Quantum Searches for Fundamental Scalars

Usama Aqeel

February 2024

Contents

1 Introduction **3**

1.1 Prerequisites and Basic Formulae . . . 3

1.1.1 Mathematical Notation . . . 3

1.1.2 General Relativity . . . 3

2 Quantum Field Theory **5**

2.1 Klein-Gordon Field . . . 5

2.1.1 K . . . 5

2.2 The Dirac Field . . . 5

2.2.1 Lorentz Invariance in Klein-Gordon . . . 5

3 Symmetries are cool but breaking them is better **8**

3.1 Groups and their representations . . . 8

3.1.1 Definition of a Group . . . 8

3.1.2 Example- Z_3 . . . 9

3.1.3 The regular representation . . . 9

3.1.4 Homomorphisms . . . 10

3.1.5 Irreducible Representations . . . 10

3.1.6 Example: $\mathbb{Z}_2$ . . . 12

3.2 Schur's Lemma . . . 14

3.2.1 Proof of Schur's Lemma . . . 14

3.3 Building Representations . . . 15

3.3.1 Dual and Conjugate Representation . . . 15

3.3.2 Tensor Product Representations . . . 16

3.3.3 Unitary Representations . . . 17

3.4 Representations of $SU(2)$ . . . 19

3.5 Young Tableu and $SU(n)$ irreps . . . 23

3.5.1 Hook length and dimensions of irreps . . . 25

3.5.2 Dual and Conjugate representations . . . 26

3.5.3 Tensor products of irreps . . . 26

4 Particle Physics and Symmetries **28**

4.1 The Poincaré Group . . . 28

4.1.1 Representations of the Poincaré group . . . 28
4.2 Quarks, leptons and $SU(3) \times SU(2) \times U(1)$ . . . 34
4.3 Breaking discrete symmetries . . . 37
4.4 Spontaneously Broken Continuous Symmetries . . . 37
4.5 Scalar Mass Matrix . . . 38
4.6 Broken Generators and Goldstone's Theorem . . . 39
4.6.1 Mass matrix for continuous symmetries . . . 41
4.7 Higgs Mechanism . . . 42
4.7.1 Abelian Higgs Model . . . 42
4.7.2 Non-Abelian Higgs . . . 44
4.8 Electroweak Symmetry Breaking and the Higgs . . . 45

5 Perturbation Theory in Gravity 46
5.1 Linearized Gravity . . . 46
5.2 Classifying the irreducible representations of the proper orthochronous Lorentz Group 48
5.3 Dynamical Degrees of Freedom . . . 49
5.3.1 Brief aside: Degrees of freedom in EM . . . 50
5.4 Back to linearized gravity . . . 52
5.4.1 Choices of gauge . . . 55
5.5 Newtonian Fields and Photon Trajectories . . . 56

6 Alternative Theories of Gravity 59
6.1 Scalar-Tensor Theories . . . 59
6.2 Brans-Dicke Theory . . . 61
6.2.1 Using Conformal Transformations . . . 62
6.3 Linearised Scalar Tensor Theories . . . 63

Appendices 65

Appendix A Shortcuts to calculate objects in General Relativity 66
A.1 Christoffel Symbols . . . 66
A.1.1 Calculating the Christoffel Symbols . . . 66

Appendix B Conformal Transformations 71

Appendix C Fifth forces in alternative theories of gravity 73
C.1 Dynamics in conformal metrics . . . 73
C.2 Conformal transformations on Connections . . . 74
C.3 Considering forces in Local Inertial Frames . . . 75
C.4 Calculating the Fifth Force . . . 76

Chapter 1

Introduction

1.1 Prerequisites and Basic Formulae

1.1.1 Mathematical Notation

- $\equiv$ - is used a lot of example calculations, this symbols means that this is the definition of the thing.
-

1.1.2 General Relativity

Metrics

- $\eta_{\mu\nu}$ - is the **Minkowski metric**. We use the *mostly-plus* signature, $\eta = \text{diag}(-1, 1, 1, 1)$, which we use.
- $g_{\mu\nu}$ - is used for the metric of a Lorentzian spacetime.
- $\partial_\mu = \frac{\partial}{\partial x^\mu}$ - denotes a **directional derivative** or the usual derivative. It can be thought of as the higher-dimensional analogue of the *gradient operator* in vector calculus. Often we will use shorthand such as $\omega_{\mu,\rho} = \partial_\rho \omega_\mu$. So something more complicated like, $\partial_\sigma \partial_\rho T^{\xi}{}_{\tau}{}^{\chi} = T^{\xi}{}_{\tau}{}^{\chi}{}_{,\sigma\rho}$.
- Christoffel Connection ($\Gamma^{\alpha}{}_{\mu\nu}$) are a system of numbers which are constructed from the derivatives of the metric,
$$\Gamma^{\alpha}{}_{\mu\nu} \equiv \frac{1}{2} g^{\alpha\beta} \left(\partial_\mu g_{\nu\beta} + \partial_\nu g_{\mu\beta} - \partial_\beta g_{\mu\nu} \right). \tag{1.1.1}$$
The Christoffel Connection allows us to construct a *covariant derivative*, as well as other covariant properties of manifolds.
- ∇_μ - will denote the *covariant derivative* of a tensor field. For (pseudo)-Riemannian manifolds, the covariant derivative on a vector field is written,
$$\nabla_\mu V^\rho = \partial_\mu V^\rho + \Gamma^{\rho}{}_{\mu\sigma} V^\sigma \tag{1.1.2}$$

On scalar fields, $\lambda(x)$, the covariant derivative takes on a particularly attractive form, $\nabla_\mu \lambda(x) = \partial_\mu \lambda(x)$. On one-form (covector) fields, the Christoffel term has a negative sign,

$$\nabla_\mu \omega_\nu = \partial_\mu \omega_\nu - \Gamma^\kappa{}_{\mu\nu}\omega_\kappa \tag{1.1.3}$$

Generally, on (n,m)-tensors, the covariant derivative generates and $(n,m+1)$-tensor:

$$\begin{aligned}\nabla_\rho T^{\mu_1\ldots\mu_n}_{\nu_1\ldots\nu_m} = \partial_\rho T^{\mu_1\ldots\mu_n}_{\nu_1\ldots\nu_m} &+ \Gamma^{\mu_1}{}_{\rho\kappa} T^{\kappa\ldots\mu_n}_{\nu_1\ldots\nu_m} + \cdots + \Gamma^{\mu_n}{}_{\rho\kappa} T^{\mu_1\ldots\kappa}_{\nu_1\ldots\nu_m} \\ &- \Gamma^{\kappa}{}_{\rho\nu_1} T^{\mu_1\ldots\mu_n}_{\kappa\ldots\nu_m} - \cdots - \Gamma^{\kappa}{}_{\rho\nu_m} T^{\mu_1\ldots\mu_n}_{\nu_1\ldots\kappa}.\end{aligned} \tag{1.1.4}$$

The rule to remember for the Christoffel terms is that Γ sums over every index separately, with a $+$ sign for *upstairs* indices, and a $-$ sign for *downstairs* indices. However, there is a unique tensor, which the covariant derivative has a special effect on:

$$\nabla_\alpha g_{\mu\nu} \equiv 0 \tag{1.1.5}$$

This is actually not a huge surprise as it comes from the definition of the $\Gamma^\rho{}_{\mu\nu}$. (1.1.5) is the ***condition of metric-compatibility***.

Chapter 2

Quantum Field Theory

2.1 Klein-Gordon Field

2.1.1 K

Here we consider the simplest of the quantum field theories known as the Klein-Gordon scalar field. Consider the action of a scalar field in Minkowski spacetime

$$S[\phi] = \int d^4x \left(\frac{1}{2}\partial_\mu \phi \partial^\mu \phi - \frac{1}{2} m^2 \phi^2 \right) = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi) \tag{2.1.1}$$

where $\phi = \phi(x)$ is a scalar function of spacetime coordinates, $x = (t, \vec{x})$.

2.2 The Dirac Field

2.2.1 Lorentz Invariance in Klein-Gordon

A Lorentz transformation on a vector is given by,

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{2.2.1}$$

for some $\Lambda^\mu{}_\nu$. How does the Klein-Gordon field transform under this transformation? Think of $\phi(x)$ as measuring the local value of some quantity at $x = \Lambda x_0$. The corresponding transformation is

$$\phi(x) \to \phi'(x) = \phi(\Lambda^{-1} x),$$

The transformed field, evaluated at the boosted point, gives the same value as the original field evaluated at the point before boosting. We see that this leaves Klein-Gordon invariant:

$$\partial_\mu \phi(x) \to \partial_\mu (\phi(\Lambda^{-1} x)) = (\Lambda^{-1})^\nu{}_\mu (\partial_\nu \phi)(\Lambda^{-1} x) \tag{2.2.2}$$

Since the metric, $\eta^{\mu\nu}$ is Lorentz invariant,

$$(\Lambda^{-1})^{\rho}{}_{\mu}(\Lambda^{-1})^{\sigma}{}_{\nu}\eta^{\mu\nu} = \eta^{\rho\sigma}. \tag{2.2.3}$$

therefore,

$$\begin{aligned}(\partial_\mu\phi(x))^2 &\to \eta^{\mu\nu}(\partial_\mu\phi'(x))(\partial_\nu\phi'(x)) \\ &= \eta^{\mu\nu}(\Lambda^{-1})^{\rho}{}_{\mu}(\Lambda^{-1})^{\sigma}{}_{\nu}\partial_\rho\phi(\Lambda^{-1}x)\partial_\sigma\phi(\Lambda^{-1}x) \\ &= \eta^{\rho\sigma}\partial_\rho\phi(\Lambda^{-1}x)\partial_\sigma\phi(\Lambda^{-1}x) \\ &= (\partial_\rho\phi)^2(\Lambda^{-1}x)\end{aligned} \tag{2.2.4}$$

Thus the whole Lagrangian is left invariant $\mathcal{L}(x) \to \mathcal{L}(\Lambda^{-1}x)$. As we see, that Lagrangian is a Lorentz scalar field. The equation of motion $(\Box + m^2)\phi(x) = 0$ is also Lorentz invariant.

