$, is not finite any longer, since of course there are infinitely many angles over which one can rotate the system. The elements $R(\phi, \psi, \rho) \in SO(3)$ depend continuously on three parameters $\phi, \psi, \rho$, namely the three angles necessary to specify the 3D rotation. A group with elements that continuously depend on a parameter, in this case 3 parameters, and that therefore are infinitely close to each other, is called a *continuous* group of *Lie* group. As said, the elements are uncountable for Lie groups, therefore some of the properties of discrete groups don't translate one-to-one into the Lie group domain.

## 4.1 General properties Lie groups

The order $|G|$ of a discrete group is a useless notion when regarding Lie groups, since the elements are uncountable and therefore would always be $\infty$. Instead the number of parameters, which were absent in discrete groups, now forms a good property with which we can identify the particular group. We call this number the *dimension* of the continuous group.

**Definition 4.1.** *Dimension of a Lie Group; d*
The *dimension* $d$ of a Lie Group $G$ is the smallest number of independent parameters necessary to define an element $g(\phi_1, \phi_2, \ldots, \phi_d) \in G$.

In the case of rotations in 3D space, $SO(3)$, the dimension of the *group* is 3, because we need 3 parameters to specify a rotation in 3D space. However the fact that the dimension of the group and target space are equal is an accidental equality, since for rotations in 2D space, we only need 1 parameter, and hence the dimension of $SO(2)$ is 1. Later we will see that the number of *generators* actually defines the dimension of a Lie group, and therefore fixes the number of parameters. In contrast, with the case of discrete groups we also talked about dimensions $n_\mu$, but in the context of discrete groups these were the dimensions of the *representations*. The dimension of a discrete *group* is not defined. The topological way the elements of a Lie group depend on the parameters classifies the type of Lie group in several ways, which are important to deduce general properties of certain types of Lie groups. We will give some important definitions now.

We can first of all regard the group as a topology in $d$-dimensional space on which we can define curves. To that extend we imagine that the $d$ parameters $\phi_i$ of $g(\phi_i)$ continuously vary by another parameter $t \in [0, 1]$, i.e. $\phi_i = \phi_i(t)$, such that $\phi_i(0)$ and $\phi_i(1)$ are the parameter values of the begin en end points of the curve: $g(\phi_i(0))$ and $g(\phi_i(1))$ respectively. Then a curve $g(\phi(t))$ is traced out in the space of all elements of $G$, for $t \in [0, 1]$. If we start at some element $g(\phi_1, \phi_2, \ldots, \phi_d) \in G$ and let the

parameters vary continuously, the area bounded by this curve is called the connected component of $g$.

**Definition 4.2.** *Connected component of a Lie Group; $G' \subset G$*
A *connected component* of $G$ is a maximal subset of elements $G' \subset G$ that can be reached by a continuous variation of the parameters $\{\phi_1, \phi_2, \ldots, \phi_d\}$ of one element $g$.

A group $G$ can therefore have more than one connected component, which by themselves are disjoint subsets of $G$.
An important theorem to identify the irreps of Lie groups that has to do with the definition of the connected component of an element $g \in G$ is the following:

**Theorem 4.3.** *Connected component of identity*
If $G$ is a Lie group with identity $e$, then the connected component of the identity $e \in G$ is an invariant subgroup.

If we let the parameters of another element $g'$ vary continuously, we can trace out a curve very close to the curve traced out by $g$. If these curves can be transformed into each other by this continuous deformation, i.e. the parameters of $g$ depend continuously on those of $g'$, then the curves are said to be *continuously deformable* into each other. If the curves can be continuously deformed into a single point, which is a special kind of curve, they are said to be *contractible*. A type of Lie groups called *simply connected* is defined with the help of such contractable curves.

**Definition 4.4.** *Simply Connected group*
A $d$-dimensional Lie group is called *simply connected* if every *closed* curve in group-space, i.e. $g_i(\phi_1, \phi_2, \ldots, \phi_d) = g_i(\phi'_1, \phi'_2, \ldots, \phi'_d)$, is *contractible to a point*. With $(\phi_1, \phi_2, \ldots, \phi_d)$ the parameter values at the beginning and $(\phi'_1, \phi'_2, \ldots, \phi'_d)$ the values at the end of the curve.

A group with a parameter space represented by a sphere is for example contractable and a torus is not. However this is a subtle definition and we will see for example that $SO(3)$ parameter space is represented by a sphere, but is not simply connected however, due to some global constraints. This will become important when talking about so-called *covers* of the group.

**Definition 4.5.** *Compact Groups*
If the set parameters $\{\phi_i\}$ of $g(\phi_1, \phi_2, \ldots, \phi_d) \in G$ is *closed* and *bounded*, then the group $G$ is called *compact*.

For example, $SO(2)$ is compact, since all elements $R(\phi) \in SO(2)$ are described by the parameter $\phi$ that ranges over $[0, 2\pi]$, which is closed and bounded, i.e. is compact.

## 4.2 Representations of Lie Groups

When regarding Lie groups, it is good to realize that the defining elements themselves can be regarded as acting on a certain vector space already, i.e. the definition of the group is given explicitly by the *defining representation* acting on the natural vector

space in the given situation. With discrete groups, this in general was not the case. When we chose a certain vector space when regarding the symmetry group of a square, for example the 3D Euclidean vector space, we were actually fixing *some representation* of $D_4$. In this case obviously the 3D representation. But the group $D_4$ itself does not depend on what vector space we let its representations act. This structure is captured in the group multiplication table and is followed by all representations of the group, since they are per definition group homomorphisms.

With continuous groups this approach is slightly different. Now the elements themselves are defined to act on a certain vector space already. Therefore these groups are matrix groups, i.e. the *group elements* themselves are matrices of fixed dimensionality. This is called the *defining* or *fundamental* representation. For $SU(2)$ (see section 4.9) it is the set of $2 \times 2$ *unitary* matrices, and for $SO(3)$ it is the set of $3 \times 3$ *orthogonal* matrices. These Lie groups however can have infinitely many *other* representations of any dimension, some irreducible, some not, depending on what vector space they are supposed to act on.

### 4.2.1 Non-Compact Lie groups and unitary representations

For non-compact simple Lie groups there is a theorem that relates the dimension of the irrep to the unitarity of the irrep (see [7], [13] [17], for proof).

**Theorem 4.6.** *Unitarity and dimensionality of irreps of Non-compact groups*
A connected, simple, non-compact Lie group admits no unitary finite-dimensional representations besides the trivial one.

Although hard to prove, the implications of this theorem are important to us. We will see that the Poincaré group for example is non-compact, hence the unitary irreps, representing the elementary particles, must be infinite dimensional. This determines the classification of the irreps by a large amount.

## 4.3 Lie Algebra; $L$

Because the elements of a $d$-dimensional Lie group depend continuously on the $d$ parameters, we saw that we can think of these group elements as vectors in some $d$-dimensional topological space. To any point in this group we can furthermore define a *tangent space*, which is called the *Lie algebra* $L$ of dimension $d$ and contains all valuable information of the group. The basis vectors of this Lie algebra are called the *generators* of the group. We will see that the generators in a particular representation are found be taking derivatives of those representations with respect to the parameters in the neighborhood of the identity of the group. The other way around, would be constructing the group elements from the generators. The group elements aren't just some ordinary linear combination of the generators, but can be constructed by *exponentiating* these generators. With this exponential mapping all group elements in the neighborhood of the identity of the group are generated. We can restrict our attention therefore to this Lie algebra and its properties. The first thing we have to investigate is what are the basis vectors or generators of this algabra.

With the notions of continuity for a Lie group and group representations, we can define an infinitesimal transformation. We regard the $n$-dimensional fundamental representation $D$ of a $d$-dimensional group $G$ and in that representation an element $D(g(\phi_1, \phi_2, \ldots, \phi_d)) \equiv D(g(\boldsymbol{\phi}))$ is expanded around the identity for which we parametrize such that $\mathbb{1} = D(e) \equiv D(g(0))$:

$$\begin{aligned} D(g(\boldsymbol{\phi})) &= D(g(0)) + \sum_{k=1}^{d} \phi_k \frac{\partial D(g)}{\partial \phi_k}|_{\phi=0} + \mathcal{O}(\phi_k^2) \\ &= D(e) + (-i) \sum_{k=1}^{d} \phi_k X_k + \mathcal{O}(\phi_k^2) \\ &= \mathbb{1} + (-i)\phi_k X_k + \mathcal{O}(\phi_k^2) \end{aligned} \tag{4.1}$$

with $\mathbb{1}$ being the $n \times n$ identity matrix. The operators $X_k$ are called the *generators* of the group and are given by

$$(-i)X_k \equiv \frac{\partial D(g)}{\partial \phi_k}|_{\phi=0}. \tag{4.2}$$

Note that the dimension of the fundamental representation $n$, being the number of linearly independent basis vectors of the vector space $D(g)$ is acting on, is not the same as the dimension of the group $d$, being the number of independent parameters or equivalently linear independent generators. For unitary groups the following holds for infinitesimal transformations:

$$\begin{aligned} \mathbb{1} &= D^\dagger(g)D(g) \\ &= (\mathbb{1} + i\phi_k X_k^\dagger)(\mathbb{1} - i\phi_k X_k) \\ &= \mathbb{1} + i\phi_k(X_k^\dagger - X_k) + \mathcal{O}(\phi_k^2) \end{aligned} \tag{4.3}$$

which, for it to hold, implies that the generators must obey $X_k = X_k^\dagger$ and hence are hermitian by construction. Any complex $N \times N$ hermitian matrix with $\det(X) = 1$ has $N^2 - 1$ degrees of freedom, hence $SU(N)$ has $N^2 - 1$ linearly independent generators, hence $d = N^2 - 1$. Take $SU(2)$ for example, here $N = 2$ since the group defining representation acts on a 2 dimensional representation space, but the dimension $d$ of the group $SU(2)$ is 3. For orthogonal groups $SO(N)$ equation (4.3) leads to $X_k = -X_k^T$, i.e. anti-symmetric generators. In this case we have $N(N-1)/2$ linearly independent generators, thereby defining the dimension $d$ of $SO(N)$.

The Taylor series expansion in equation (4.1) of course comes down to an infinite series, which can be identified with the exponential series. Equation (4.1) can therefore be extended to

$$\begin{aligned} D(g(\boldsymbol{\phi})) &= \mathbb{1} + (-i)\phi_k X_k + (-i)^2 \frac{(\phi_k X_k)^2}{2!} + (-i)^3 \frac{(\phi_k X_k)^3}{3!} + \ldots \\ &= \sum_{l=0}^{\infty} \frac{(-i\phi_k X_k)^l}{l!} \\ &= e^{-i\phi_k X_k} \end{aligned} \tag{4.4}$$

So we conclude that every element in the neighborhood of the identity of a continuous group can be written as a complex exponent of the generators times the parameters. Written like this we easily see that unitarity of the group elements, i.e. $\left(e^{-i\phi_k X_k}\right)^\dagger = \left(e^{-i\phi_k X_k}\right)^{-1}$, implies hermiticity for the generators, i.e. $X_k = X_k^\dagger$. The orthogonal analog can also easily be checked.

These generators $X_k$ obviously depend, with equation (4.2), on the chosen representation $D$. The generators themselves therefore don't fully characterize the algebra or group, but the vector space they form, given their commutation relations, do.

## 4.4 Lie product; Structure Constants

As mentioned we can regard a $d$-dimensional group as a topology that can locally, i.e. in the neighborhood of the identity, be described by its $d$-dimensional tangent space. This tangent space forms a vector space or Lie algebra when certain conditions are met by the *Lie product* $[\ ,\ ]$ between two of its members:

1. (*Closure*) $[X_1, X_2] \in L,\ \forall X_1, X_2 \in L$
2. (*Linearity*) $[\alpha X_1 + \beta X_2, X_3] = \alpha[X_1, X_3] + \beta[X_2, X_3]\ ,\ \forall X_1, X_2, X_2 \in L,\ \forall \alpha, \beta \in \mathbb{R}$
3. (*Anti Symmetry*) $[X_1, X_2] = -[X_2, X_1]$
4. (*Jacobi Identity*) $[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0\ ,\ \forall X_1, X_2, X_3 \in L$

If the $X_i$ are generators of any matrix representation, this Lie product can be regarded as the ordinary matrix commutation relation. Since we won't need a Lie product other than the commutator, we will consider the Lie product always as being represented by the commutator from now on.