The transformation law for $\phi(x)$ is the simplest possible one. There are other examples of fields transforming in more complicated ways, such as vectors. For the vector potential, $A^\mu(x)$, a quantity distributed in space also carries an *orientation*, which must be rotated or boosted. The orientation must be forward as the point of evaluation of the field is changed!

- under 3-dimensional rotation $\to v^i(x) \to R_{ij}v^j(R^{-1}x)$,
- under Lorentz transformation $\to v^\mu(x) \to \Lambda^{\mu}{}_{\nu}v^\nu(\Lambda^{-1}x)$

We can use tensor products to build tensors of arbitrary rank by simply adding more indices and more factors of Λ. Using such tensor fields we can write a variety of Lorentz invariant equations, e.g Maxwell theory,

$$\partial_\mu F^{\mu\nu} = 0; \qquad \Box A_\nu - \partial_\nu\partial^\mu A_\mu = 0 \tag{2.2.5}$$

which follow from the Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}; \qquad F = \mathrm{d}A \tag{2.2.6}$$

How do we classify all of the possible tensor product representations of the Lorentz group, instead of guessing. If Φ_a is an $n \times n$-component multiplet, the Lorentz transformation law is given by an $n \times n$ matrix $M(\Lambda)$ is,

$$\Phi_a \to M_{ab}(\Lambda)\Phi_b(\Lambda^{-1}x) \tag{2.2.7}$$

It can be shown that the most general linear transformation laws can be built from these linear transformations, so there is no advantage in considering transformations more general than this. We suppress the change in field. Writing instead,

$$\Phi \to M(\Lambda)\Phi \tag{2.2.8}$$

What are the possible forms of

Chapter 3

Symmetries are cool but breaking them is better

In this chapter we focus on the theory of behind *spontaneous symmetry breaking* building up to the Higgs Mechanism.

3.1 Groups and their representations

3.1.1 Definition of a Group

We begin with a gentle primer group theory. Definition: A group in mathematics is a set along with some composition law $(G, \cdot\ )$. Under the composition law, the elements of the group obey the following axioms:

1. (*Completeness*) If $a \cdot b = c \in G$ if $a, b \in G$.
2. (*Identity*) There exists a unique element $e \in G$ such that, for any element g, $g \cdot e = e \cdot g = g$.
3. (Inverse) For every element, g, there exists a multiplicative inverse $g^{-1} \cdot g = g \cdot g^{-1} = e$
4. (*Associativity*) The group composition law is associative, if $a, b, c \in G$, then $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

For **abelian** (commutative) groups there is an additional axiom:

5. (*Commutivity*) For any two elements $a, b \in G$ we have $a \cdot b = b \cdot a$.

A **representation** of G is a mapping, D of the elements of G onto a set of linear operators with the following properties:

1. $D(e) = 1$, where 1 is the identity operator in the space on which the linear operators act.
2. $D(g_1)D(g2) = D(g_1 g_2)$, in other words the group multiplication law is mapped onto the natural multiplication in the linear space on which linear operators act.

$\mathbb{Z}_3$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

Table 3.1: The group multiplication table for $\mathbb{Z}_3$.

3.1.2 Example- Z_3

A group is **finite** if it has a finite number of elements. Otherwise it is **infinite**. The number of elements in a finite group G is called the order of G. Here is a finite group of order 3 shown in (**??**),

This is the multiplication table for $\mathbb{Z}_3$, the *cyclic group of order 3*. Each element of the group appears in each row and column only once. This is typical of group multiplication because the inverse of each element exists. We see that the cyclic group is Abelian.

The following is a representation for $\mathbb{Z}_3$,

$$D(e) = 1, \quad D(a) = e^{2\pi i/3}, \quad D(b) = e^{4\pi i/3} \tag{3.1.1}$$

The dimension of the representation is the dimension of the space it acts on, for example, the above representation is a 1-dimensional representation.

3.1.3 The regular representation

Another representation of $\mathbb{Z}_3$ is:

$$D(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D(a) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad D(b) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \tag{3.1.1}$$

This representation was constructed directly from the multiplication table by the following trick. take the group elements themselves to form an orthonormal basis for a vector space, $|e\rangle$, $|a\rangle$ and $|b\rangle$. Now define,

$$D(g_1)\,|g_2\rangle = |g_1 g_2\rangle\,, \tag{3.1.2}$$

Which can be seen as a representation. It is called the *regular representation*. Clearly, the dimension of the regular representation is the order of the group. The matrix elements are thus calculated from the following construction,

$$\begin{aligned} |e_1\rangle \equiv |e\rangle\,, \quad |e_2\rangle \equiv |a\rangle\,, \quad |e_3\rangle \equiv |b\rangle \\ [D(g)]_{ij} = \langle e_i|\, D(g)\, |e_j\rangle \end{aligned} \tag{3.1.3}$$

This works for any representation, not just the regular one. The basic idea is just the insertion of a complete set of intermediate states. The matrix corresponding to a product of operators is the matrix product of the matrices corresponding to the operators,

$$\begin{aligned}
[D(g_1g_2)]_{ij} &= [D(g_1)D(g_2)]_{ij} \\
&= \langle e_i| D(g_1)D(g_2) |e_j\rangle \\
&= \sum_k \langle e_i| D(g_1) |e_k\rangle \langle e_k| D(g_2) |e_j\rangle \\
&= \sum_k [D(g_1)]_{ik}[D(g_2)]_{kj}
\end{aligned} \tag{3.1.4}$$

Note that the construction of the regular representation is completely general for any finite group. For any finite group, we can define a vector space in which the basis vectors are labeled by the group elements. Then (3.1.2) defines the regular representation.

3.1.4 Homomorphisms

What can we do with group? We can define some kind of function to bend the group into different shapes, or to represent the group in a different form. The way we do this is using group homomorphisms which are a type of map that preserve relationships between elements. They are particularly of interest in the representation of a group. For example, let's define a map ρ, from the group elements to the real matrices of dimension n,

$$\begin{aligned}
\rho : G &\to GL(n, \mathbb{R}) \\
g &\mapsto \rho(g) = A
\end{aligned} \tag{3.1.1}$$

Suppose that under the map, there was some difference between the composition law in G and in $\rho(G)$, we will denote the composition law in $\rho(G)$ by $\times$. Given that $a \cdot b = c$, under the homomorphism, we must have $\rho(a) \times \rho(b) = \rho(a \cdot b) = \rho(c)$. So we can see that the in order for the homomorphism to preserve the group product, we need $\rho(a) \times \rho(b) = \rho(a \cdot b)$ for all $a, b \in G$

3.1.5 Irreducible Representations

What makes the idea of group representations so powerful is the fact that they live in linear spaces. And the wonderful thing about linear spaces is we are free to choose to represent the states in more convenient way by making linear transformations. As long as the transformations are invertible, the new states are as good as the old ones. Such a transformation on the states produces a *similarity transformation* on the linear space, so that we can always make a new representation of the form,

$$D(g) \to D'(g) = S^{-1}D(g)S \tag{3.1.1}$$

Because of the form of the similarity transformation, the new set of operators has the same multiplication rules as the old ones, for example, if $g_1, g_2 \in G$, then,

$$\begin{aligned} D'(g_1)D'(g_2) &= (S^{-1}D(g_1)S)(S^{-1}D(g_2)S) \\ &= S^{-1}D(g_1)(SS^{-1})D(g_2)S \\ &= S^{-1}D(g_1)D(g_2)S \\ &= S^{-1}D(g_1g_2)S = D'(g_1g_2) \end{aligned} \tag{3.1.2}$$

which suffices as a representation. D and D' are said to be equivalent up to choice of basis, $D \sim D'$.

Unitary operators (O such that $O^\dagger = O^{-1}$) are particularly important in Quantum Mechanics. A representation is is unitary if all the $D(g)$s are unitary. Both the representations we have discussed are unitary. It turns out that all representations of finite groups are equivalent to unitary representations.

A representation is *reducible* if it has an *invariant subspace* which means that the action of any $D(g)$ on any vector in the subspace is still in the subspace. In terms of a projection operator P onto the subspace this condition can be written as,

$$PD(g)P = D(g)P \quad \text{for all } g \in G \tag{3.1.3}$$

For examples, the regular representation of $\mathbb{Z}_3$ has an invariant subspace projected by,

$$P = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{3.1.4}$$

because $D(g)P = P$ for all g. The restriction of the representation of the invariant subspace is itself a representation. In this case, it is the *trivial representation* for which, $D(g) = 1$ (the trivial representation $D(g) = 1$, is always a representation - every group has one).

- A representation is *irreducible* if it is not reducible
- A representation is *completely reducible* is it is equivalent to a representation whose matrix has the following form

$$\begin{pmatrix} \boxed{D_1(g)} & & \cdots & \\ & \boxed{D_2(g)} & \cdots & \\ \vdots & & \ddots & \vdots \\ & \cdots & & \boxed{D_d(g)} \end{pmatrix} \tag{3.1.5}$$

where the $D_i(g)$ is an irreducible block. This is called *block diagonal form.*

A representation in block diagonal form is said to be the *direct sum* of the subrepresentations, $D_j(g)$,

$$D_1 \oplus D_2 \oplus ... \oplus D_d \tag{3.1.6}$$

$\mathbb{Z}_2$	e	a
e	e	a
a	a	e

Table 3.2: The group multiplication table for $\mathbb{Z}_2$.

In transforming a representation to block diagonal form, we are decomposing the original representation into a direct sum of its irreducible representations (*irreps*). Another way of defining complete reducibility is to say that a completely reducible representation can be decomposed into a direct sum of irreps.

Consider the $\mathbb{Z}_3$ representation (3.1.1) with a similarity transformation,

$$S = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{pmatrix} \tag{3.1.7}$$

where $\omega = e^{2\pi i/3}$. Then,

$$D'(e) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad D'(a) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega & 0 \\ 0 & 0 & \omega^2 \end{pmatrix}, \quad D'(b) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \omega^2 & 0 \\ 0 & 0 & \omega \end{pmatrix} \tag{3.1.8}$$

which clearly shows that this representation is reduced.

3.1.6 Example: $\mathbb{Z}_2$

Consider the cyclic group of order 2, $\mathbb{Z}_2$.

Here are a few $\mathbb{Z}_2$ representations:

1. $\rho_0(e) = 1$ and $\rho_0(a) = 1$ which is the trivial representation.
2. $\rho_1(e) = 1$ and $\rho_1(a) = -1$
3. $\rho_2(e) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\rho_2(a) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$
4. $\rho_4(e) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\rho_4(a) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

- Notice that ρ_2 is just a combination of ρ_1 on $\begin{pmatrix} x \\ 0 \end{pmatrix}$ and ρ_1 acting on $\begin{pmatrix} 0 \\ y \end{pmatrix}$
- Also notice that ρ_4 is really a ρ_0 acting on $\begin{pmatrix} x \\ 0 \end{pmatrix}$ and ρ_1 acting on $\begin{pmatrix} 0 \\ y \end{pmatrix}$

Alternatively, consider the $U(1)$ representation,

$$\rho'(\theta) = \begin{pmatrix} \cos n\theta & \sin n\theta & 0 \\ -\sin n\theta & \cos n\theta & 0 \\ 0 & 0 & e^{i\theta} \end{pmatrix} \tag{3.1.1}$$

Now we have three decompositions,

1. $\begin{pmatrix} a \\ ia \\ 0 \end{pmatrix} \mapsto e^{in\theta} \begin{pmatrix} a \\ ia \\ 0 \end{pmatrix}$

2. $\begin{pmatrix} a \\ -ia \\ 0 \end{pmatrix} \mapsto e^{-in\theta} \begin{pmatrix} a \\ -ia \\ 0 \end{pmatrix}$

3. $\begin{pmatrix} 0 \\ 0 \\ b \end{pmatrix} \mapsto e^{i\theta} \begin{pmatrix} 0 \\ 0 \\ b \end{pmatrix}$

A *left-invariant G-module* (and representation) is reducible if there exists a proper, invariant subspace $W \subset V$ that is

- $a \cdot w \in W$ for $a \in G$, $w \in W$ (invariant)
- and $W \neq V$ and $W \neq 0$ (proper).

If V is reducible, then $W \subset V$ forms a *left-invariant G-module*.

In terms of representation, matrices $\rho(a)$ is a representation is reducible, it can be put into triangular form. We can find a matrix S, such that

$$S^{-1}\rho(a)S = \begin{pmatrix} \hat{\rho}(a) & A(a) \\ 0 & B(a) \end{pmatrix} \tag{3.1.2}$$

where, $\hat{\rho}$, A and B are matrices. W is spanned by vectors $\begin{pmatrix} w \\ 0 \end{pmatrix}$, then $\hat{\rho}(a)$ is an *induced representation* on W.

A *decomposible* left G-module V is one such that $V = W_1 \oplus W_2$ where W_1 and W_2 are each proper, invariant subspaces.

We then have for the representation

$$\rho = \rho_1 \oplus \rho_2 \tag{3.1.3}$$

where ρ_i are the induced representations on W_1. There exists and S such that

$$S^{-1}\rho(a)S = \begin{pmatrix} \rho_1(a) & \\ & \rho_2(a) \end{pmatrix} \tag{3.1.4}$$

with $\begin{pmatrix} v_1 \\ 0 \end{pmatrix}$ spanning W_1 and $\begin{pmatrix} 0 \\ v_2 \end{pmatrix}$ spanning W_2

Generally, we can write

$$\begin{aligned} V &= W_1 \oplus W_2 \oplus \ldots \oplus W_n \\ \rho &= \rho_1 \oplus \rho_2 \oplus \ldots \oplus \rho_n \end{aligned} \tag{3.1.5}$$

In our $\mathbb{Z}_2$ example, $\rho_2 = \rho_1 \oplus \rho_1$ and $\rho_4 = \rho_0 \oplus \rho_1$, while the $U(1)$ rep can be built up as $\rho' = \rho_{(n)} \oplus \rho_{(-n)} \oplus \rho_{(1)}$

An *irreducible (simple) G-module* (or irredcible representation) is one which has no proper, invariant subspaces.

- We can build general representations out of irreducible building blocks.
- Each group has many (usually infinite) numbers of different irreps.

3.2 Schur's Lemma

Let V and W be irreducible G-modules and $f : V \to W$ be a homomorphism, that is:

$$f(av) = af(v) \text{ for } a \in G, v \in V \tag{3.2.1}$$

then either,

- f is invertible, thus V and W define equivalent representations ($V \sim W$)
- $f = 0$

3.2.1 Proof of Schur's Lemma

There exist two invariant subspaces under the map f,

- $\ker f = \{v \in V : \; f(v) = 0\}$ (kernel). If $v \in \ker f$ then $f(av) = af(v) = 0$ then $av \in \ker f$.
- $\text{im } f = \{w \in W : \exists\, v \in V : f(v) = W\}$ (image). If $w \in \text{im} f$ then $aw = af(v) = f(av)$ then $aw \in \text{im} f$.

This shows that $\ker f$ and im f are both invariant subspaces. Since V and W are irreducible, either

- ($\ker f = 0$ or $\ker f = V$) and (im $f = 0$ or im $f = W$)

Thus we have two possibilities

(a) $\ker f = 0$, im $f = W$, thus f is invertible and f is an isomorphism.

(b) $\ker f = V$, im $f = 0$, this $f = 0$

Corollary to Schur's Lemma

Let V and W be a finite-dimensional G-module and $f : V \to V$ be a homomorphism (endomorphism) then f is proportional to the identity.

$$f = \lambda \mathbb{I} \quad (\lambda \in \mathbb{C}) \tag{3.2.1}$$

For the proof, we take f to be a square matrix. Let λ be some eigenvalue of f and consider $f - \lambda\mathbb{I}$:

(a) $\det(f - \lambda\mathbb{I}) = 0$ so, it is not invertibl

(b) $f - \lambda\mathbb{I}$ us a homomorphism therefore $f - \lambda\mathbb{I} = 0$, hence by Schur's Lemma $f = \lambda\mathbb{I}$

This corollary implies that *all irreps of Abelian groups are one-dimensional.* To see this, note that given $a' \in G$ then since G is Abelian,

$$a'(av) = a(a'v) \text{ for } a \in G \tag{3.2.2}$$

So a' itself is a homomorphism. Thus if V is irreducible, $a = \lambda(a)\mathbb{I}$ for all $a \in G$ (this is a diagonal matrix). But this is manifestly reducible since V is one-dimensional. Therefore, we can already classify the irreps of $U(1)$:

$$\rho_{(n)} = e^{in\theta};\ n \in \mathbb{Z} \tag{3.2.3}$$

where n is the *charge* of the $U(1)$ representation.

3.3 Building Representations

3.3.1 Dual and Conjugate Representation

if $\rho : G \to GL(n, \mathbb{C})$ is a representation, then dual, ρ^*, and conjugate, $\overline{\rho}$, representations are defined as,

$$\rho^*(a) = \rho(a^-1)^T \text{ for } a \in G \tag{3.3.1}$$

$$\overline{\rho}(a) = \rho(a)_{\text{c.c}} \text{ for } a \in G \tag{3.3.2}$$

where c.c denotes the complex conjugate. They can be shown to be homomorphisms. By definition,

$$\dim \rho = \dim \rho^* = \dim \overline{\rho} \tag{3.3.3}$$

In terms of the modules ρ^* acts on the *dual* vector space V^* and $\overline{\rho}$ acts on the *complex conjugate* vector space $\overline{V}$. We use the following notation: $v^i \in V$, $\omega_i \in V^*$ and $u^{\bar{i}} \in \overline{V}$. Then under these choices we the group action on each of these components are as follows:

$$v^i \mapsto \rho(a)^i{}_j v^j, \quad \omega_j \mapsto \rho^*(a)_i{}^j \omega_j = \omega_j \rho(a^{-1})_i{}^j, \quad u^{\bar{i}} \mapsto \bar{\rho}(a)^{\bar{i}}{}_{\bar{k}} u^{\bar{j}} \tag{3.3.4}$$

If $\bar{\rho} \sim \rho$ is real the representation is **real**. We have seen that we can for *direct sums* $\rho_i \oplus \rho_2$ representations via the module, $W = V_1 \oplus V_2 \ni (v_1, v_2)$. The group action on this element of the module shows we gave a representation,

$$g \cdot (v_1, v_2) = (gv_1, gv_2) \tag{3.3.5}$$

or as a representation,

$$\rho(a) = \begin{pmatrix} \rho_1(a) & 0 \\ 0 & \rho_2(a) \end{pmatrix}; \quad w = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{3.3.6}$$

3.3.2 Tensor Product Representations

We can also define *tensor product representation*

Given the vector spaces V and W the tensor product vector space $V \oplus W$ is defined by

- $(v_1 + v_2) \otimes w = v_1 \otimes w + v_2 \otimes w$ (linear)
- $v \otimes (w_1 + w_2) = v \otimes w_1 + v \otimes w_2$
- $(\lambda v) \otimes w = v(\lambda \otimes w) = \lambda(v \otimes w)$

In terms of a basis, we can write:

$$v = v^i e_i \in V, \quad w = w^a f_a \in W \tag{3.3.1}$$

Then we have,

$$U = U^{ia} e_i \otimes f_a = v^i w^a e_i \otimes f_a \in V \otimes W \tag{3.3.2}$$

where U^{ia} are the elements of a $\dim V \times \dim W$ matrix.