Since this Lie algebra $L$ has a finite dimension $d$ any vector in this space must be expandable in some basis $\{X_1, X_2, \ldots, X_d\}$. Since the product $[X_1, X_2]$ must itself be an element of $L$ we can expand it in this particular basis:

$$[X_i, X_j] = \sum_{k=1}^{d} f_{ij}^k X_k, \qquad i, j = 1, \ldots, d \tag{4.5}$$

The tensors $f_{ij}^k$ are the *structure constants* of the group and are of fundamental importance, since unlike the generators, they are *in*dependent of the chosen representation, apart from a basis transformation of the algebra. These structure constants define the multiplication rules for the Lie algebra. They furthermore obey, because of the constraints on Lie algebra, the following conditions:

$$f_{ij}^k = -f_{ji}^k \tag{4.6}$$

$$f_{ij}^k f_{mn}^j + f_{mj}^k f_{ni}^j + f_{nj}^k f_{im}^j = 0 \tag{4.7}$$

## 4.5 Additional structures of the Lie Algebra

Since the Lie algebra is a vector space with the generators as its basis, this vector space can itself have some substructure like the groups did. Furthermore other structures can be defined acting on the Lie product, being a part of the Lie algebra, itself.

**Definition 4.7.** *Subalgebra's: ideal, center*

A subset $N \subset L$ is called a *subalgebra* if $[N, N] \subset N$, i.e. it is closed under the Lie product. $N$ is called an *ideal* if $[L, N] \subset N$, i.e. $N$ is invariant under the Lie product $\forall X \in L$ and is therefore just an invariant subalgebra. The *center* of $L$ is the maximal ideal $N$ for which $[L, N] = 0$ holds and is always abelian, since automatically also $[N, N] = 0$ holds.

Since the algebra identity $0 \in L$ always commutes with all other algebra elements, the identity is always in the set $N$. If $N = \{0\}$, then $N$ is called *trivial*, if $N$ contains more elements then just the identity, it is called *non-trivial*. Since every group has an identity, every algebra has a 0 as the identity generator and hence $N = \{0\}$ at least.

In terms of the ideals and centre of an algebra, the algebra's can be classified as being *simple*, *semi-simple* or *non semi-simple*:

**Definition 4.8.** *Simple groups*

A group is called *simple* if its algebra has no proper ideals. The only ideal the algebra may have is therefore the trivial centre: $N = \{0\}$, since it is non proper.

**Definition 4.9.** *Semi-simple groups*

A group is called *semi-simple* if it has no proper abelian ideals, i.e. it can have proper ideals, but they can't be abelian. Since the centre is always abelian, a semi-simple group therefore must have a non-proper, i.e. trivial centre $N = \{0\}$. But, contrary to simple groups, a semi simple group can have additional proper ideals, as long as they aren't commutative.

**Definition 4.10.** *Non Semi-simple groups*

A group is called *non semi-simple* if it has abelian ideals. If its centre is non trivial, the group is immediately non semi-simple. If however the centre is trivial, the algebra contains at least one other ideal which is commutative.

In algebra's just as in groups, we can define equivalence relations that partitions the algebra into equivalence classes. Just as with groups we can derive the *quotient algebra's* in this sense.

We now consider what is called a direct sum of algebra's. This has everything to do with the direct product of groups, since a direct sum of algebra's is simply the algebra belonging to a direct product group as defined above.

**Definition 4.11.** *Direct sum of Algebra's*

Consider the set of Lie algebra's $\{M_i\}$ and let $L$ be the collection of vectors such that every $l \in L$ can be written in a *unique* way $l = \sum_i m_i$ with $m_i \in M_i$. Then $L$ is called the *direct sum* of $M_i$, written as $L = \dotplus M_i$. If furthermore the $M_i$ are ideals of $L$, i.e. $[M_i, M_i] \subset M_i$ and $[M_i, M_j] = 0$, then $L$ is said to be *decomposable* and the direct sum is written as $L = \bigoplus M_i$.

A direct product group $G = H \times H'$ can be written in term of its elements as $g = h_i h'_j$. Because of the definition for direct product groups, we know that $\forall h \in H, h' \in H'$ the elements commute $hh'h^{-1}h'^{-1} = e_G$ and that $ghg^{-1} \in H$ and $gh'g^{-1} \in H'$. However when we write the elements as an exponential mapping $exp\colon L \to G$, i.e.

$$g = \exp l \tag{4.8}$$

We now consider two algebra's $T$ and $M$ and let $D$ be a homomorphism $D\colon M \to T$, with $T$ the set of linear operators, such that the operator $D(X \in M)$ is a derivation of $T$. Now take the direct sum algebra of $M$ and $T$, i.e. $M \dot{+} T$, and define a Lie product for the mixed elements $m, t \in M \dot{+} T$ as:

$$[m, t] \equiv (D(m))(t) \tag{4.9}$$

If the product $[l_1, l_2]$ is calculated and $l_1, l_2 \in M$ ór $l_1, l_2 \in T$, the respectable Lie products are used.

**Definition 4.12.** *Semidirect sum of algebra's*
A Lie algebra $L$ is a *semidirect sum* of subalgebra's $T$ and $M$ if $L = M \dot{+} T$ and $[T, T] \subset T$, $[M, M] \subset M$ and $[M, T] \subset T$. The semidirect sum is written as $L = T \oplus M$, $T$ being an ideal of $L$ and $M$ a subalgebra of $L$.

## 4.6 Casimir Operators, irrep labels and CSCO

The set of generators which commute with all other generators in the Lie algebra, is called the set of *Casimir* operators $\{C_i\}$.

**Definition 4.13.** *Casimir operator $C_i$*
Given an set of generators $\{X_j\}$ that constitute a Lie algebra $L$, a Casimir operator is a combination of $\{X_j\}$ that commutes with all generators $X_j$, i.e. $\forall X_j \in L$

$$[C_i, X_j] = 0 \tag{4.10}$$

Because of this defining property, the Casimir operators represent observables that can be measured simultaneously with all other observables. The complete set of Casimir's therefore generates the maximal commuting set of eigenvalues and since every *eigenvalue* of a Casimir operator is unaltered by any transformation in that group, they thereby *label* the eigenvector in a group invariant way.

There is a general and technically difficult way to derive all Casimir operators for a given group, but we don't need that. We only need to verify that the given Casimir's indeed commute with the whole algebra.

The *Cartan method* is used for classifying irreps of Lie groups (see [7] for further details). It uses Casimir operators to do so. The basis vectors of a representation space $V$ are taken to be the eigenvectors of a set of mutually commuting operators, here being the generators of the group. Commuting operators can have complete sets of simultaneous eigenvectors. By the definition of the Casimir operators $C_i$, Schurs lemma states that any Casimir operator is mapped to a multiple of the unit matrix in an irreducible representation. Therefore all vectors of the irreducible representation are

eigenvectors of that Casimir operator, with the same eigenvalue $c_i$. These eigenvalues $c_i$ therefore label the particular irrep, and hence can be identified with the invariant properties of elementary particles.

**Theorem 4.14.** *Number of Casimir's*
Given a certain group, the number of independent Casimir operators equals the rank of the algebra. The rank of a Lie algebra is equal to the dimension of a Cartan sub-algebra. The Cartan sub-algebra is the maximal nilpotent sub-algebra that is self-normalizing

For proof see [7].
A quick reminder of the idea how we specify our quantum states is appropriate here. First of all one uses the eigenvalues of the Casimir operators corresponding to the symmetry group concerning the system. Then, given the generators of the group, one seeks to find the maximal or complete set of commuting operators (CSCO) to gather all information available. This is done by choosing one generator $X_i$ and then find out what other generators $X_j$ commute with $X_i$. If the system is degenerate, meaning that two states $|\psi_1\rangle$ and $|\psi_2\rangle$ have the same eigenvalue $x_k$ for some generator $X_k$, then there must be another generator $X_j$ commuting with $X_k$, such that the degenerate state $|x_k\rangle$ can be written in terms of the eigenvalues $|x_k, x_j\rangle$. Thereby solving the degeneracy of the system. This process continues until no degeneracy is left. The set of operators with non-degenerate eigenvalues is called the complete set of commuting operators (CSCO) and determines the system completely.

## 4.7 Representations of Lie Algebras

With discrete groups we found it interesting to find representations that would copy the group behavior, i.e. were group homomorphisms. With Lie groups however we are more interested in representing the *algebra* rather then the group itself. This is simply because one can generate the group representation easily from the algebra reps, by using the exponential mapping. The algebra is just the infinitesimal hence local description of the group around the identity. One however has to make sure that the global properties of the group are taken into account when generating the group representations. This algebra is a vector space with a finite $d$-dimensional basis, endowed with a group specific commutation relation and has therefore a finite representation obeying these Lie algebra multiplication rules, unlike the group representation which can have infinite dimension for Lie groups. Any representation of the algebra therefore has to respect equation (4.5).

### 4.7.1 Adjoint Representation of Lie Algebra

A frequently used and very important example is the *adjoint* representation of the Lie algebra. Consider a $d$ dimensional Lie algebra $L$, with generators $\{X_i\}$ which obey the commutation relations (4.5). We then define the adjoint representation $ad_{X_i}$ as:

$$\sum_{k=1}^{d} X_k [ad_{X_i}]_{km} = [X_i, X_m] \tag{4.11}$$

From this we can identify the structure constants with the matrix entries of the adjoint representation:

$$f_{im}^{k} = [ad_{X_i}]_{km}. \tag{4.12}$$

The importance of the adjoint representation is that the interaction fields, and more generally *any* operator fields transform in the adjoint representation.

## 4.8 What Discrete Group properties carry over to Lie groups?

To ensure that the theorems that hold for the Discrete groups are translated well into the domain of Continuous Lie groups, we have to make sure that summing over all elements is well defined. Since this was used in the GOT (see section 3.7.1) for discrete groups and with which all important properties could be derived. For Discrete groups this was no problem, since the group elements where countable, i.e. $\sum_g^{|G|}$ made sense. For Lie groups of dimension $n$ however, the elements are continuous functions of the $n$ parameters, so summing over all group elements turns into integrating over the whole group $\sum_g^{|G|} \to \int_G dG$, which in turn is translated into an integral over the parameter space depending on some density function $\rho$ which tells us "how many" elements are there within a certain range of the parameters:

$$\int_G dG \to \int \rho(\theta_i) d\theta_i \tag{4.13}$$

Were $d\theta_i$ is called an *invariant measure*, see [7] and [9] for details. For compact groups, for which the $\theta_i$ per definition are bounded, this is no problem and the integral converges. For non compact groups, this is general not true and one has to be careful to adopt the discrete theorems to quickly.

## 4.9 $SO(3)$ and $SU(2)$; Important compact groups

Thus far the only rotations we mentioned where discrete ones, i.e. rotations over fixed angles. If we want to consider systems which are invariant under any rotation, in say 3D space, we have to consider the group $SO(3)$. This can be easily seen be regarding a 3D rotation $R(\phi_k)$, $k = 1, 2, 3$. A rotation in 3D is defined as being an operations which leaves the real 3D inner product $(x, y) = x^T \cdot y$ invariant, thereby demanding orthogonality of $R(\phi_k)$, i.e.

$$R^T(\phi_k)R(\phi_k) = \mathbb{1}, \tag{4.14}$$

which therefore by definition constitute the group $O(3)$ of orthogonal $3 \times 3$ matrices. By taking the determinant of the left and right side we see that $\det[R(\phi_k)] = \pm 1$. Restricting to the proper rotations with determinant $+1$, thereby excluding reflections which change the handedness of the system or equivalently have determinant $-1$, the rotations $R(\phi_k)$ constitute the group $SO(3)$ of $3 \times 3$ orthogonal matrices with determinant $+1$. The dimension of $SO(3)$ is by accident also 3, this is because in general any $SO(N)$ group is generated by $N \times N$ antisymmetric matrices, which have $\frac{N(N-1)}{2}$ degrees of freedom. As can be derived from the analogon of equation (4.3) for the $SO(N)$ groups.

Since rotations in 3D are parametrized by three *bounded* parameters $\phi_k$, with $k = 1, 2, 3$, $SO(3)$ is a 3 dimensional *compact* group and therefore the unitary irreps can be taken to be *finite* dimensional, as we will see shortly. Note that the dimension of the group is a fixed quantity and the dimension of the representation depends on the representation space involved. This was the same for discrete groups, where it was not necessary for the order of the group to be equal to the dimension of the irreps, although it did put restraints on them by equation (3.43).

The fact that its irreps can be taken to be finite dimensional holds in fact for any compact group, including all $SO(N)$ groups. Because of its compactness, many of the tools and theories constructed for discrete groups cary over. For example the decomposition of representations of discrete groups into irreps, by making use of the character technique, can be used with compact groups as well.