Given two G-modules V and W then the tensor product G-module $V \otimes W$ is given by

$$a(v \otimes w) = av \otimes aw \tag{3.3.3}$$

In terms of a representation,

$$U^{ia} \mapsto U'^{ia} = \rho(a)^{ia}{}_{jb} U^{jb} = \rho(a)^i{}_j \rho(a)^a{}_b U^{jb} \tag{3.3.4}$$

In other words if $d = \dim V \times \dim W$, then $\rho(a)^{ia}{}_{jb}$ is a $d \times d$ matrix acting on a d-dimensional *vector* U^{ia}. It might seem strange to call an object with two indices a *vector*, but we look at the pair (ia) as labelling the d-components of the following:

$$U = \begin{pmatrix} U^{11} \\ U^{12} \\ U^{21} \\ \vdots \\ U^{nn} \end{pmatrix} \tag{3.3.5}$$

3.3.3 Unitary Representations

For physics, a key refinement of a complex representation is a *unitary representation.* A unitary representation is a homomorphism $\rho : G \to U(d)$; that is to say, we expect symmetries in a quantum theory to be related as unitary representations. They require the inner product between states to be preserved. Suppose we have a Hilbert space $\mathcal{H}$ with some symmetry group G. This symmetry group maps physical states into physical states so we expect for each $a \in G$:

$$\begin{aligned} G : \mathcal{H} &\longrightarrow \mathcal{H}, \\ |\psi\rangle &\mapsto |\psi'\rangle = S(a)\,|\psi\rangle\,. \end{aligned} \tag{3.3.1}$$

Since this is a symmetry it should preserve the norm between the states is:

$$\begin{aligned} \langle\chi'|\psi'\rangle &= \langle S(a)\chi|S(a)\psi\rangle \\ &= \langle\chi|\, S(a)^\dagger S(a)\,|\psi\rangle \\ &= \langle\chi|\psi\rangle \end{aligned} \tag{3.3.2}$$

where $|\chi\rangle\,, |\psi\rangle \in \mathcal{H}$, we require $S(a)^\dagger S(a) = \mathbb{I}$ for $S(a)$ to be a symmetry. This is the condition that the operators be *unitary.* This actually gives a slightly more general definition. Let $U(\mathcal{H}) = \{$unitary operators on a Hilbert space $\mathcal{H}\}$. The point is that this allows $\mathcal{H}$ to be infinite-dimensional. Note that the unitary condition is actually a slightly strong condition. For physical states, we actually just require,

$$|\langle\psi'|\chi'\rangle\,|^2 = |\langle\psi|\chi\rangle\,|^2 \tag{3.3.3}$$

for all states $|\chi\rangle\,, |\psi\rangle \in \mathcal{H}$. But there are two possibilities

(a) Linear and unitary

$$\begin{aligned} S(a)\,[\alpha\,|\psi\rangle + \beta\,|\chi\rangle] &= \alpha S(a)\,|\psi\rangle + \beta S(a)\,|\chi\rangle \\ \langle S(a)\chi|S(a)\psi\rangle &= \langle\chi|\psi\rangle \end{aligned} \tag{3.3.4}$$

(b) anti-linear and anti-unitary

$$\begin{aligned} S(a)\,[\alpha\,|\psi\rangle + \beta\,|\chi\rangle] &= \alpha^* S(a)\,|\psi\rangle + \beta^* S(a)\,|\chi\rangle \\ \langle S(a)\chi|S(a)\psi\rangle &= \langle\chi|\psi\rangle^* \end{aligned} \tag{3.3.5}$$

case (b) requires the symmetry to include *time-reversal*, so we can argue on physical grounds that we should ignore it and assume that the symmetry is a linear and unitary operator on states in Hilbert space.

If a unitary representation is reducible then it is also decomposible. Hence all (finite-dimensional) unitary representations

$$\rho = \rho_1 \oplus \rho_2 \oplus \dots \rho_n \tag{3.3.6}$$

This is easy to see since a reducible rep is written,

$$\rho(a) = \begin{pmatrix} \hat{\rho}(a) & A(a) \\ 0 & B(a) \end{pmatrix} \tag{3.3.7}$$

Then if we impose the unitary condition $\rho(a)^\dagger \rho(a) = \mathbb{I}$ implies that $A(a) = 0$,

$$\begin{aligned} \rho(a)^\dagger \rho(a) &= \begin{pmatrix} \hat{\rho}(a)^\dagger & 0 \\ A(a)^\dagger & B^\dagger(a) \end{pmatrix} \begin{pmatrix} \hat{\rho}(a) & A(a) \\ 0 & B(a) \end{pmatrix} \\ &= \begin{pmatrix} \hat{\rho}(a)^\dagger \hat{\rho}(a) & \hat{\rho}(a)^\dagger A(a) \\ A(a)^\dagger \hat{\rho}(a) & A(a)^\dagger A(a) + B(a)^\dagger B(a) \end{pmatrix} = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & \mathbb{I} \end{pmatrix} \end{aligned} \tag{3.3.8}$$

This we require $\hat{\rho}(a)^\dagger A(a) = 0$ which implies that $A(a) = 0$. Thus we have the decomposition

$$\rho(a) = \begin{pmatrix} \hat{\rho}(a) & 0 \\ 0 & \hat{\hat{\rho}}(a) \end{pmatrix} \tag{3.3.9}$$

Another useful property is Wigner's unitarian trick:

If G is a compact Lie group then every complex representation is equivalent to a unitary representation. This means that for a compact Lie group, classifying the complex irrep classifies all the representations of physical interest. The proof involves considering G as a manifold (which it is) and integrating. This requires a *integral measure* (Haar measure) which only necessarily exists when G is compact.

(*Rough proof*): We start with a complex representation ρ. Choose some norm over the vector space - for example, $\langle v|v\rangle = v^\dagger v$. Then construct a new norm:

$$\langle v|w\rangle_\rho = \int da \, \langle \rho(a)v|\rho(a)w\rangle \tag{3.3.10}$$

where we integrate over the group (as a manifold). Then by construction,

$$\begin{aligned}\langle\rho(a)v|\rho(a)w\rangle_\rho &= \int_G da\,\langle\rho(a)\rho(b)v|\rho(a)\rho(b)w\rangle \\ &= \int_G da\,\langle\rho(ab)v|\rho(ab)w\rangle \\ &= \int_G da\,\langle\rho(a')v|\rho(a')w\rangle = \langle v|w\rangle_\rho\end{aligned} \tag{3.3.11}$$

Therefore $\rho(a)$ is unitary with respect to the new norm, and by definition $\exists T$ such that $\langle v|w\rangle_\rho = \langle Tv|Tw\rangle_\rho$, so $T^{-1}\rho(a)T$ is a unitary representation with respect to $\langle v|w\rangle$. The other point is that we immediately note that $U(n)$ is compact.

Additionally, we note that that *there are no faithful, finite-dimensional unitary representations of non-compact Lie groups.* The non-faithful condition is fine, consider $x \in \mathbb{R}$ under $+$, $\rho(x) = e^{i\alpha x}$ is unitary and non-faithful representation. For some more detail on this embedding, we impose an equivalence relation, $\sim$, on $\mathbb{R}$, such that $x \sim x + \frac{2\pi}{\alpha}$, this is topologically equivalent to $\mathbb{R}/\sim\ \simeq S^1$. The reason why this is a non-faithful representation is because $\rho(x) \sim \rho(x + 2\pi/\alpha)$ thus as map, ρ is not a bijection.

3.4 Representations of $SU(2)$

Let's find irreducible representations of $SU(2)$. By definition

$$SU(2) = \left\{ a = \begin{pmatrix} x & -y^* \\ y & x^* \end{pmatrix} : x^*x + y^*y = 1 \right\} \tag{3.4.1}$$

Since this is compact, we know that every complex representation will be equivalent to a unitary representation. Also note that all reps will be equivalent to reps of $SL(2, \mathbb{C})$.

trivial representation: the ρ_1 singlet

The module is one-dimensional $V \simeq \mathbb{C}$ and

$$\rho_1 : SU(2) \to GL(1, \mathbb{C}) = \mathbb{C} \tag{3.4.2}$$

$\rho_1(a) = 1$ for all $a \in SU(2)$.

defining representation: the ρ_2 doublet

The module is two-dimensional $V \simeq \mathbb{C}^2$ and

$$\rho_2 : SU(2) \to GL(2, \mathbb{C}) \tag{3.4.3}$$

$$\rho_2 = \begin{pmatrix} x & -y^* \\ y & x^* \end{pmatrix} \tag{3.4.4}$$

so $v^i \in V$, in components,

$$v^i \mapsto v'^i = \rho_2(a)^i{}_j v^j, \quad i,j = 1,2 \tag{3.4.5}$$

dual and complex representations: ρ_2^* and $\overline{\rho}_2$

The dual representation is defined as

$$\rho_2^* : SU(2) \to GL(2,\mathbb{C}) \tag{3.4.6}$$

Explicitly, the matrix acting on the dual space is,

$$\rho_2^*(a) = \rho_2(a^{-1})^T = \begin{pmatrix} x^* & -y \\ y^* & x \end{pmatrix} \tag{3.4.7}$$

and we have the complex conjugate representation,

$$\overline{\rho}_2 : SU(2) \to GL(2,\mathbb{C}) \tag{3.4.8}$$

$$\overline{\rho}_2(a) = \begin{pmatrix} x^* & -y \\ y^* & x \end{pmatrix} \tag{3.4.9}$$

We immediately see that $\overline{\rho}_2 \simeq \rho_2^*$. We also see that $T^{-1}\rho_2^*(a)T = T^{-1}\overline{\rho}_2(a)T = \rho_2 : T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. Hence all three representations are equivalent.

Tensor product representation: $\rho_2 \otimes \rho_2$

Consider the tensor product module: (4-dimensional), $W = V \otimes V$ with

$$v = \begin{pmatrix} v^{11} \\ v^{12} \\ v^{21} \\ v^{22} \end{pmatrix} \in W \tag{3.4.10}$$

such that the action of the group representation is

$$\begin{aligned} v^{ij} &\mapsto [\rho_2 \otimes \rho_2](a)^{ij}{}_{kl} v^{kl} \\ &= \rho_2(a)^i{}_k \; \rho_2(a)^j{}_l v^{kl} \\ &= (\rho(a)_2 \; v \; \rho_2(a)^T)^{ij} \end{aligned} \tag{3.4.11}$$

$$\rho_2(a)^i{}_k\ \rho_2(a)^j{}_l v^{kl} = \begin{pmatrix} x^2 & -xy^* & -y^*x & y^{*2} \\ xy & xx^* & -y^*y & y^*x^* \\ yx & -yy^* & x^*x & -x^*y^* \\ y^2 & yx^* & x^*y & x^{*2} \end{pmatrix} \begin{pmatrix} v^{11} \\ v^{12} \\ v^{21} \\ v^{22} \end{pmatrix} \tag{3.4.12}$$

We can decompose into *invariant subspaces*, if we denote a d-dimensional left-invariant G-module by W_d

$$V \otimes V = W_1 \oplus W_3 \tag{3.4.13}$$

$$\begin{aligned} W_1 &= \{v^{ij} \in V \otimes V : v^{ij} + v^{ji} = 0\} \text{ (anti-symmetric)} \\ W_3 &= \{v^{ij} \in V \otimes V : v^{ij} - v^{ji} = 0\} \text{ (symmetric)} \end{aligned} \tag{3.4.14}$$

then $v^{ij} = \lambda \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in W_1$. Acting with $\rho_2 \otimes \rho_2(a)$:

$$\begin{aligned} v^{ij} &\mapsto \lambda \begin{pmatrix} x & -y^* \\ y & x^* \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x & y \\ -y^* & x^* \end{pmatrix} \\ &= \lambda \begin{pmatrix} x & -y^* \\ y & x^* \end{pmatrix} \begin{pmatrix} -y^* & x* \\ -x & -y \end{pmatrix} \\ &= \lambda \begin{pmatrix} 0 & (x^*x + y^*y) \\ -(x^*x + y^*y) & 0 \end{pmatrix} = v^{ij} \end{aligned} \tag{3.4.15}$$

Thus W_1 is a *trivial* representation. Thus we can write

$$\rho_2 \otimes \rho_2 \simeq \rho_1 \oplus \rho_3 \tag{3.4.16}$$

For the ρ_3, we require $v^{12} = v^{21}$:

$$\begin{pmatrix} v^{11} \\ v^{12} = v^{21} \\ v^{22} \end{pmatrix} \mapsto \begin{pmatrix} x^2 & -2xy^* & y^{*2} \\ xy & (x^*x + y^*y) & -x^*y \\ y^2 & 2x^*y & x^{*2} \end{pmatrix} \begin{pmatrix} v^{11} \\ v^{12} = v^{21} \\ v^{22} \end{pmatrix} \tag{3.4.17}$$

This new representation is *irreducible*.

tensor product representation: $\rho_2 \otimes \rho_2 \otimes \rho_2$

Now we have an 8-dimensional vector space:

$$W = V \otimes V \otimes V \text{ with } v^{ijk} \in W \tag{3.4.18}$$

and

$$v^{ijk} \mapsto \rho_2 \otimes \rho_2 \otimes \rho_2(a)^{ijk}{}_{lmn} v^{lmn} \tag{3.4.19}$$

The invariant subspaces are

$$\begin{aligned}
&W_2 = \{v^{ijk} \in W : v^{ijk} + v^{jik} = 0\} \text{ (anti-symmetric on } ij\text{, 2-dimensional)}\\
&W_2' = \{v^{ijk} \in W : v^{ijk} - v^{ikj} = 0\} \text{ (anti-symmetric on } jk\text{, 2-dimensional)}\\
&W_4 = \{v^{ijk} \in W : \text{ total symmetric on } ijk\} \text{ (4-dimensional)}
\end{aligned} \tag{3.4.20}$$

$$\rho_2 \otimes \rho_2 \otimes \rho_2 \simeq \rho_2 \oplus \rho_2 \oplus \boxed{\rho_4} \tag{3.4.21}$$

The representation in the box is a new 4-dimensional representation (also an irrep).

$SU(2)$ **irreps:** ρ_n

Extending this pattern, we have the irreps:

$$\rho_n : SU(2) \to GL(n, \mathbb{C}) \tag{3.4.22}$$

which can be realised by *symmetric tensors*,

$$V_{n-1} \quad \subset \quad \bigotimes_{i=1}^{n-1} V_i = W \tag{3.4.23}$$

$$= \quad \{v^{i_1 \dots i_{n-1}} \in W \text{ symmetric on } i_1 \dots i_{n-1}\} \tag{3.4.24}$$

thus on action from the group, v transforms as:

$$\begin{aligned}
v^{i_1 \dots i_{n-1}} &\mapsto \hat{\rho}(a)^{i_1 \dots i_{n-1}}{}_{j_1 \dots j_{n-1}} v^{j_1 \dots j_{n-1}}\\
&= \rho_2(a)^{i_1}{}_{j_1} \dots \rho_2(a)^{i_{n-1}}{}_{j_{n-1}} v^{j_1 \dots j_{n-1}}
\end{aligned} \tag{3.4.25}$$

Looking at independent components, $v^{111\dots11}, v^{111\dots12}, v^{111\dots22}, \dots, v^{222\dots22}$, it is easy to see that since $(\rho_2^* \sim \overline{\rho}_2 \sim \rho_2)$, then this property is inherited by higher dimensional irreps,

$$\rho_n^* \sim \overline{\rho}_n \sim \rho_n. \tag{3.4.26}$$

If n is odd, then we have a real representation. We recall that $SO(3) \simeq SU(2)/\mathbb{Z}_2$ (identifying $a \sim -a$). Thus since, $\rho_2(-a) = -\rho_2(a) \neq \rho_2(a)$ we have:

(a) if n even ρ_n if *not* a representation of $SO(3)$,

(b) if n is odd ρ_n *is* a representation of $SO(3)$

Since $SO(3)$ is the rotation symmetry of the $n = 3$ representation, then the ρ_n representation corresponds to a state of spin, $s = \frac{1}{2}(n-1)$.

3.5 Young Tableu and $SU(n)$ irreps

We can generate the construction of $SU(2)$ irreps as symmetric tensors to general irreps of $SU(n)$ (or equally $SL(n,\mathbb{C})$). We start with the defining representation which is n-dimensional:

$$\rho_n : SU(n) \to GL(n,\mathbb{C}) \tag{3.5.1}$$

with the module $V \simeq \mathbb{C}^n$ with $v^i \in V,\ i = 1,\dots,n$. Then we have the *tensor product reps*,

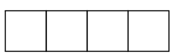
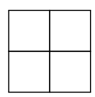
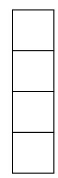
$$\rho_n \otimes \dots \otimes \rho_n : SU(n) \to GL(n^p,\mathbb{C}) \tag{3.5.2}$$

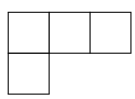
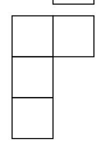
with the module $W = V \otimes \dots \otimes V$ with $v^{i_1\cdots i_p} \in W$. We can decompose W into irreps by *symmetrising* and *antisymmetrising* on indices. The combinatorics are encoded in *Young tableu*.