The general way to find the generators in a certain finite dimensional representation of a group is to write down the representation and use equation (4.2). In the finite dimensional fundamental representation of $SO(3)$, i.e. the representation acting on vectors, the matrix elements of the hermitian $SO(3)$ generators $X_k$, are given by

$$(X_k)^i{}_j = -i\epsilon^i{}_{jk}, \tag{4.15}$$

This can be seen by writing out an infinitesimal 1st order rotation $R_{\mathbf{n}}(\phi)$ around a unit axis $\mathbf{n}$ and over angle $\delta\phi$ giving

$$\begin{aligned} x^i &\to x'^i \\ &= (R_{\mathbf{n}}(\phi))^i{}_j x^j \\ &= x^i + \delta x^i \\ &= x^i + (\delta\phi)(\mathbf{n}\times\mathbf{x})^i \end{aligned} \tag{4.16}$$

By defining a new axis $\tilde{\mathbf{n}} = (\delta\phi)\mathbf{n}$, we create a vector around which the rotation takes place and which length gives the angle of rotation. After renaming $\tilde{\mathbf{n}} \to \mathbf{n}$, we leave out $\delta\phi$ all together. Written in tensor notation the infinitesimal change $\delta x^i$ due to the rotation becomes

$$\begin{aligned} \delta x^i &= \epsilon^i{}_{jk} n^j x^k \\ &= -\epsilon^i{}_{jk} n^k x^j \\ &\equiv -i(X_{\mathbf{n}})^i{}_j x^j \end{aligned} \tag{4.17}$$

using equation (4.2). We therefore see from $(X_{\mathbf{n}})^i{}_j \equiv \mathbf{n}\cdot(\mathbf{X})^i{}_j = (X_k)^i{}_j n^k = -i\epsilon^i{}_{jk} n^k$ that the generators $X_k$ for this rotation take the form of $(X_k)^i{}_j = -i\epsilon^i{}_{jk}$, which proves equation (4.15). The $SO(3)$ generators satisfy the following commutation relations

$$[X_i, X_j] = i\epsilon_{ijk} X_k. \tag{4.18}$$

Equation (4.15) was derived by regarding the infinitesimal rotation of a vector $x^i$ and is therefore the fundamental representation of $SO(3)$. As mentioned in section 4.3 the generators of any $SO(N)$ group must be fully antisymmetric matrices and equation (4.15) explicitly shows that the finite dimensional representation of $SO(3)$ indeed is. It also is finite dimensional, since the vector space on which it acts, i.e. 3D space, is.

We can also construct a *infinite* dimensional representation of any Lie group by letting it act on a infinite dimensional vector space, for which we can take the vector space of continuous functions $\Phi(x^j)$. This representation space is infinite dimensional because any scalar function $\Phi(x^j)$ can be expanded in the infinite dimensional complete orthogonal basis of its Fourier components, as we saw in section 3.4. For $SO(3)$ this procedure gives

$$\begin{aligned}\Phi(x^j) &\to \Phi'(x'^j)\\ &= R_{\mathbf{n}}\left[\Phi(x^j)\right]\\ &= \Phi(x^j) + \delta\left(\Phi(x^j)\right)\end{aligned} \tag{4.19}$$

Now, by making use of the expression (4.17) for the change in $x^i$, the infinitesimal change $\delta\Phi$ takes the following form

$$\begin{aligned}\delta\left(\Phi(x^j)\right) &= \partial_i\Phi(x^j)\delta x^i\\ &= \left[-n^k\epsilon^i{}_{jk}x^j\partial_i\right]\Phi(x^j)\\ &\equiv \left[-in^kX_k\right]\Phi(x^j)\end{aligned} \tag{4.20}$$

again by identification using equation (4.2). We conclude that the generators for the infinite dimensional representation take the form

$$X_k = -i\epsilon^i{}_{jk}x^j\partial_i \tag{4.21}$$

which, for the quantum mechanical definition of the momentum operator $p_i = -i\partial_i$ with $\hbar \equiv 1$, equals the familiar expression for the angular momentum operator $J_k$

$$J_k = \epsilon^i{}_{jk}x^jp_i = (\mathbf{x}\times\mathbf{p})_k \tag{4.22}$$

However as we saw in equation (4.3) any generator of an $SO(N)$ group must in general be an $N\times N$ antisymmetric tensor, also known as a 2-form $M_{ij}$. In general a p-form is just a fully antisymmetric tensor with p indices, hence the requirement of antisymmetry is also automatically fulfilled in the case of $SO(N)$. We conclude that therefore any $SO(N)$ has 2-forms $M_{ij}$ as generalized generators. So what do these generators $M_{ij}$ have to do with $J_k$ for $SO(3)$? They are, what is called each others 'dual operator'.

**Definition 4.15.** *Hodge Dual;* $*T$

Let $T_p$ be a p-form, acting on a $N$-dimensional space. Furthermore let $\epsilon_{\mu_1\mu_2\ldots\mu_p\nu_1\nu_2\ldots\nu_q}$, with $p+q=N$ be the unique fully antisymmetric $N$-tensor. Then the dual of $T_p$ is the $(N-p)$ tensor $*T$, with components

$$(*T)_{\nu_1\nu_2\ldots\nu_q} = \frac{1}{p!}\epsilon_{\nu_1\nu_2\ldots\nu_q}{}^{\mu_1\mu_2\ldots\mu_p}T_{\mu_1\mu_2\ldots\mu_p} \tag{4.23}$$

The Hodge dual $*T$ represents the same physical quantity as the original operator $T$. This can be seen by the fact that $*T$ transforms in the same way and has the same amount of degrees of freedom as $T$ does.

Now taking the Hodge dual of $J_k$ in a 3 dimensional space by

$$\begin{aligned}M_{ij} &\equiv \epsilon_{ij}{}^kJ_k\\ &= -i\epsilon_{ij}{}^k\epsilon_{km}{}^nx^m\partial_n\\ &= -i(\delta_{im}\delta_{jn}-\delta_{in}\delta_{jm})x^m\partial_n\\ &= -i(x_i\partial_j - x_j\partial_i)\end{aligned} \tag{4.24}$$

and we thus get the antisymmetric 2-form $M_{ij}$ as the generators, as requested by the group properties of $SO(3)$. We conclude that the natural generators of any $SO(N)$ in any dimension $N$, are given by the generalization of equation (4.24). For $N = 3$, the dual of this 2 form turns out to be a 1 form that can be represented by the angular momentum vector $J_k$.

The equations (4.15) and (4.21) are respectively the finite and the infinite representations of the algebra of $SO(3)$ characterized by the commutation relations (4.18). Upon exponentiation, any group element of $SO(N)$ can be written as

$$R(\omega_{ij}) = e^{-i\omega_{ij}M^{ij}} \tag{4.25}$$

and for $N = 3$ this can be written in terms of the angular momentum operators $X_k$ as

$$R(n_k) = e^{-in_kX^k}. \tag{4.26}$$

We can reparametrize the 3 parameters or rotation angles $\phi_k$ as $\phi\mathbf{n}$, with $\mathbf{n}$ being the normalized 3D axis of rotation and $\phi$ the angle of rotation around $\mathbf{n}$. After this reparametrization equation (4.25) becomes

$$R_{\mathbf{n}}(\phi) = e^{-i\phi\mathbf{n}\cdot\mathbf{X}} = \mathbb{1}\cos\phi + i\mathbf{n}\cdot\mathbf{X}\sin\phi. \tag{4.27}$$

This reparametrization later will help us to find a relation between $SO(3)$ and the important group $SU(2)$, which is the group of $2\times 2$ unitary matrices with determinant 1. It also clearly shows, by identifying the parameter space of SO(3) with a sphere $S_2$ of radius $\pi$, that two rotations around the same axis $\mathbf{n}$ over $\phi = \pi$ and $\phi = -\pi$ must be identified, i.e. every pair of opposite points on the surface of $S_2$ represent the same rotation. Since there exists no contractible curve between two such points, we conclude that $SO(3)$ is not simply connected. A bit more on the relation between $SU(2)$ and $SO(3)$ is next.

## 4.10 $SU(2)$ is the universal covering group of $SO(3)$

As we saw one choice of parametrizing any $SO(3)$ rotation can be made by taking two parameters to specify a 3D unit vector $\hat{n}$, around which is to be rotated, and a third parameter $\theta, 0 \leqslant \theta \leqslant \pi$, the angle of rotation. Following the line of [11] constructing out of those parameters a vector $\mathbf{n} = \theta\hat{n}$ one can see that the corresponding parameter space spanned by $\mathbf{n}$ therefore can be represented by a solid sphere of radius $\pi$. Every point in this sphere represents some rotation in 3D space but importantly, two opposite points on the surface, i.e. $-\pi\hat{n}$ and $\pi\hat{n}$, are identified $-\pi\hat{n} \equiv \pi\hat{n}$. They represent one and the same 3D rotation, since it does not matter if one rotates over $\pi$ radians around $-\hat{n}$ or $+\hat{n}$. This important property of the $SO(3)$ parameter space causes $SO(3)$ to be connected, but not simply connected. This fact can be demonstrated by noting that there are two types of closed curves in this parameter space. Letting the curves start at the identity, which in the parameter space is represented by the origin of the sphere, we can distinguish two types of closed curves: for the first type we let the curve stay inside the sphere, not touching its surface and ending at the identity. This type of closed curves is clearly contractable to a point in a smooth way.

The second type is obtained by starting the curve at the origin and than reach for the surface in the $\hat{n}$ direction, by letting $\theta$ reach $\pi$, continue at the identified opposite of that surface point and then let $\theta$ increase to $2\pi$ which means that the curve starts goes from the surface at $\theta = \pi$ reaches the identity where $\theta = 2\pi$ again. This expresses the equivalency of rotating around $\hat{n}$ for angles $\theta \geqslant \pi$ and rotating around $-\hat{n}$ over $\theta - \pi$, i.e. the domain $\theta(\text{mod } \pi)$. We see that, as expected, the identity at the origin is parametrized by $\theta = 0 = 2\pi$. This line therefore just represents all rotations up to $2\pi$ around some axis $\hat{n}$. This second type of closed curves simply looks like a line, straight or curved does not matter for being closed, through the origin of the sphere from one point on the surface to the exact opposite of this point on the sphere. Hence this last type of closed curves cannot be contracted to a point in a smooth way, without breaking it up. It means that for these curves a rotation over $2\pi$ is continuously connected to the identity but not simply connected to it and therefore one cannot identify the identity with rotation over $2\pi$ for these curves: They are not continuously deformable into eachother, since the curve that connects them is closed, rotation over $2\pi$ and 0 are both represented by the origin, but not contractable.

However since passing the origin was only required to close the curve, we can let the curve hit the surface a second time before returning to the origin. This means that we let $\theta$ go from 0 to $4\pi$. This curve is still of the second type since it reaches the sphere's surface, but now it *is* contractable in a smooth way, by pulling the non-diametrically opposed surface points two by two towards each other until the curve is contracted towards two diametrically opposed surface points, which are identical. Hence the closed curve is smoothly contracted to one point and these curves become contractable after a rotation over $4\pi$, meaning that in this case a rotation over $4\pi$ and the identity *can* be identified.

So we conclude that for any closed curve if $\theta$ is increased by $k \times 2\pi$ for $k$ even all curves are contractable and for $k$ odd, some of them are not. The fact that the parameter space of $SU(2)$ is simply connected, means that there is just one class of contractable curves. This means that all representations are single valued. Whereas for SO(3), which is not simply connected, has two classes of closed curves and hence its irreps are double valued.

## 4.11 Irreps of $SO(3)$

In order to derive the irreps of $SO(3)$ we use the fact that the commutation relations of $X_k$, the $SO(3)$ generators and $J_k$, the angular momentum operators in quantum mechanics, can be identified. Starting from the commutation relations of $J_k$,

$$[J_i, J_j] = i\epsilon_{ijk}\hbar J_k \tag{4.28}$$

thereby taking them as fundamental, we will see that the eigenvalues of the operators allow for integer and half-integer values. In general these commutation relations can be derived for the orbital angular momentum operator $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ and the fundamental commutation relations of position and momentum $[r_i, p_j] = i\hbar\delta_{ij}$. This is needed to describe the intrinsic angular momentum, or spin, of all physical particles. Since the operators don't commute, the eigenvalues can't be assigned to all $J_k$ at once. As always we choose the $z$-direction as our preferred direction of quantization. It is easily verified

that $\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2$ does commute with every $J_k$ and that we therefore can define states $|m, \beta\rangle$, which are simultaneous eigenstates of $J_z$ and $\mathbf{J}^2$, as

$$\begin{aligned} \mathbf{J}^2|m,\beta\rangle &= \beta^2\hbar^2|m,\beta\rangle \\ J_z|m,\beta\rangle &= m\hbar|m,\beta\rangle \end{aligned} \tag{4.29}$$

To derive the possible values of $m$ and $\beta$ we introduce the ladder operators, $J_\pm = J_x \pm iJ_y$ and derive the commutation relations

$$[\mathbf{J}^2, J_\pm] = 0 \tag{4.30}$$

$$[J_z, J_\pm] = \pm\hbar J_\pm \tag{4.31}$$

We then have

$$\begin{aligned} J_z(J_\pm|m,\beta\rangle) &= (J_\pm J_z + [J_z, J_\pm])|m,\beta\rangle \\ &= (J_\pm m\hbar \pm \hbar J_\pm)|m,\beta\rangle \\ &= (m \pm 1)\hbar(J_\pm|m,\beta\rangle) \end{aligned} \tag{4.32}$$

Hence the name, ladder operators.