Given an ordered partition $\lambda = (p_1, p_2, \dots, p_s)$ of positive integers $p_i \in \mathbb{N}$ satisfying

$$p = p_1 + p_2 + \cdots + p_n, \quad p_i \in \mathbb{N}, \quad p_1 \geq p_2 \geq \dots p_s \tag{3.5.3}$$

For example for $p = 4$ there are 5 possibilities. To each possibility, we can attribute a *Young tableu*

$\lambda = (4)$  $\lambda = (2,2)$  $\lambda = (1,1,1,1)$ 

$\lambda = (3,1)$  $\lambda = (2,1,1)$  (3.5.4)

The finite-dimensional *irreducible representations* of $SU(n)$ are in *one-to-one correspondence* with Young tableu with $s < n$. The tableu encodes the symmetry properties of the irrep as an invariant subspace of $W = \bigotimes_{i=1}^p V_i$ with the rules

- antisymmetrize on columns (denoted by the operation a)
- symmetrize on rows (denoted by the operation s).

For example:

- ($p = 1$): $W = V \ni v^i \quad \boxed{i} \quad = v^i$ (defining representation has no sym or asym properties because there's only one index).

- ($p=2$): $W = V\otimes V \ni v^{ij}$

$$\begin{array}{|c|}\hline i \\ \hline j \\ \hline \end{array} : (av)^{ij} = \frac{1}{2}(v^{ij} - v^{ji}) \text{ antisymmetric on } (ij)$$
$$\begin{array}{|c|c|}\hline i & j \\ \hline \end{array} : (sv)^{ij} = \frac{1}{2}(v^{ij} + v^{ji}) \text{ symmetric on } (ij) \tag{3.5.5}$$

- ($p=3$): $W = V\otimes V\otimes V \ni v^{ijk}$

$$\begin{array}{|c|}\hline i \\ \hline j \\ \hline k \\ \hline \end{array} :$$

$$(a\cdot v)^{ijk} = \frac{1}{6}(v^{ijk} - v^{ikj} + v^{jki} - v^{jik} + v^{kij} - v^{kji}) \text{ antisymmetric on } (ijk) \tag{3.5.6}$$

$$\begin{array}{|c|c|}\hline i & k \\ \hline j & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array} :$$

$$\begin{aligned} (av)^{ij} &= \frac{1}{2}(v^{ij} - v^{ji}) \text{ antisymmetric on } (ij) \\ (s\cdot a\cdot v)^{ij} &= \frac{1}{4}(v^{ij} + v^{ji}) \text{ symmetric on } (ik) \end{aligned} \tag{3.5.7}$$

$$\begin{array}{|c|c|c|}\hline i & j & k \\ \hline \end{array} :$$

$$(s\cdot v)^{ijk} = \frac{1}{6}(v^{ijk} + v^{ikj} + v^{jki} + v^{jik} + v^{kij} + v^{kji}) \text{ symmetric on } (ijk) \tag{3.5.8}$$

- As a final example consider $W = V\otimes V\otimes V\otimes V \ni v^{ijkl}$

$$\begin{array}{|c|c|}\hline i & k \\ \hline j & l \\ \hline \end{array} :$$

$$\begin{aligned} (a\cdot v)^{ijkl} &= \frac{1}{4}(v^{ijkl} - v^{jikl} - v^{ijlk} + v^{jilk}),\ \text{antisymmetric on } (ij)(kl) \\ (s\cdot a\cdot v)^{ijkl} &= \frac{1}{4}(v^{ijkl} + v^{kjil} + v^{ilkj} + v^{klij} \\ &\quad - v^{jikl} - v^{jkil} - v^{likj} - v^{lkij} \\ &\quad - v^{ijlk} - v^{kjli} - v^{iljk} - v^{klji} \\ &\quad + v^{jilk} + v^{jkli} + v^{lijk} + v^{lkji}),\ \text{symmetric on } (ik)(jl) \end{aligned} \tag{3.5.9}$$

Note that antisymmetrization on n indices

$$v^{[i_1 \dots i_n]} = \lambda \varepsilon^{i_1 \dots i_n} \tag{3.5.10}$$

But,

$$\begin{aligned} v^{[i_1 \dots i_n]} &\mapsto \rho(a)^{i_1}_{\ j_1} \dots \rho(a)^{i_n}_{\ j_n} \lambda \varepsilon^{j_1 \dots j_n} \\ &= (\det \rho(a))\, \lambda \varepsilon^{i_1 \dots i_n} = v^{[i_1 \dots i_n]} \end{aligned} \tag{3.5.11}$$

where the final equality is obtained due to the fact that $SU(n)$ have $\det \rho = 1$ hence $v^{[i_1 \dots i_n]}$ is the trivial representation. These guys have no columns in the Young tableu.

3.5.1 Hook length and dimensions of irreps

We would also like to know the dimension of each module. Let $x = (ij)$ be the j^{th} box in the i^{th} row of a Young tableu $[\lambda]$. Then

$$\text{hook}(i, j) = \#\text{ boxes to the right} + \#\text{ number of boxes below} + 1 \tag{3.5.1}$$

Or pictorially,

$$\begin{array}{|c|c|c|c|c|c|c|c|} \hline \ & \ & \ & \ & \ & \ & \ & \ \\ \hline & (ij) & \rightarrow & \rightarrow & \rightarrow & \rightarrow \\ \cline{1-6} & \downarrow & & & \\ \cline{1-5} & \downarrow & \\ \cline{1-3} \ \\ \cline{1-1} \end{array}$$

Let W be the irreducible $SU(n)$ module corresponding to the Young tableau $[\lambda]$, then

$$\dim W = \prod_{x=(ij)\in[\lambda]} \frac{n + j - i}{\text{hook}(ij)} \tag{3.5.2}$$

$SU(7)$ examples

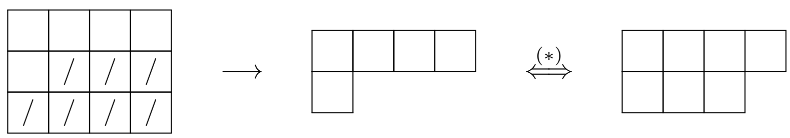
- $\dim \left(\begin{array}{|c|c|c|} \hline \ & \ & \ \\ \hline \ \\ \cline{1-1} \end{array} \right) = \dfrac{\begin{array}{ccc} 7 & 8 & 9 \\ 6 \end{array}}{\begin{array}{ccc} 4 & 2 & 1 \\ 1 \end{array}} = \frac{7\times8\times9\times6}{4\times2\times1\times1} = 378$

- $\dim \left(\begin{array}{|c|c|c|} \hline \ & \ & \ \\ \hline \ & \ & \ \\ \hline \ \\ \cline{1-1} \end{array} \right) = \dfrac{\begin{array}{ccc} 7 & 8 & 9 \\ 6 & 7 & 8 \\ 5 \end{array}}{\begin{array}{ccc} 5 & 3 & 2 \\ 4 & 2 & 1 \\ 1 \end{array}} = 7^2 \times 8 \times 9 = 3528$

3.5.2 Dual and Conjugate representations

We would also like a nice way to find *dual* and *conjugate* representations. Since $\rho_d^* \sim \overline{\rho}_d$ for the defining rep we have, $\rho^* \sim \overline{\rho}$ for all irreps (i.e dual and conjugate reps).

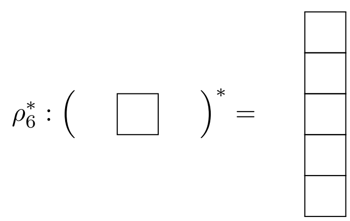
Dual (and equivalently conjugate) $SU(n)$ irreps correspond to Young Tableau that fit together as a rectangle with columns of length, n. For example for $SU(3)$,



(3.5.1)

where the two diagrams on the right are conjugate to one another.

For the defining representation of $SU(6)$, the conjugate rep is:

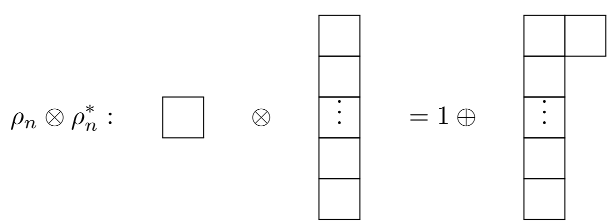


(3.5.2)

the dual is antisymemtric on $(n-1)$ indices. We can write $v^{i_1\cdots i_{n-1}} = \varepsilon^{i_1\cdots i_{n-1}} v_{i_n}$.

3.5.3 Tensor products of irreps

We can build higher dimensional irreps from smaller irreps,



(3.5.1)

where 1 denotes the singlet. The dimension of this new representation can be found with the usual calculation of hook lengths,

$$\dim\left(\begin{array}{|c|c|}\hline & \\ \hline & \multicolumn{1}{c}{} \\ \cline{1-1} \vdots & \multicolumn{1}{c}{} \\ \cline{1-1} & \multicolumn{1}{c}{} \\ \cline{1-1} & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array}\right) = \frac{\begin{pmatrix} n & n+1 \\ n-1 & \\ \vdots & \\ 2 & \end{pmatrix}}{\begin{pmatrix} n & 1 \\ n-2 & \\ \vdots & \\ 1 & \end{pmatrix}} = (n+1)(n-1) = n^2 - 1 \tag{3.5.2}$$

This representation in particular is interesting because it has the same dimension as the group $SU(n)$. Such a representation always exists called the *adjoint representation.* If we write adjoint as

$$v^{i_1 \ldots i_{d-1} j} = \varepsilon^{i_1 \ldots i_{d-1} i} a^j{}_i \tag{3.5.3}$$

then if $a^j{}_j = 0$ and

$$a^j{}_i \longrightarrow [\rho_d(a)]^j{}_k \, a^k{}_l \left[\rho_d(a)^{-1}\right]^l{}_i \qquad (\textit{adjoint transformation}) \tag{3.5.4}$$

Chapter 4

Particle Physics and Symmetries

4.1 The Poincaré Group

We have argues that in a quantum mechanical theory with Hilbert space $\mathcal{H}$ that physical symmetries correspond to unitary representations of a group G on $\mathcal{H}$ (unless the symmetry involves times-reversal). Since symmetries play a vital role in particle physics we will try and identify:

- which groups G are relevant to particle physics.
- what representations appear.