It is furthermore easy to derive that necessarily $\beta^2 \geqslant m^2$, meaning that the raising and lowering of the $J_z$ eigenvalues must somewhere come to an end. The maximum value of $m$ we call $j$, and one can calculate the eigenvalues of $\mathbf{J}^2$ , by writing $\mathbf{J}^2$ succinctly as $J_-J_+ + J_z(J_z + \hbar)$, which gives $\beta^2 = j(j+1)$ and as $J_+J_- + J_z(J_z - \hbar)$, which gives $\beta^2 = m(m-1)$. It follows that and $m = \{-j, -j+1, \ldots, j-1, j\}$, having $2j+1$ values. The number $2j+1$ must therefore be an integer, and $j$ integer or half-integer as we claimed in the beginning. The equations (4.29) can thus be rewritten as the more familiar

$$\begin{aligned} \mathbf{J}^2|m,j\rangle &= j(j+1)\hbar^2|m,j\rangle \\ J_z|m,j\rangle &= m\hbar|m,j\rangle \end{aligned} \tag{4.33}$$

And by normalizing we get for the ladder operators

$$\begin{aligned} J_+|m,j\rangle &= \hbar\sqrt{(j-m)(j+m+1)}|m+1,j\rangle \\ J_-|m,j\rangle &= \hbar\sqrt{(j+m)(j-m+1)}|m-1,j\rangle \end{aligned} \tag{4.34}$$

By the identification of the algebra's of $X_k$ and $J_k$ we conclude that the irreps of $SO(3)$ follow the same pattern, just derived for the angular momentum operators: The irreps $|m, j\rangle$ of the $SO(3)$ generators $X_k$ are labeled by $j$ and are $2j+1$ dimensional, which is, as was mentioned, finite for $j < \infty$. The irreps of the $SO(3)$ group are now identified with the representation space $|m, j\rangle$ on which the operators $X_k$ act. This is common use; representations of a group are defined by the operators representing the group, together with the representation space on which the operators act and it is in this space that the wave functions live. Within a particular irrep $|m, j\rangle$, the eigenvalues of $X_z$ and the actions of $X_\pm$ are therefore given by

$$\begin{aligned} X_z|m,j\rangle &= m|m,j\rangle \\ X_+|m,j\rangle &= \sqrt{(j-m)(j+m+1)}|m+1,j\rangle \\ X_-|m,j\rangle &= \sqrt{(j+m)(j-m+1)}|m-1,j\rangle \end{aligned} \tag{4.35}$$

Since a finite group element of $SO(3)$ is given by exponentiating the algebra, it is sufficient to regard the irreps of the algebra instead of the whole group, as we did here. Again, the analysis led us to the fact that $2j+1$, being the multiplicity of $m$ or the dimensionality of irrep $|m,j\rangle$, must be an integer and hence $j$ is an integer or half-integer.

The difference between general rotation generator say $J_i$, which is the operator associated with *total* angular momentum and the orbital angular momentum generator $L_i$, which generates *translations around a circle* explicitly not being rotations. They have exactly the same algebra, but because $L_i$ generates translations its global restrictions are somewhat different than those of $J_i$, resulting in the allowed integer for eigenvalues the orbital angular momentum generator $L_i$ and the integer and half-integer eigenvalues of the total angular momentum generator $J_i$, leaving room for the Spin angular momentum to be integer and half integer valued.

However we can realize the commutation relations (4.18) of the generators in a basis of $2\times 2$ Pauli matrices $\sigma_k$, instead of the integer $j$ matrix representations given by equation (4.35), with $j=\{0,1,2,\ldots\}$. The Pauli matrices are given by

$$\sigma_1=\begin{pmatrix}0&1\\1&0\end{pmatrix},\sigma_2=\begin{pmatrix}0&i\\-i&0\end{pmatrix},\sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}. \tag{4.36}$$

This can be done by equating

$$X_k=\frac{1}{2}\sigma_k \tag{4.37}$$

and then showing that for these generators equation (4.18) still holds. This factor $\frac{1}{2}$ turns out to be rather important: it distinguishes between fermion and boson representations. For the case $SO(3)$, with $j$ restricted to the be an integer, the dimensions of the irreps where $\{1,3,5,\ldots\}$. Therefore only bosonic state vectors, such as $|0,0\rangle, |1,-1\rangle$ or $|1,0\rangle$ could be represented. Now however, with the generators given by $2\times 2$ Pauli matrices $\sim\sigma_k$, also 2 dimensional state vectors, like $|\frac{1}{2},-\frac{1}{2}\rangle$ and $|\frac{1}{2},\frac{1}{2}\rangle$ can be represented. The generators $X_k=\frac{1}{2}\sigma_k$ are hermitian and traceless and therefore form a basis for the group of unitary $2\times 2$ matrices $U$ with $\det U=+1$, i.e. the group $SU(2)$. Exponentiating these generators like we did for $SO(3)$ and can in general be done for any compact group, we generate all $SU(2)$ group elements $U$

$$U(\theta_k)=e^{-i\theta_k X^k}=e^{-i\frac{1}{2}\theta_k\sigma^k}. \tag{4.38}$$

Which can be rewritten as

$$U_{\mathbf{n}}(\theta)=\mathbb{1}\cos\frac{1}{2}\theta+i\mathbf{n}\cdot\boldsymbol{\sigma}\sin\frac{1}{2}\theta, \tag{4.39}$$

by making use of $(\mathbf{n}\cdot\boldsymbol{\sigma})^2=\mathbb{1}$. We immediately recognize the $4\pi$ periodicity of the $SU(2)$ elements in contrast to the $2\pi$ periodicity of the $SO(3)$ group elements.

Since both the elements of SO(3) and SU(2) are parametrized by the same parameters $\phi,\mathbf{n}$, we can find a homomorphism $f\colon SU(2)\to SO(3)$. However for $R\in SO(3)$ we have that $R_{\mathbf{n}}(2\pi)=R_{\mathbf{n}}(0)=1$, but for $U\in SU(2)$ we get for the same angles $U_{\mathbf{n}}(0)=1$, but $U_{\mathbf{n}}(2\pi)=-1$. Therefore the angles $0,2\pi$ are mapped by $f$ onto the identity of $SO(3)$, whereas in $SU(2)$ they parametrize two distinct elemetents. Hence the kernel of $f$ is

non trivial and discrete, i.e. $\text{Ker} f = \{\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right), \left(\begin{smallmatrix}-1&0\\0&-1\end{smallmatrix}\right)\} = Z_2$ and $SU(2)$ is the universal cover of $SO(3)$ with the precise relation $SU(2)/Z_2 \simeq SO(3)$. Furthermore $SU(2)$ is simply connected, since the parameterspace of $SU(2)$ is isomorphic to a 2 sphere $S_2$ with radius $2\pi$ but in this case *all* points on the surface are identified, instead of two opposite points as was the case with $SO(3)$. They share the same algebra, i.e. equation (4.18), hence they are locally isomorphic, but the parameter space of $SU(2)$ is twice the size of that of $SO(2)$, so globally homomorphic.

**Theorem 4.16.** *$SO(3)$, $SU(2)$ Casimirs and Invariant labels*
The Casimir operator for $SO(3)$ and because they share the same algebra, also $SU(2)$, is

$$J^2 = J_{\mu\nu} J^{\mu\nu} \tag{4.40}$$

and the related invariant physical quantity is the total angular momentum, which eigenvalues are given by $j(j+1)$. It is easy to show that $J^2$ commutes with all generators $J_i$.

# 5 The Lorentz and Poincaré Group

By deriving the Killing vector fields for the Minkowski metric at the end of section 2.5 we got generators for the Poincaré group, see equation (2.58). Knowing this is the isometry group of Minkowski space, we are interested in its irreps, because they represent the elementary particles.

Special Relativity requires that two inertial observers would observe the same physical laws. The Minkowski spacetime then dictates the transformations relating the two observers and the homogeneous part of it are called the Lorentz transformations. Apart from the Killing vector method the Lorentz transformations relating two inertial frames $F$ and $F'$ can equivalently be derived using the experimental fact that the speed of light has the same value $c$ in both frames. Regarding the infinitesimal distance travelled by the light as the pythagorean $d\mathbf{x}^2 = c^2dt^2$ and $d\mathbf{x}'^2 = c^2dt'^2$ in $F$ and $F'$ respectively, this leads to

$$-c^2dt^2 + d\mathbf{x}^2 = 0 = -c^2dt'^2 + d\mathbf{x}'^2, \tag{5.1}$$

where the invariant distance $-c^2dt^2 + d\mathbf{x}^2$ is called the infinitesimal *spacetime interval* and is denoted by $ds^2$. Of course $ds^2$ is only equal to zero for the trajectories of objects that travel with the speed of light. In general there are three types of trajectories

$$ds^2 = \begin{cases} < 0 & \text{(timelike)} \\ = 0 & \text{(lightlike)} \\ > 0 & \text{(spacelike).} \end{cases} \tag{5.2}$$

However $ds^2$ is conserved under this transformation, no matter what its value is. This is because equation (5.1) *defines* how this transformation should be executed in order to leave the speed of light invariant. We therefore can leave out the $= 0$ part and just state that any two events separated by a spacetime interval $ds^2$ in inertial frame $F$ will be separated by the same spacetime interval $ds^2$ in the other inertial frame $F'$, which moves with a relative speed with respect to $F$, i.e.

$$-c^2dt^2 + d\mathbf{x}^2 = -c^2dt'^2 + d\mathbf{x}'^2. \tag{5.3}$$

In order to write the Lorentz transformations in common notation, define the contravariant four vector components as $x^\mu = (cx^0, x^i) = (ct, \mathbf{x})$ and the covariant components $x_\mu = (cx_0, x_i) = (-ct, \mathbf{x})$. Equation (5.3) then, which just says that light has the same speed in every inertial frame in $3D$ space, can be rewritten in a compact way as the inner product on the $4D$ Minkowski space

$$dx_\mu dx^\mu = dx_{\mu'}dx^{\mu'} \tag{5.4}$$

The inner product in equation (5.4) or the finite interval version $x_\mu x^\mu$ defines the Minkowski metric $\eta_{\mu\nu}$ on the Minkowski space, by writing

$$x_\mu x^\mu \equiv \eta_{\mu\nu}x^\nu x^\mu = -(x^0)^2 + (\mathbf{x})^2. \tag{5.5}$$

In Minkowski space this metric therefore can be interpreted as the diagonal matrix $\text{diag}(-1, 1, 1, 1)$, and from $x_\mu = \eta_{\mu\nu}x^\nu$ we see that it can be used to lower indices.

Its inverse, written as $\eta^{\mu\nu}$, so happens to have the same diagonal entries as $\eta_{\mu\nu}$ in Minkowski space and is subsequently used to raise indices. Being the inverse of $\eta_{\mu\nu}$ they obey

$$\eta_{\mu\nu}\eta^{\nu\rho} = \delta_{\mu}{}^{\rho} \tag{5.6}$$

where $\delta_{\mu}{}^{\rho}$ is merely the Kronecker delta, in Minkowsky notation, used because one sums over upper and lower indices in the Einstein summation convention, which explains why the inverse of $\eta_{\mu\nu}$ was written as $\eta^{\mu\nu}$ in the first place. The invariance of the finite spacetime interval $s^2$ is then expressed as

$$\eta_{\mu\nu}x^{\mu}x^{\nu} = \eta_{\mu'\nu'}x^{\mu'}x^{\nu'} \tag{5.7}$$

Writing out equation (5.7) one can derive the linear Lorentz transformations for an observer in the primed frame $F'$ moving with velocity $v$ in the $x^1$-direction with respect to the observer in the unprimed frame $F$, they are then given by:

$$\begin{aligned} x^{0'} &= \gamma(x^0 - x^1 v) \\ x^{1'} &= \gamma(x^1 - v x^0) \\ x^{2'} &= x^2 \\ x^{3'} &= x^3, \end{aligned} \tag{5.8}$$

where $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ and then putting $c \equiv 1$.
Equations (5.8) are the Lorentz transformations when $F'$ moves in the $x^1$-direction of $F$ and are therefore not the most general transformations. In general any transformation $x^{\mu} \to x^{\mu'}$ which satisfies equation (5.3) or equivalently (5.4) is of the form

$$x^{\mu'} = \Lambda^{\mu'}{}_{\nu}x^{\nu} + a^{\mu} \tag{5.9}$$

which is exactly the transformation generated by the Killing vectorfield for the Minkowski metric. These are the Poincaré transformations, also known as the *inhomogeneous* Lorentz transformations and they leave, by definition, the spacetime interval $s^2$ invariant. The *homogeneous* Lorentz transformations is the subset of these transformations that leave the spacetime interval invariant but furthermore, by putting $a^{\mu} = 0$, leave the origin invariant and therefore this subset of transformations only consist of the constant matrices $\Lambda^{\mu'}{}_{\nu}$. These transformations $x^{\mu} \to x^{\mu'} = \Lambda^{\mu'}{}_{\nu}x^{\nu}$ can therefore be regarded as some special rotations in $4D$ spacetime. Combining these homogeneous Lorentz transformations with equation (5.7) gives us the transformation of the metric itself

$$\begin{aligned} \eta_{\mu\nu}x^{\mu}x^{\nu} &= \eta_{\mu'\nu'}x^{\mu'}x^{\nu'} \\ &= \eta_{\mu'\nu'}\Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}x^{\mu}x^{\nu} \end{aligned}$$

which results in

$$\eta_{\mu\nu} = \eta_{\mu'\nu'}\Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}. \tag{5.10}$$

Any set of transformation matrices $\Lambda$ that leaves the metric $\eta$ invariant and thereby conserves the inner product, i.e. for which $\eta_{\mu'\nu'} = \eta_{\mu\nu}$, forms the pseudo-orthogonal rotation group $O(3,1)$. Now to find the proper notation for the inverse Lorentz transformation $(\Lambda^{-1})^{\mu}{}_{\nu}$ we combine equations (5.6) and (5.10) to get

$$\eta_{\mu\nu}\eta^{\nu\rho} = \delta_{\mu}{}^{\rho} = \eta_{\mu'\nu'}\Lambda^{\mu'}{}_{\mu}\Lambda^{\nu'}{}_{\nu}\eta^{\nu\rho}. \tag{5.11}$$

Looking at the right hand side of equation(5.11) we see that after rearranging the order a bit, $\eta_{\mu'\nu'}\Lambda^{\nu'}{}_{\nu}\eta^{\nu\rho}$ must represent the inverse of $\Lambda^{\mu'}{}_{\mu}$, i.e. $(\Lambda^{-1})^{\rho}{}_{\mu'}$. We therefore have

$$(\Lambda^{-1})^{\rho}{}_{\mu'} = \eta_{\mu'\nu'}\Lambda^{\nu'}{}_{\nu}\eta^{\nu\rho} \equiv \Lambda_{\mu'}{}^{\rho}, \tag{5.12}$$

where explicit use was made of the raising and lowering of the indices by the metric. Therefore $\Lambda_{\mu'}{}^{\rho}$ is just a shorthand notation for the larger expression $\eta_{\mu'\nu'}\Lambda^{\nu'}{}_{\nu}\eta^{\nu\rho}$, which can be calculated since all factors in this product are actually well defined and can be interpreted as matrices. $\Lambda^{\mu\nu}$ for example only has meaning, because it can be interpreted as a product of two well defined matrices $\Lambda^{\mu\nu} = \eta^{\mu\rho}\Lambda^{\mu}{}_{\rho}$. Furthermore for any $SO(p,q)$ group, the inverse of a group element per definition always equals the transpose of that group element in the matrix representation. So we can state that when interpreting $\Lambda$ as a matrix, for the entries of its transpose $\Lambda^T$

$$(\Lambda^T)^{\rho}{}_{\mu'} = \eta_{\mu'\nu'}\Lambda^{\nu'}{}_{\nu}\eta^{\nu\rho} \equiv \Lambda_{\mu'}{}^{\rho}, \tag{5.13}$$

is the valid expression. Taking the determinant leaves us with $\det\Lambda = \pm 1$, and choosing $+1$ gives us the special orthogonal group $SO(3,1)$. Taking the $\mu = \nu = 0$ component of the left side of equation (5.10), we can derive that the whole $SO(3,1)$ group is disconnected and splits up into four regions depending on the sign of the determinant and the value of $\Lambda^0{}_0$:

1. $\Lambda^0{}_0 \geqslant 1$, $\det\Lambda^{\mu}{}_{\nu} = 1$: Proper Orthochronous Lorentz group
2. $\Lambda^0{}_0 \geqslant 1$, $\det\Lambda^{\mu}{}_{\nu} = -1$: Non-Proper Orthochronous Lorentz group
3. $\Lambda^0{}_0 \leqslant -1$, $\det\Lambda^{\mu}{}_{\nu} = 1$: Proper Non-Orthochronous Lorentz group
4. $\Lambda^0{}_0 \leqslant -1$, $\det\Lambda^{\mu}{}_{\nu} = -1$: Non-Proper Non-Orthochronous Lorentz group

We are only interested in the Proper Orthochronuous Lorentzgroup and from now on we will refer to it as simply the Lorentz group and denote it by $SO(3,1)$. The inhomogeneous Lorentz group or Poincaré group is denoted by $ISO(3,1)$. For infinitesimal rotations we saw in the previous section how to derive the generators for the rotations. We can do the same for the boosts. We first rewrite a general Lorentz coordinate transformation in terms of an angle $\zeta$. One can rewrite equations (5.8) as some sort of rotation, by defining an angle $\zeta$ as $\gamma = \cosh\zeta$. This implies that $\tanh\zeta = \frac{v}{c}$ and we get

$$\begin{aligned} x'^0 &= -x^1\sinh\zeta + x^0\cosh\zeta \\ x'^1 &= x^1\cosh\zeta - x^0\sinh\zeta \\ x'^2 &= x^2 \\ x'^3 &= x^3 \end{aligned} \tag{5.14}$$

They look a lot like the normal rotation would, except with cos and sin replaced by their hyperbolic counterparts. This however is crucial, since for proper rotations the angle $\phi$ over which to rotate is a *bound* parameter, since $\sin\phi$ and $\cos\phi$ are periodic functions identifying $\phi = 0$ with $\phi = 2\pi$, making the group of rotations *compact*. With the Lorentz transformations however, $\zeta$ is unbounded, because $\cosh\zeta$ and $\sinh\zeta$ lack this periodicity, resulting in a *non compact* group.

The generators of the ordinary rotations in $4D$ are then simply given by those of $SO(3)$ with an extra row and column of 0's to make it 4D, i.e. $X_k^4 = \begin{pmatrix} 0 & \rightarrow \\ \downarrow & X_k^3 \end{pmatrix}$ and the generators for the boosts $Y_k$; by writing equations (5.14) in infinitesimal form, we obtain the generators $Y_k$ for the boosts. It turns out that they can be written in terms of their matrix entries as

$$(Y_k)^\mu{}_\nu = -i(\eta_{\mu 0}\eta^\nu{}_k - \eta_{\mu k}\eta^\mu{}_0), \tag{5.15}$$

where $\eta^\nu{}_k = \eta_{k\mu}\eta^{\nu\mu}$, etc. The group $SO(3,1)$ is related by the group $SO(4)$ by a Wick 'rotation' $t \to it$, which has the minus sign in the $SO(3,1)$ metric as a consequence and consequently tranfsorms the compact group $SO(4)$ to a non-compact group $SO(3,1)$, because of the aforementioned appearing of the hyperbolic functions. However for the *number* of generators of a group this has no consequence. Therefore the number of generators of $SO(3,1)$ equals that of $SO(4)$, i.e. since $SO(4)$ generators must be anti-symmetric and therefore traceless, which follows also from orthogonality and $\det(O \in SO(4)) = 1$ respectively, we have $\frac{4\times 3}{2} = 6$ independent generators for the Lorentz group, 3 boosts and 3 rotations.

The Wick rotation, which merely is some complexification of the algebra, however does turn some hermitian generators into anti-hermitian ones, i.e. rotations into boosts. Which has the very important and immediate consequence that the defining finite dimensional vector representation cannot be *unitary*, for this demands hermiticity of all generators.

The total algebra of the Lorentz group is given by

$$\begin{aligned} [X_i, X_j] &= i\epsilon_{ijk}X_k \\ [X_i, Y_j] &= i\epsilon_{ijk}Y_k \\ [Y_i, Y_j] &= -i\epsilon_{ijk}X_k \end{aligned} \tag{5.16}$$

which shows that the rotations form a subgroup within $SO(3,1)$, in contrast to the boosts. One can derive that there exists no *real* basis transformation $X_k, Y_k \to X'_k, Y'_k$, with these new generators such that some of them constitute an invariant subalgebra. Therefore we conclude that $SO(3,1)$ must be a *simple* group.

## 5.1 Finite dimensional non-unitary irreps of Lorentz group

By making a complex basis transformation of the Lie algebra of the Lorentz group

$$M_k = \frac{1}{2}(X_k + iY_k) \tag{5.17}$$

and

$$N_k = \frac{1}{2}(X_k - iY_k) \tag{5.18}$$

we can rewrite equations (5.16) as

$$\begin{aligned} [M_i, M_j] &= i\epsilon_{ijk}M_k \\ [N_i, N_j] &= i\epsilon_{ijk}N_k \\ [M_k, N_l] &= 0 \end{aligned} \tag{5.19}$$

We see that under this Wick rotation, the $SO(3,1)$ algebra splits up into two $SU(2)$ subalgebra's which commute and, hence form the algebra of the direct product group $SU_A(2) \times SU_B(2)$. However because of this complex basis transformation, we cannot conclude that $SO(3,1) \simeq SU(2) \times SU(2)$, but rather $SO(3,1) \simeq SL(2,\mathbb{C})$ which is its double cover, just like $SU(2)$ is the double cover of $SO(3)$. This means that $SO(3,1)$, $SU_A(2) \times SU_B(2)$ and $SL(2,\mathbb{C})$ all share the same complex algebra, but are not all single valued. The whole classification of these algebra's can be found in [7]. For the labeling of the irreps however this distinction is irrelevant, since it merely means that multiple valued irreps of $SO(3,1)$ are single valued irreps of $SL(2,\mathbb{C})$ and we therefore can label the $SO(3,1)$ irreps by the labels $j$ of the $SU(2)$'s, i.e. by the combination $(j_A, j_B)$. The irreps of $SO(3,1)$ therefore have dimensions $(2j_1+1)(2j_2+1)$. The trivial irrep is labeled by $(0,0)$ and has dimension 1, and for example scalar fields transform in this representation. The second lowest dimensional irrep is given by the so called Weyl spinors and are labeled by $(0, \frac{1}{2})$ and $(\frac{1}{2}, 0)$, they describe the transformations of neutrino's. An electron for example is a spin 1/2 particle that is represented by $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ acting on a 4 dimensional spinor, since it needs to be symmetric in $j_1$ and $j_2$, because parity must be included. See [9] for clear details.

Importantly enough, these *finite* $(2j_1 + 1)(2j_2 + 1)$ dimensional representations of $SO(3,1)$ look unitary, because of the fact that all irreps of $SU(2)$ are. However there is one big difference, all generators of $SU(2)$ are hermitian, which does not hold for $SO(3,1)$ and was the reason for the *groups* $SO(3,1)$ *not* to be isomorphic to $SU(2) \times SU(2)$. So these finite dimensional irreps cannot be unitary, although they can be labeled by $(j_1, j_2)$, the labels of two unitary groups. The anti-Hermitian boost generators $Y_k$ in fact produce *anti*-unitary group elements, which implies that they cannot describe the Lorentz transformation of the particle states after all. We need infinite dimensional irreps to describe the transformation of the particle states. However all physical observables and particle fields *do* transform under the finite dimensional irreps of the Lorentz group [9].

The Hilbert space of states that furnish this finite dimensional representation

$$\{|j_A, j_B; m_A, m_B\rangle; m_A = -j_A, \ldots, j_A; m_B = j_B \ldots, j_B\} \equiv \{|m_A, m_B\rangle\} \tag{5.20}$$

can therefore be written as the direct product the representation spaces of the the $SU(2)$'s: $\{|j_A, m_A\rangle\}$ and $\{|j_B, m_B\rangle\}$. We can also choose as an equivalent basis

$$\{|j_0, j_1; j, m\rangle; j = j_0, \ldots, j_1; m = -j, \ldots, j\} \equiv |j, m\rangle, \tag{5.21}$$

with $j_0 = |j_A - j_B|$ and $j_1 = j_A + j_B$.