There are three types of symmeties:

- *spacetime symmetries*: Poincaré group
- *gauge symmetries* describe interactions
  - EM
  - strong force
  - weak force
- global symmetries
  - exact: lepton number, baryon number, ...
  - approximate: isospin, quark model, ...

4.1.1 Representations of the Poincaré group

The Poincaré is a matrix group

$$ISO(3,1) = \left\{ \begin{pmatrix} \Lambda & a \\ 0 & 1 \end{pmatrix} \in GL(5,\mathbb{R}) : \Lambda \in O(3,1) \right\} \tag{4.1.1}$$

acting $v = \begin{pmatrix} x \\ 1 \end{pmatrix}$

$$x^\mu \mapsto \Lambda^\mu{}_\nu x^\nu + a^\mu \tag{4.1.2}$$

Then Lorentz group is the subgroup with $a = 0$. The Lorentz group has two subgroups

- *orthochronous*: $O^+(3,1)$,

$$\begin{aligned} &\textit{future timelike} \longrightarrow \textit{future timelike (no time reversal)} \\ &\Lambda^0{}_0 \geq 1 \end{aligned} \tag{4.1.3}$$

- proper: $SO(3,1)$: $\det\Lambda = 1$ (no reflections)

It is natural to consider the intersection of these Lorentz subgroups:

$$\begin{aligned} SO^+(3,1) : SO(3,1) \cap O^+(3,1) &\quad \textit{proper and orthochronous}\text{ Lorentz group} \\ ISO^+(3,1) : &\quad \textit{proper and orthochronous}\text{ Poincaré group} \end{aligned} \tag{4.1.4}$$

The other *bits* of the Poincaré group can be reached with

- time reversal $\Lambda = \mathrm{diag}(-1,1,1,1)$
- parity: $\Lambda = \mathrm{diag}(1,-1,-1,-1)$

or the combinations of the above.

A theorem by Wigner classified the unitary representations of $ISO^+(3,1)$ with *non-negative energies* and definite *mass*. We can roughly prove this focusing on the essence of the proof, rather than the mathematical details. Note that $ISO^+(3,1)$ is *non-compact* (boosts and translations are not bounded) and so:

- unitary representations are ∞-dimensional $S(\Lambda, a)$

(The Lorentz group is also non-compact thus has no finite dimensional unitary representations).

Wigner starts with the *translation subgroup*:

$$x^\mu \mapsto x^\mu \tag{4.1.5}$$

$G = \mathbb{R}^4$ which is four copies of Abelian group. We know that the Abelian irreps are one-dimensional. For $G = \mathbb{R}$ under addition, irrep is labelled by the *charge*,

$$\rho_q = e^{-iqa} \tag{4.1.6}$$

where $q \in \mathbb{R}$ which is no faithful. For $G = \mathbb{R}^4$ we have four charges,

$$\rho_{p_n} = e^{-ip\cdot a} \tag{4.1.7}$$

where $p \cdot a = p_\mu a^\mu$ and $p^\mu = (E, \vec{p})$. In terms of the Hilbert space,

- Hilbert space: each irrep is one-dimensional, $|p^\mu\rangle$ and labels the basis,

$$\mathcal{H}_{p^\mu} = \{\lambda\,|p^\mu\rangle : \lambda \in \mathbb{C}\} \tag{4.1.8}$$

- unitary rep:

$$S(1,a)\,|p^\mu\rangle = e^{-ip\cdot a}\,|p^\mu\rangle \tag{4.1.9}$$

and *non-negative* energy means $E > 0$. *Momentum* of the state p^μ just labels *translation irrep*.

Next we consider the Lorentz group part. For this rep, we require

$$S(\Lambda,0)S(1,a)S(\Lambda,0)^{-1} = S(\Lambda\Lambda^{-1},\Lambda a) = S(0,\Lambda a) \tag{4.1.10}$$

Then we can form a *unitary rep* of the Poincaré group by

$$\begin{aligned} &S(\Lambda,0)S(1,a)S(\Lambda,0)^{-1}\left|\Lambda p\right\rangle \\ &\qquad = S(\Lambda,0)S(1,a)\left|p\right\rangle = e^{-ip\cdot a}S(\Lambda,0)\left|p\right\rangle \\ &\qquad = e^{ip\cdot a}\left|\Lambda p\right\rangle \\ &\qquad = S(1,\Lambda a)\left|\Lambda p\right\rangle = e^{-i\Lambda p\cdot\Lambda a}\left|\Lambda p\right\rangle \end{aligned} \tag{4.1.11}$$

Actually we don't have to take all states: only the ones with fixed p^2

$$\mathcal{H} = \bigoplus_{p^\mu:\ p^2=m^2} \mathcal{H}_p \tag{4.1.12}$$

where in general,

$$S(\Lambda,a)\left|p\right\rangle = S(1,a)S(\Lambda,0)\left|p\right\rangle = e^{-i\Lambda p\cdot a}\left|\Lambda p\right\rangle \tag{4.1.13}$$

In the simplest case: $p^\mu = 0$

$$\mathcal{H} = \{\lambda\left|0\right\rangle : \lambda\in\mathbb{C}\} \tag{4.1.14}$$

$S(\Lambda,a)\left|0\right\rangle = \left|0\right\rangle$ which transforms under the trivial representation, this is the *ground state* of QFT. But for $m^2 \geq 0$

$$\begin{aligned} \mathcal{H} = \{\left|\psi\right\rangle &= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2E_{\vec{p}}}\psi(\vec{p})\left|p\right\rangle : p^2 = m^2 \\ &= \text{one-particle Hilbert space of QFT (scalar)} \end{aligned} \tag{4.1.15}$$

Now note that we have,

$$\mathcal{H} = \{\text{positive energy solutions of the Klein-Gordon equation}\} \tag{4.1.16}$$

So, one way to view the ∞-dimensional irrep of the Poincaré group: **field representation**. This construction only gives spin-0 rep. To see the other spins, consider

1. $p^2 = m^2 > 0$
2. $p^2 = 0$

We are able to perform boosts in case 1 so we move to the intertial frame of our particle, $p^\mu = (E,0,0,0)$. One can define

$$H \subset SO^+(3,1) = \{\Lambda^\mu{}_\nu p^\nu = p^\mu\} \tag{4.1.17}$$

this is the *little group* or the *stabalizer*. It is the set of group elements that leave the momentum of the particle unchanged. In this case $H = SO(3)$ which is the group of rotations in $\mathbb{R}^3$. Thus for the scalar rep: $S(\Lambda, 0)\,|p\rangle = |p\rangle$ if $\Lambda \in SO(3)$. This means we are free to consider an $SO(3)$ irrep for each $|p\rangle$; or more generally $SU(2)$ (we showed that odd-dimensional irreps of $SU(2)$ are irreps of $SO(3)$).

$$\mathcal{H}_{p,s} = \mathcal{H}_p \otimes \mathcal{H}_s \tag{4.1.18}$$

where $\mathcal{H}_s = SU(2)$ irrep space of dimension $2s+1$ (we use the fact that since $SU(2)$ is compact its complex irreps are equivalent to unitary reps). Recall that $SU(2)$ irreps are labelled by dimension $2s+1$

- $s = 0$: singlet
- $s = \frac{1}{2}$: doublet v^i : $|\frac{1}{2}\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $|-\frac{1}{2}\rangle \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
- $s = 1$: triplet $v^{(ijk)}$: $|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\quad |0\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $|-1\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
- and so on.

Then for the irreps of Poincaré:

$$m^2 > 0, \text{ spins} \quad \mathcal{H} = \bigoplus_{p^\mu:\ p^2 = m^2} \mathcal{H}_{p^\mu, s} \tag{4.1.19}$$

We really need to give,

$$S(\Lambda, a)\,|p^\mu, m\rangle = ? \tag{4.1.20}$$

Choose a *standard boost*

$$p^\mu = L^\mu{}_\nu(p) p_0{}^\nu \tag{4.1.21}$$

where $p_0{}^\nu = (m, 0, 0, 0)$. This is not unique. Given a general Λ then,

$$L(\Lambda p)^{-1}\Lambda \cdot L(p) \in H(p_0) \tag{4.1.22}$$

where $H(p_0)$ is the little group for p_0. We define

$$S(\Lambda, a)\,|p^\mu, m\rangle = e^{-i\Lambda p \cdot a} \rho_s(W(\Lambda, p))^{m'}{}_m\,|\Lambda, m'\rangle \tag{4.1.23}$$

where $\rho_s(W(\Lambda, p))^{m'}{}_m$ is the spin-s rep. Next we embed the spin-s rep into the *non-unitary* $SO^+(3,1)$ rep. However, in each case

- *spin-s* Hilbert space $\simeq$ *solutions of wave-equation*

$$\begin{array}{ccc} s = 0 & \textit{Klein-Gordon equation} & (\Box + m^2)\psi(x) \\ s = 1/2 & \textit{Dirac equation} & (i\partial\!\!\!/ - m)\psi(x) \\ s = 1 & \textit{Proca equation} & \partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) + m^2 A_\nu = 0 \\ \vdots & \vdots & \vdots \end{array} \tag{4.1.24}$$