We can derive the representation for the Lorentz generators on the $\{|m_A, m_B\rangle\}$ basis by inverting equations(5.17) and (5.18). We get for the rotation generators $X_k$:

$$\begin{aligned} X_3|m_A, m_B\rangle &= (M_3 + N_3)|m_A, m_B\rangle = (m_A + m_B)|m_A, m_B\rangle \\ X_\pm|m_A, m_B\rangle &= (M_\pm + N_\pm)|m_A, m_B\rangle = [j_A(j_A+1) - m_A(m_A \pm 1)]^{1/2}|m_A \pm 1, m_B\rangle + \\ &\quad + [j_B(j_B+1) - m_B(m_B \pm 1)]^{1/2}|m_A, m_B \pm 1\rangle \end{aligned} \tag{5.22}$$

and for the boost generators $Y_k$ we get:

$$\begin{aligned} Y_3|m_A, m_B\rangle &= i(M_3 - N_3)|m_A, m_B\rangle = i(m_A - m_B)|m_A, m_B\rangle \\ Y_\pm|m_A, m_B\rangle &= i(M_\pm - N_\pm)|m_A, m_B\rangle = i[j_B(j_B+1) - m_B(m_B \pm 1)]^{1/2}|m_A, m_B \pm 1\rangle - \\ &\quad - i[j_A(j_A+1) - m_A(m_A \pm 1)]^{1/2}|m_A \pm 1, m_B\rangle \end{aligned} \tag{5.23}$$

Thereby defining finite dimensional non-unitary irreps of the Lorentz group, labeled by the two $SU(2)$ Casimir eigenvalues $(j_A, j_B)$. However since we are interested in elementary particle states in Hilbert space, we must *demand* that the representation is unitary.

## 5.2 Infinite dimensional unitary irreps of the Lorentz group

Since we want our particle states to transform under unitary irreps, something must give. We will start from the finite dimensional representations labeled by the eigenvalues $(j_A, j_B)$ of the $SU(2)$ Casimirs $\mathbf{M}^2$ and $\mathbf{N}^2$, but this time in the $|j_0, j_1; j, m\rangle \equiv |j, m\rangle$ basis. But now, since we want our irreps to be unitary, we wont restrict the values of the labels $(j_0, j_1)$ any longer. The demand for unitarity together with the Wigner-Eckart theorem lead to the following result, of which the extensive derivation can be found in [9]:

**Theorem 5.1.** *Unitary irreps for the Lorentz group*
There are two classes of unitary irreducible representations:

- The *Principal Series*: characterized by two parameters $(\nu, j_0)$, with $\nu = -i\omega; \omega$ real and $j_0 = 0, 1/2, 1, \ldots$.
- The *Complementary Series*: characterized by $(\nu, j_0)$, but now $-1 \leqslant \nu \leqslant 1$ and $j_0 = 0$.

The parameters $(\nu, j_0)$ are related to the eigenvalues $j_A(j_A+1)$ and $j_B(j_B+1)$ of the Casimir operators $\mathbf{M}^2$ and $\mathbf{N}^2$ respectively in the finite dimensional representation by

$$\begin{aligned} j_0 &= |j_A - j_B| \\ \nu &= j_A + j_B. \end{aligned} \tag{5.24}$$

We see however that instead of integer values like $j_1$ in the finite dimensional representation, $\nu$ takes purely imaginary values in the Principal series and continuous values between +1 and -1 in the Complementary series. These values are very different from the more familiar finite dimensional values and I found it difficult to interpret them as familiar physical quantities.

The rotation generators $X_k$ act on these vectors $|j_0, \nu; j, m\rangle \equiv |j, m\rangle$ in the canonical way, i.e. as SO(3) operators. But for the action of the boost operators $Y_k$, we need the Wigner-Eckhart theorem, which comes in two subtheorems, see [9] for an elaborate derivation and discussions:

**Theorem 5.2.** *Irreducible set of Operators*
If a set of operators $\{O_i^{(\mu)}; i = 1, \ldots, d_\mu\}$, acting on a vector space $V$ transforms under a symmetry group $G$ as

$$U(g) O_i^{(\mu)} U(g)^{-1} = O_i^{(\mu)} D^{(\mu)}(g)^j{}_i, \tag{5.25}$$

where $D^{(\mu)}(g)^j{}_i$ is an irreducible representation of dimension $d_\mu$, then the set $\{O_i^{(\mu)}\}$ is said to transform irreducibly under $G$.

So this is how operators transform irreducibly under a group G, instead of the normal basisvectors $|e_j^\nu\rangle$, which span the invariant vector space and transform according to

$$U(g)|e_j^\nu\rangle = |e_i^\nu\rangle D^{(\nu)}(g)^i{}_j \tag{5.26}$$

Then we look at how the combined system $O_i^{(\mu)}|e_j^\nu\rangle$ transforms under $U(g)$:

$$\begin{aligned} U(g) O_i^{(\mu)} |e_j^\nu\rangle &= U(g) O_i^{(\mu)} U^{-1}(g) U(g) |e_j^\nu\rangle \\ &= O_k^{(\mu)} |e_l^\nu\rangle D^{(\mu)}(g)^k{}_i D^{(\nu)}(g)^l{}_i \end{aligned} \tag{5.27}$$

were we see that it transforms as the direct product representation $D^{(\mu\times\nu)}$. Hence it can be decomposed using the Clebsch-Gordan decomposition and we end up with the following theorem [9]

**Theorem 5.3.** *Wigner-Eckart*
Let $\{O_i^{(\mu)}\}$ be a set of irreducible operators and $|e_j^\nu\rangle$ be a basis for the irrep $\nu$, then

$$\begin{aligned} \langle e_\lambda^l | O_i^{(\mu)} | e_j^\nu \rangle &= \sum_\alpha \langle \alpha, \lambda | l(\mu, \nu) | i, j \rangle n_\lambda^{-1} \sum_k \langle e_\lambda^k | w_{\alpha k}^\lambda \rangle \\ &\equiv \sum_\alpha \langle \alpha, \lambda | l(\mu, \nu) | i, j \rangle \langle \lambda | O^\mu | \nu \rangle_\alpha \end{aligned} \tag{5.28}$$

were $\langle \lambda | O^\mu | \nu \rangle_\alpha$ is the *reduced matrix element* and $\langle \alpha, \lambda | l(\mu, \nu) | i, j \rangle\!\rangle$ are the Clebsch-Gordan coefficients. Now applying this theorem to the Boost generators $Y_k$, which transform irreducibly like a 3-vector under rotations and hence form a spin 1 irrep as can be seen from the Lorentz algebra (5.16), we get

$$\begin{aligned} \langle j'm' | Y_3 | jm \rangle &= A^{j'}{}_j \langle j'm' | (1, j) | 0m \rangle \\ \langle j'm' | Y_\pm | jm \rangle &= \mp\sqrt{2} A^{j'}{}_j \langle j'm' | (1, j) | \pm 1 m \rangle \end{aligned} \tag{5.29}$$

were $A^{j'}{}_{j}$ correspond to the reduced matrix elements $\langle\lambda|O^{\mu}|\nu\rangle_{\alpha}$ defined in the Wigner-Eckart theorem 5.3. They depend on $\nu$ and $j_0$ defining the Principal and Complementary series in non-trivial ways, and for their explicit form and derivations I would refer to [9] for a clear explanation. The important thing here is that these irreps depend on these two Casimir labels.

## 5.3 Poincaré Algebra and the Casimir operators

The Lorentz group can easily be extended to the Poincaré group by including spacetime translations, with generators $P^{\mu}$. See for a detailed discussion [12], [13], [14]. As we saw in section 2.5 a general Poincaré transformation looks like

$$x^{\mu} \rightarrow x^{\mu'} = \Lambda^{\mu'}{}_{\nu} x^{\nu} + a^{\mu} \tag{5.30}$$

and we shall denote it by $T(\Lambda, a)$.

Two subsequent Poincaré transformations result in to the following composition rule for the Poincaré group

$$\begin{aligned} x^{\mu''} &= \Lambda^{\mu''}{}_{\mu'} x^{\mu'} + a^{\mu''} \\ &= \Lambda^{\mu''}{}_{\mu'} \left[ \Lambda^{\mu'}{}_{\mu} x^{\mu} + a^{\mu'} \right] + a^{\mu''} \\ &= \left[ \Lambda^{\mu''}{}_{\mu'} \Lambda^{\mu'}{}_{\mu} \right] x^{\mu} + \left[ \Lambda^{\mu''}{}_{\mu'} a^{\mu'} + a^{\mu''} \right] \end{aligned} \tag{5.31}$$

The Poincaré group composition law for is therefore

$$T(\Lambda'', a'')T(\Lambda', a') = T(\Lambda''\Lambda', \Lambda'' a' + a'') \tag{5.32}$$

and hence the inverse of $T(\Lambda, a)$ is $T(\Lambda^{-1}, -\Lambda^{-1}a)$, with components $(\Lambda^{-1})^{\rho}{}_{\sigma} \equiv \Lambda_{\sigma}{}^{\rho} = \eta_{\mu\sigma}\eta^{\nu\rho}\Lambda^{\mu}{}_{\nu}$.

Now, when dealing with quantum mechanics, any Poincaré transformation in spacetime $T(\Lambda, a)$ has a unitary representation of linear operators $U(\Lambda, a)$ acting on the Hilbert space of wavefunctions $\Psi$. Being a representation the operators satisfy the same composition law given by equation (5.32). To derive the Poincaré algebra, we look at infinitesimal transformations. The identity transformation is given by putting $\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu}$ and $a^{\mu} = 0$, i.e. $T(\mathbb{1}, 0)$ so for an infinitesimal translation $\Lambda^{\mu}{}_{\nu}$ and $a^{\mu}$ are given by

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu}, \qquad\qquad a^{\mu} = \epsilon^{\mu} \tag{5.33}$$

Applying this to a general Poincaré transformation we get

$$\Lambda^{\mu}{}_{\nu} x^{\nu} + a^{\mu} \rightarrow (\delta^{\mu}{}_{\nu} + \omega^{\mu}{}_{\nu}) x^{\nu} + \epsilon^{\mu} = \ = x^{\mu} + \omega^{\mu}{}_{\nu} x^{\nu} + \epsilon^{\mu} = x^{\mu} + V^{\mu}(x) \tag{5.34}$$

were we recognize the Killing vectors (2.58) derived from the Minkowski metric. So we're on the right track.

Using equation (5.10), which defines the Lorentz group and even more general the Poincaré group, we deduce that the infinitesimal spacetime rotation or Lorentz transformation $\omega_{\mu\nu}$ must be anti symmetric in its indices, i.e. $\omega_{\mu\nu} = -\omega_{\nu\mu}$. Together with

$\epsilon^\mu$ this leaves us with a total of 10 parameters, hence the Poincaré group is a 10 dimensional Lie group and we verify again that Minkowski space is a maximal symmetric spacetime for $D = 4$. The unitary operator $U(\mathbb{1} + \omega, \epsilon)$ corresponding to the infinitesimal transformation is given by the following equation of first order in its parameters

$$U(\mathbb{1} + \omega, \epsilon) = \mathbb{1} + \frac{1}{2} i\omega_{\mu\nu} J^{\mu\nu} - i\epsilon_\rho P^\rho + \dots \tag{5.35}$$

where $J^{\mu\nu}$ and $P^\rho$ of course are the generators, which will turn out to obey the commutation relations of the spacetime rotation and translation operators respectively and can thus be interpreted as such.

The result for the Poincaré algebra is given by the set of equations

$$[P_\mu, P_\nu] = 0 \tag{5.36}$$
$$[P_\mu, J_{\nu\lambda}] = i(P_\nu g_{\mu\lambda} - P_\lambda g_{\mu\nu}) \tag{5.37}$$
$$[J_{\mu\nu}, J_{\lambda\sigma}] = i(J_{\lambda\nu} g_{\mu\sigma} - J_{\sigma\nu} g_{\mu\lambda} + J_{\mu\lambda} g_{\nu\sigma} - J_{\mu\sigma} g_{\nu\lambda}) \tag{5.38}$$

from which we immediately deduce that the set of translation operators $\{P^\mu\} = T_4$ forms an abelian invariant subgroup.

It's not difficult to show [9], [12] that the group has two Casimir operators

$$C_1 \equiv -P_\mu P^\mu = P_0^2 - \mathbf{P}^2 = M^2 \tag{5.39}$$

and

$$C_2 \equiv W_\lambda W^\lambda. \tag{5.40}$$

$C_1$ commutes with $J_{\mu\nu}$, because it is a Lorentz scalar and $C_1$ commutes with $P^\mu$ because $T_4$ is abelian. $W^\lambda = \epsilon^{\lambda\mu\nu\sigma} J_{\mu\nu} P_\sigma$ is the *Pauli-Lubanski* vector and is related to the helicity $\lambda$. This can be seen [14] by first noting that $W_\mu P^\mu = 0$ holds and then

$$W^0 = \frac{\mathbf{WP}}{P^0} = \mathbf{JP} \sim \lambda \mathbf{p} \tag{5.41}$$

By definition these Casimir's label the irreps and especially $C_1$ is important, since it is the Casimir corresponding to the *abelian* invariant subgroup $T_4$. Which means that besides the eigenvalues of $C_1$, also the eigenvalues of the generators of translations $P_\mu$ can be used to label the so called *standard vectors*, since they are invariant under $T_4$, because of its abelian nature. The fact that $T_4$ is invariant is important for the method of induced representations, explained below.

## 5.4 Unitary irreps of the Poincaré group

For deriving the unitary irreps for the Poincaré group we use the method of induced representations:

### 5.4.1 Induced representation method in general

An induced representation [8] is a representation of a group $G$, that by induction is derived from a representation of a subgroup $H \subset G$. Let $H$ be a subgroup of group $G$. Then we can decompose $G$ formally in its distinct left cosets with respect to $H$, as

$$G = \sum_{i=1}^{|G|/|H|} g_i H \tag{5.42}$$

with $g_i$ being the coset representative, i.e. $g_1 = e$, $g_2$ is such that $g_2 H \cap g_1 H = \varnothing$ and $g_3 H \cap (g_1 H \cup g_2 H) = \varnothing$ etc., i.e. $i = 1, 2, \ldots |G|/|H|$. Now let $H$ have an irrep $D^{(\mu)}$ with dimension $d_\mu$ and let $|\mu m\rangle$ be the basis of the representation space of $D^{(\mu)}$, with $\mu$ simply labeling the representation and $m = 1, 2, \ldots, d_\mu$. Now define the acting of a coset representative $g_i$ on a basis vector $|\mu m\rangle$ as

$$g_i |\mu m\rangle \equiv |\mu m i\rangle. \tag{5.43}$$

Take any element $g \in G$ and $g_i$ as a coset representative, then for some $h \in H$, $g g_i$ can be expressed as

$$g g_i = g_j h \tag{5.44}$$

with $g_j$ as some other coset representative. This is true, because first of all $g g_i \in G$, because $G$ is a group. Secondly since the union of all cosets of $H$ must include all elements of $G$, there must be some $g_j$ for which equation (5.44) is true. Therefore for any $g \in G$

$$\begin{aligned} g|\mu m i\rangle &= g g_i |\mu m\rangle \\ &= g_j h |\mu m\rangle \\ &= g_j \sum_{m'} D^{(\mu)}_{m'm}(h) |\mu m'\rangle \\ &= \sum_{m'} D^{(\mu)}_{m'm}(h) |\mu m' j\rangle \end{aligned} \tag{5.45}$$

where the third equality just expresses the operator $h \in H$ in its $d_\mu$-dimensional irreducible basis $|\mu m\rangle$ as a matrix $D^{(\mu)}$, the irrep of $H$. This means that the $|G|/|H| \cdot d_\mu$ vectors $|\mu m i\rangle$ are closed under the group $G$ and therefore form a basis for a representation of $G$. This particular representation is called the *induced representation* (induced by $H$) and is sometimes denoted as $D^{(\mu)} \uparrow G$. It is in general reducible and can be decomposed into irreps of $H$:

$$D^{(\mu)} \uparrow G = \bigoplus_{\nu} a_\nu^{(\mu)} D^{(\nu)} \tag{5.46}$$

where $a_\nu^{(\mu)}$ is the multiplicity of the irrep $D^{(\nu)}$ in the induced representation $D^{(\mu)} \uparrow G$. So for any group $G$, with subgroup $H$ with known irreps $D^{(\mu)}$, representations of $G$ itself, in general reducible, can always be induced by letting coset representatives $g_i$ act on the basis $|\mu m\rangle$ of the representation space of the irreps $D^{(\mu)}$ of $H$, given by equation (5.43).

### 5.4.2 Induced representation method for Poincaré

To apply the general technique described above to the Poincaré group, we again mostly follow [9] but may also refer to [12] and the closely related and clear lecture notes [13] for an alternative and clear approach.

We can describe this method to derive the irreps of some group, which has an abelian invariant subgroup, in words by the following steps, directly applied to the Poincaré group, which I will call $\tilde{P}$

- Define an eigenvector of the generators of the abelian invariant subgroup $P^\mu$ to get a standard vector $|\mathbf{p_0}\rangle$. This can be done since, because of the abelian nature of $T_4$, the eigenvalues of $P^\mu$ are invariant under $T_4$ itself and in this way all possible eigenvalues of $P^\mu$ and $C_1$, defining the *orbits*, can be scanned systematically.
- Given some standard vector $|\mathbf{p_0}\rangle$, determine its *little group*: The set of generators from the factor group $\tilde{P}/T_4$ that leave standard vector invariant. Expand the standard vector by adding the eigenvalues of the generators of the little group.
- Following the discussion of section 5.4.1, every irrep of the little group induces a representation of the whole group by acting on the invariant subspace with the remaining generators of the factor group to construct new basis vectors. Now the fact that $T_4$ is an *invariant* subgroup is important, since this implies that one can act with one of the non-commuting operators from the factor group on $|\mathbf{p_0}\rangle$ and still be sure to create a new eigenvalue for $P_\mu$, thereby generating the rest of the invariant subspace.
- The basis constructed as such is an irrep because it is closed under all the generators of the group, hence no smaller invariant subspace exists.

The fact that $C_1$ is not positive definite, but is a Casimir operator splits the whole representation space into 6 *orbits* (see [12]) and their corresponding Casimir labels:

1. Null vector $m^2 = 0, p_0 = \mathbf{p} = 0$, little gp: full Lorentz $(j_0, \nu)$
2. Time-like $m^2 > 0, p_0 > 0$, little gp: SO(3) $(m, j)$
3. Light-like $m^2 = 0, \mathbf{p} \neq 0$, little gp: ISO(2) $(m = 0, \lambda)$
4. Space-like $m^2 < 0$, little gp: SO(2,1)
5. $m^2 = 0, p_0 < 0$, little gp: ISO(2)
6. $m^2 > 0, p_0 < 0$, little gp: SO(3)

of which only the first four orbits correspond to physically interesting cases, since negative energy particles are non-physical. We pick out the time-like orbit $m^2 > 0, p_0 > 0$ for further inspection: it can be represented by the standard vector $(m, 0)$ from which it is clear that the only generators from the factor group $\tilde{P}/T_4 \simeq SO(3,1)$ that leave this standard vector invariant are indeed the 3D rotations $SO(3)$, which therefore forms its little group. Now every irrep of $SO(3)$, labeled by $j$ (See section

4.11), induces an irrep of $\tilde{P}$. The corresponding standard vector can be labeled by $|j, m^2; \mathbf{p} = 0, \lambda\rangle$, where $\lambda = (-j \ldots, j)$ with the defining operator equations

$$P^\mu |j, m^2; \mathbf{p} = 0, \lambda\rangle = |j, m^2; \mathbf{p} = 0, \lambda\rangle p_t^\mu, \quad p_t^\mu = (m, \mathbf{0}) \tag{5.47}$$

$$\mathbf{J}^2 |j, m^2; \mathbf{p} = 0, \lambda\rangle = |j, m^2; \mathbf{p} = 0, \lambda\rangle j(j+1) \tag{5.48}$$

$$J_3 |j, m^2; \mathbf{p} = 0, \lambda\rangle = |j, m^2; \mathbf{p} = 0, \lambda\rangle \lambda \tag{5.49}$$

Now we can generate the whole irrep by acting with the remaining generators of the Lorentz group. Now since in general any Lorentz transformation $\Lambda$ can be decomposed [9] as $\Lambda = R(\alpha, \beta, 0) L_3(\zeta) R(\phi, \theta, \psi)^{-1}$, where $L_3(\zeta)$ is a pure Lorentz boost in the $z$-direction and $R$'s are pure rotations, which were both in the little group. Hence we first only need to focus on the pure boost $L_3(\zeta)$, since $R(\phi, \theta, \psi)^{-1}$ leaves the standard vector invariant per definition. We now define the action of $L_3(\zeta)$ on the standard vector as

$$L_3(\zeta) |j, m^2; \mathbf{p} = 0, \lambda\rangle \equiv |j, m^2; p\hat{\mathbf{z}}, \lambda\rangle, \tag{5.50}$$

with $p = m \sinh(\zeta)$. Again this is sensible, because $T_4$ is invariant and hence the vector $L_3(\zeta)|j, m^2; \mathbf{p} = 0, \lambda\rangle$ must represent some *new* eigenvalue for $P^\mu$. The action of the second rotation $R(\alpha, \beta, 0)$ on $|j, m^2; p\hat{\mathbf{z}}, \lambda\rangle$ is now defined to be

$$R(\alpha, \beta, 0) |j, m^2; p\hat{\mathbf{z}}, \lambda\rangle \equiv |j, m^2; \mathbf{p}, \lambda\rangle \tag{5.51}$$

Defining an operator $H(p) \equiv R(\alpha, \beta, 0) L_3(\zeta)$, we can write the vector $|j, m^2; \mathbf{p}, \lambda\rangle$ as

$$|j, m^2; \mathbf{p}, \lambda\rangle = H(p) |j, m^2; \mathbf{p} = 0, \lambda\rangle \tag{5.52}$$

indicating that $H(p)$ is the Lorentz transformation which transforms the standard vector into an arbitrary vector.

Having used all generators of the Poincaré group, we conclude that the basis vectors $|j, m^2; \mathbf{p}, \lambda\rangle$ span an invariant subspace of the Hilbert space and hence furnish an irrep labeled by the Casimir eigenvalues $(j, m)$ which for this massive elementary particle represented by $|j, m^2; \mathbf{p}, \lambda\rangle$ correspond to its mass and spin, which was what we expected. We now derived these general results using only group theoretical arguments. The same can be done (see [9], [12]) for all other orbits, to get the full classification for all elementary particles in Minkowski spacetime.

# 6 Conclusion

We tried to find a general recipe to classify the elementary particles in a given spacetime. Furthermore we wanted to know what the invariant physical observables are that characterize these elementary particles and how they in general can be found. We ended up with the following results:

- First we need to define what spacetime we are talking about by solving the modified Einstein equations(2.38) for some given value for $\Lambda$ to obtain the metric tensor $g_{\mu\nu}(x)$ .
- Then we want to determine the isometry group for this spacetime by solving the Killing equations (2.53) to find the Killing vector fields $V^\mu(x)$ which generate the isometry group.
- Then the elementary particles can be identified with the unitary irreducible representations (UIR's) of this isometry group.
- If the isometry group has an abelian invariant subgroup, the method induced representations can be used to find all unitary irreps. For simple and semi-simple groups, the classification of the irreps has been documented.
- De Casimir operators, constructed out of the isometry group generators, label the UIR's: for example $(j, m)$ for the massive particles in Minkowski as expected.

As was mentioned in the introduction, measurements show that the universe actually *is* curved [1],[2]. The universe turns out to be a solution to the modified Einstein equations(2.38) with $\Lambda > 0$, hence it is a de Sitter spacetime. Moschella [15] describes this space as an embedded space using the metric

$$-(x^0)^2 + x_i x^i = R^2,\ i = 1, \ldots, 4 \tag{6.1}$$

with $R = \sqrt{\frac{3}{\Lambda}}$ related to the cosmological constant. From this metric it is easy to see that the isometry group should be $SO(4, 1)$. Having 10 generators, this too is a maximal symmetric spacetime in $D = 4$. We can therefore try to compare the dS Casimirs with the Poincaré Casimirs to try to find some sensible statements about what we mean by an elementary particle having mass $m$ in dS spacetime.

Garidi [16] follows this path, and uses *contraction* of the generators: one rescales the generators of de Sitter isometry group and then let the scale parameter go to infinity, making the metric flat again, i.e. going from dS spacetime towards Minkowski spacetime. It would be interesting to look at the effect of this contraction operation on the dS Casimir operators and their eigenvalues. It might turn out for example that some of the "non physical" orbits of the Poincaré group could correspond to physically valid orbits in dS, resulting in the possible existence of new, previously disqualified, elementary particles.

# A (In)finite dimensional Vector Spaces and the concept of length

A vector space $V$ over a field $F$, in our case taken to be $\mathbb{C}$, is in general a set of entities that obeys the following eight axioms for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in F$

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
3. There exists a unique vector $\mathbf{0} \in V$ such that $\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$
4. For every $\mathbf{u} \in V$ there exists a vector $-\mathbf{u} \in V$, such that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
5. $a(b\mathbf{u}) = (ab)\mathbf{u}$
6. $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$
7. $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$
8. $1 \in F$ such that $1\mathbf{u} = \mathbf{u}$

This vector space $V$ can be anything, as long as it obeys the familiar axioms for a vector space. For finite dimensional vector spaces each individual vector $\mathbf{v} \in V$ can be expressed as a linear combination of a finite set of basis vectors $\{e_{(i)}\}$: $\mathbf{v} = v^i e_{(i)}$, where $v^i$ are the base-dependent *components* of the vector $\mathbf{v}$ in the particular basis $\{e_{(i)}\}$. The parenthesis around the index $i$ in $e_{(i)}$ implies the fact that $i$ indicates $i$ distinct *vectors*, not components thereof. The Cartesian product $\mathbb{R}_1 \times \mathbb{R}_2 \times \ldots \times \mathbb{R}_n = \{(x_1, x_2, \ldots, x_n) | x_i \in \mathbb{R}_i)\}$, better known as $\mathbb{R}^n$, is a common example of a finite dimensional vector space where the basis consists of the linear independent vectors $\{e_{(i)}\}$ with components $(e_{(i)})_j = \delta_{ij}$.
As another important example of a vector space, consider the *in*finite dimensional Hilbert space of continuous square integrable functions $L^2$. In general a Hilbert space is a vector space for which every Cauchy sequence converges to an element of that space. To understand this better we need to define a few more concepts:

**Definition A.1.** *Inner Product;* $(\mathbf{u}, \mathbf{v})$
The inner product $(\ ,\ )$, also known as *scalar* product, is a map that assigns an complex number to an ordered pair of vectors:

$$(\ ,\ ) : V \times V \to \mathbb{C} \tag{A.1}$$

and obeys the following rules for $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $a, b \in \mathbb{C}$:

1. (*Hermiticity*) $(\mathbf{u}, \mathbf{v}) = (\mathbf{v}, \mathbf{u})^*$
2. (*Linearity in the second vector*) $(\mathbf{w}, a\mathbf{v} + b\mathbf{u}) = a(\mathbf{w}, \mathbf{v}) + b(\mathbf{w}, \mathbf{u})$
3. (*Positivity*) $(\mathbf{v}, \mathbf{v}) \geqslant 0$

Sometimes a fourth axiom is added, but this simply is a corollary of 1 and 2:

4. (*Anti-Linearity in the first vector*) $(a\mathbf{w} + b\mathbf{v}, \mathbf{u}) = a^*(\mathbf{w}, \mathbf{u}) + b^*(\mathbf{v}, \mathbf{u})$

A vector space endowed with an inner product as additional structure is called an *inner product space*. One can easily prove that for any inner product space the following inequality, known as the *Cauchy-Schwarz inequality*, holds:

$$|(\mathbf{u}, \mathbf{v})|^2 \leqslant (\mathbf{u}, \mathbf{u}) \cdot (\mathbf{v}, \mathbf{v}) \tag{A.2}$$

The lefthand side of this equation is called the modulus (squared) of the inner product and for any complex number $z = a+bi$ this is taken to be $|z|^2 = z \cdot z^* = (a+bi)(a-bi) = a^2+b^2$. In a complex finite dimensional vector space we take the inner product between two vectors as

$$(\mathbf{u}, \mathbf{v}) \equiv \sum_i^n u_i^* v_i \tag{A.3}$$

In an infinite dimensional vector space of continuous functions $f$ on an interval $[a, b]$, we take the inner product to be

$$(f, g) = \int_a^b f^*(x) g(x) dx \tag{A.4}$$

An natural extension to this structure is the concept of a *norm*, which in turn can be used to define the concept of *length* by the definition of a *metric* in a vector space.

**Definition A.2.** *Norm of a vector* $\mathbf{v}$*;* $||\mathbf{v}||$
The norm maps a vector to a real number, i.e.

$$||\ ||: V \to \mathbb{R} \tag{A.5}$$

In general any norm must obey the following rules for $\mathbf{u}, \mathbf{v} \in V$ and $\lambda \in \mathbb{C}$:

1. $||\mathbf{v}|| \geqslant 0$, with equality iff $\mathbf{v} = 0$ (*Positivity*)
2. $||\lambda \cdot \mathbf{v}|| = |\lambda| \cdot ||\mathbf{v}||$ (*Positive scalability*)
3. $||\mathbf{v} + \mathbf{u}|| \leqslant ||\mathbf{v}|| + ||\mathbf{u}||$ (*Triangle inequality*)

In general any mapping obeying these rules would represent a norm, but in our case we will take the 2-norm or *Euclidean* norm given by

$$||\mathbf{v}|| = (\mathbf{v}, \mathbf{v})^{1/2} = \sqrt{(\mathbf{v}, \mathbf{v})}. \tag{A.6}$$

The norm of a vector is in this case therefore said to be *induced by the inner product* and a vector space endowed with a norm is likewise called a *normed vector space*. Another reason for choosing this norm is that the *parallelogram law* holds, which reads:

$$||\mathbf{u} + \mathbf{v}|| + ||\mathbf{u} - \mathbf{v}|| = 2||\mathbf{u}|| + 2||\mathbf{v}||. \tag{A.7}$$

Any norm satisfying this equality can be proven to be associated with or induced by an inner product.

This inequality can in turn be used to prove the triangle inequality, that together with positivity and positive scalability, which simply follow from the requirements on the inner product itself, prove that the induced norm indeed is a norm.

The modulus of a complex number is an example of a norm on the complex plain, after identifying a complex number with a vector in $\mathbb{R}^2$ and applying the usual Euclidean norm.

On the other hand, taking the vector space $\mathbb{R}^n$ and adding the inner product (A.3) and induced norm $\sqrt{(\mathbf{v},\mathbf{v})}$, we have the famous Euclidean space where the norm of a vector is equal to its Pythagorean length. This is a good example of how the norm naturally defines some concept of *length* in a vector space: Having defined an inner product and accompanying norm, we can identify the length of a vector simply with that norm and it will provide for a length which has all the requirements expected from a length, such as positivity. In other words, the length of vector $\mathbf{v}$ is equal to $||\mathbf{v}||$. Together with length, *distance* between two vectors $\mathbf{u},\mathbf{v} \in V$ is likewise defined by making use of the norm and introducing the *metric* $d(\ ,\ )$ as

$$d(\mathbf{u},\mathbf{v}) = ||\mathbf{u}-\mathbf{v}|| = \sqrt{(\mathbf{u}-\mathbf{v},\mathbf{u}-\mathbf{v})}. \tag{A.8}$$

One says that the metric is induced by the norm. Introducing the difference vector $\boldsymbol{\Delta} = \mathbf{u}-\mathbf{v}$, equation (A.8) can be can be rewritten as

$$d(\mathbf{u},\mathbf{v}) = ||\boldsymbol{\Delta}|| = \sqrt{(\boldsymbol{\Delta},\boldsymbol{\Delta})} \tag{A.9}$$

which, in infinitesimal form $\boldsymbol{\Delta} \to dx^\mu$, this metric is often written as

$$ds^2 = (dx^\mu, dx^\mu) \equiv g_{\mu\nu} dx^\mu dx^\nu \tag{A.10}$$

where $g_{\mu\nu}$ obviously is defined by the inner product in the particular space. For 3D Euclidean space for example

$$ds^2 = dx^2 + dy^2 + dz^2 \tag{A.11}$$

and hence $g_{\mu\nu} = \mathrm{diag}(1,1,1)$. For 4D Minkowski space the inner product is however defined as

$$g_{\mu\nu}x^\mu x^\nu = x_\mu x^\mu = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 \tag{A.12}$$

and hence $g_{\mu\nu} = \mathrm{diag}(-1,1,1,1)$. We see here that the metric is not positive definite any more. A vector space with such a metric is called a pseudo-Riemannian mannifold[4], whereas a vector space with a positive definite metric is called Riemannian. Therefore ultimately defining an inner product on a vector space, which induces a norm, which in turn induces a metric, creates the notions of length and distance in that vector space by identifying the norm with the metric. But there is more: by introducing an inner product *angles* between vectors can also be defined, ultimately using the Cauchy-Schwarz inequality (A.2). Writing this equation in terms of norms we get:

$$|(\mathbf{u},\mathbf{v})| \leqslant ||\mathbf{u}|| \cdot ||\mathbf{v}|| \tag{A.13}$$

[4]A mannifold is a topological space which tangent space resembles Euclidean space. Topological means that ideas like continuity and connectedness are defined for that space.

or

$$\frac{|(\mathbf{u},\mathbf{v})|}{||\mathbf{u}||\cdot||\mathbf{v}||}\leqslant 1 \qquad \Leftrightarrow \qquad -1\leqslant\frac{(\mathbf{u},\mathbf{v})}{||\mathbf{u}||\cdot||\mathbf{v}||}\leqslant 1 \tag{A.14}$$

By identifying $\frac{(\mathbf{u},\mathbf{v})}{||\mathbf{u}||\cdot||\mathbf{v}||}$ with $\cos(\theta)$, we have a measure for the angle $\theta$ between two vectors $\mathbf{u}$ and $\mathbf{v}$:

$$\cos(\theta)=\frac{(\mathbf{u},\mathbf{v})}{||\mathbf{u}||\cdot||\mathbf{v}||} \tag{A.15}$$

**Definition A.3.** *Orthonormal Basis*
An orthonormal basis is a set of linearly independent vectors $\{e_{(i)}\}$ of unit norm that span the space:

$$(e_{(i)},e_{(j)})=\delta_{ij} \tag{A.16}$$

with $\delta_{ij}=\begin{cases}1 & \text{iff } i=j\\ 0 & \text{otherwise}\end{cases}$, being the Kronecker delta.

If a chosen basis is not orthonormal, we can always construct an orthonormal basis from it by applying the well known *Gram-Schmidt orthogonalisation* technique. In the text every basis is considered to be properly normalized by this procedure.

# References

[1] A.G. Riess, *et al.*, *Observational Evidence from Supernovea For An Accelerating Universe And A Cosmological Constant.* (Astron.J.116:1009-1038,1998)

[2] S. Perlmutter, *et al.*, *Discovery Of Supernova At Half The Age Of The Universe And Its Cosmological Implications.* (LBNL, Berkeley, CA(US), 1997)

[3] D. Tong, *Classical Dynamics* [PDF Document]. (Lecture notes retrieved from http://www.damtp.cam.ac.uk/user/tong/dynamics/clas.pdf)

[4] S.T. Thornton, J.B. Marion, *Classical Dynamics Of Particles And Systems*, 5th edition. (Brooks/Cole, 2004)

[5] S.M. Carrol, *Lecture Notes on General Relativity* [PDF Document]. (University of California, Santa Barbara, 1997)

[6] M.Blau, *Lecture Notes on General Relativity* [PDF Document]. (Universitt Bern, Switzerland, 2012)

[7] A.O. Barut and R. Raczka, *Theory Of Group Representations And Applications.* (World Scientific Publishing Co. Pte. Ltd., Singapore, 1986)

[8] M.A. Naimark and A.I. Stern, *Theory Of Group Representations.* (Springer-Verlag New York Inc., 1982)

[9] W. Tung, *Group Theory In Physics.* (World Scientific Publishing Co. Pte. Ltd., Singapore, 1985)

[10] H.F. Jones, *Groups, Representations And Physics*, 2nd edition. (Taylor and Francis group, New York, 1998)

[11] J.M. Evans, *PQM Supplementary Notes: Spin, Topology,* $SU(2) \rightarrow SO(3)$ *etc.* (Lecture notes found at http://www.damtp.cam.ac.uk/user/examples/)

[12] S. Weinberg, *The Quantum Theory Of Fields; Volume I.* (Press Syndicate of the University of Cambridge, 1995)

[13] X. Bekaert and N. Boulanger, *The Unitary Representations Of The Poincaré Group In Any Spacetime Dimension.* (arXiv: hep-th/0611263v1, 2006)

[14] B. Bakker, D. Boer, *Group Theory*, 8th edition. (Lecture notes, Vrije Universiteit, Amsterdam, 2008)

[15] U. Moschella, *The De Sitter And Anti-de Sitter Sightseeing Tour.* (Università dell'Insubria, Milano, 2005)

[16] T. Garidi, *What Is Mass In Desitterian Physics?.* (Université Paris, Paris, 2008)

[17] B.C. Hall, *An Elementary Introduction To Groups And Representations.* (University of Notre Dame, 2000)