Take the *Proca equation*:

$$\partial^\mu(\partial_\mu A_\nu - \partial_\nu A_\mu) + m^2 A_\nu = 0 \tag{4.1.25}$$

the solution $A_\mu = \varepsilon_\mu e^{-ip\cdot x}$ with $p^2 = m^2$ and $p^\mu \varepsilon_\mu = 0$. The condition on the polarization vector shows

$$p^\mu = \begin{pmatrix} E, & 0, & 0, & 0 \end{pmatrix} \text{ and } \varepsilon_\mu = \begin{pmatrix} 0, & \varepsilon_x, & \varepsilon_y, & \varepsilon_z \end{pmatrix} \tag{4.1.26}$$

Next we need to consider $m^2 = 0$. We can have,

$$p^\mu = \begin{pmatrix} E & E & 0 & 0 \end{pmatrix}^\mu \tag{4.1.27}$$

then the little group has the explicit form,

$$\Lambda^\mu{}_\nu = \begin{pmatrix} 1 + \frac{1}{2}(a^2+b^2) & -\frac{1}{2}(a^2+b^2) & a & b \\ \frac{1}{2}(a^2+b^2) & 1 - \frac{1}{2}(a^2+b^2) & a & b \\ a\cos\theta + b\sin\theta & -(a\cos\theta + b\sin\theta) & 1 & 0 \\ -a\sin\theta + b\cos\theta & a\sin\theta - b\cos\theta & 0 & 1 \end{pmatrix} \tag{4.1.28}$$

This can be split up in an interesting way,

$$\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta & \sin\theta \\ & & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 + \frac{1}{2}(a^2+b^2) & -\frac{1}{2}(a^2+b^2) & a & b \\ \frac{1}{2}(a^2+b^2) & 1 - \frac{1}{2}(a^2+b^2) & a & b \\ a & -a & 1 & 0 \\ b & -b & 0 & 1 \end{pmatrix} \tag{4.1.29}$$

What is this? Take a vector defined by

$$v^\mu = \begin{pmatrix} 1 + \frac{1}{2}(x^2+y^2) \\ \frac{1}{2}(x^2+y^2) \\ x \\ y \end{pmatrix} \tag{4.1.30}$$

then acting Λ on v (defining $r^2 = \frac{1}{2}(x^2+y^2)$)

$$\begin{aligned} &\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \cos\theta & \sin\theta \\ & & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 + \frac{1}{2}(a^2+b^2) & -\frac{1}{2}(a^2+b^2) & a & b \\ \frac{1}{2}(a^2+b^2) & 1 - \frac{1}{2}(a^2+b^2) & a & b \\ a & -a & 1 & 0 \\ b & -b & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 + r^2 \\ r^2 \\ x \\ y \end{pmatrix} \\ &= \begin{pmatrix} 1 + \frac{1}{2}(x+a)^2 + \frac{1}{2}(y+b)^2 \\ \frac{1}{2}(x+a)^2 + \frac{1}{2}(y+b)^2 \\ (x+a)\cos\theta + (y+b)\sin\theta \\ -(x+a)\sin\theta + (y+b)\cos\theta \end{pmatrix} \end{aligned} \tag{4.1.31}$$

which is just

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \tag{4.1.32}$$

Thus the little group is just $ISO(2)$ which is the group of symmetries on $\mathbb{R}^2$. The analysis is just as for $ISO^+(3)$: first consider **translations**,

- $\rho_{(k)} = e^{-i\vec{k}\cdot\vec{a}}$
- The little group for $\begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$ is *trivial* so finished.

Unless $k = 0$,

- $\rho_{(\alpha)} = e^{-i\alpha\theta}$

We start by considering translations, with $\vec{a} = (a, b)$ then irreps are 1-dimensional labelled by choice $\vec{k} = (k_1, k_2)$

- $\rho_k = e^{-i\vec{k}\cdot\vec{a}}$
- Hilbert space $\mathcal{H}_k : M(0, \vec{a})\,|\vec{k}\rangle = e^{-i\vec{k}\cdot\vec{a}}\,|\vec{k}\rangle$

  For the full *scalar* rep of $ISO(2)$:

- $\mathcal{H} = \bigoplus_{\vec{k};\ k^2=t^2} \mathcal{H}_{\vec{k}}$
- $M(\theta, \vec{a})\,|\vec{a}\rangle = e^{-iR(\theta)\vec{k}\cdot\vec{a}}\,|R(\theta)\vec{k}\rangle$

What about the *little group*? Stabilizer of $\vec{k}$ is $\mathbb{I}$ (trivial little group). However if we take the trivial rep: $\vec{k} = 0$, then the stabilizer of $\vec{k} = 0$ is $SO(2)$. So we have a family of irreps (not faithful)

$$M_\alpha(\theta, \vec{a}) = e^{-i\alpha\theta}; \quad \alpha \in \frac{1}{2}\mathbb{Z} \tag{4.1.33}$$

we allow reps when $\rho_\alpha(2\pi, \vec{a}) = -\rho_\alpha(2\pi, \vec{a})$. So there are two possibilities:

1. continuous spin labelled by t^2,

$$\mathcal{H} = \bigoplus_{\vec{k};k^2=t^2} \mathcal{H}_{\vec{k}} \qquad M(\theta, \vec{a})\,|\vec{a}\rangle = e^{-iR(\theta)\vec{k}\cdot\vec{a}}\,|R(\theta)\vec{k}\rangle \tag{4.1.34}$$

2. *helicity* labelled by $\alpha \in \frac{1}{2}\mathbb{Z}$

$$\textit{1-dimensional, } \mathcal{H}_\alpha \qquad \rho_\alpha(\theta, \vec{0}) = e^{-i\alpha\theta} \tag{4.1.35}$$

Physically, we do not see the continuous spin reps, so we define

$$\mathcal{H}_{p,\alpha} = \mathcal{H}_p \otimes \mathcal{H}_\alpha \tag{4.1.36}$$

Then the Poincaré irrep has,

- $m^2 = 0$, with helicity, α
- $\mathcal{H} = \oplus_{p^\mu;p^2=0}\mathcal{H}_{p_\alpha}$

- $S(\Lambda, a)\, |p^\mu, \alpha\rangle = e^{-i\Lambda p \cdot a} \rho_\alpha(W(\Lambda, p))\, |\Lambda p, \alpha\rangle$

Usually, one combines $\pm\alpha$ as ± 1 helicity reps of spin $|\alpha|$. Again $\mathcal{H}$ can be viewed as the space of solutions of a *massless equation*.

4.2 Quarks, leptons and $SU(3) \times SU(2) \times U(1)$

Let's now summaries the particles of the Standard Model and the *gauge symmetries*. Recall that the *matter* particles are all spin-$\frac{1}{2}$. We have 3 families of leptons with masses:

$$\begin{array}{cccc} e^-\ (0.5MeV) & \mu^-\ (106MeV) & \tau^-\ (1.8GeV) & \\ \nu_e\ (0. < 10eV) & \nu_\mu\ (< 0.16MeV) & \nu_\tau\ (< 18MeV) & \textit{neutrinos} \end{array} \tag{4.2.1}$$

and 3 families of quarks:

$$\begin{array}{ccc} u\ (\approx 0.4GeV) & c\ (1.6GeV) & t\ (175GeV) \\ d\ (\approx 0.4GeV) & s\ (0.5GeV) & b\ (4.5GeV) \end{array} \tag{4.2.2}$$

We cannot observe quarks as free particles (*confinement*) but only as bound stated (*the hadrons*). The masses are *effective* - the rough contribution of each quark to the bound state.

In SM mass for quarks and lepton comes from *spontaneous symmetry breaking* via the Higgs particle. So to describe symmetries let's start by considering very high energies (*electroweak scale*) where all fermions are effectively massless. We see that

$$\begin{array}{ccc} m^2 = 0 & \text{spin-}\frac{1}{2} & \text{2 different helicity reps} \\ & \mathcal{H}_{-\frac{1}{2}} & \mathcal{H}_{+\frac{1}{2}} \\ & \textit{left-handed helicity} & \textit{right-handed helicity} \end{array} \tag{4.2.3}$$

so we can decompose for each particle into two reps,

$$l^- \to l^-_L,\ l^-_R \quad \text{and} \quad q \to q_L,\ q_R \tag{4.2.4}$$

where l^- represents *leptons* and q represents the *quarks*. Focusing on the *Electroweak symmtry* via Salam and Weinberg.

$$\textit{electroweak}: \quad U(1)_Y \times SU(2) \tag{4.2.5}$$

we know the irreps

$$\begin{array}{ll} U(1)_Y: & \rho_q = e^{iq\theta} \\ SU(2): & \underline{1}, \underline{2}, \underline{3}, \ldots \end{array} \tag{4.2.6}$$

So, how do the quarks and leptons form representations of $U(1)_Y \times SU(2)$?

Let's start with $SU(2)$, we have only *doublets* and *singlets* but the two different helicities are treated differently:

- doublets: $\underline{2}$

$$\begin{pmatrix} \nu_{e,L} \\ e^-_L \end{pmatrix}, \quad \begin{pmatrix} \nu_{\mu,L} \\ \mu^-_L \end{pmatrix}, \quad \begin{pmatrix} \nu_{\tau,L} \\ \tau^-_L \end{pmatrix}$$
$$\begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad \begin{pmatrix} c_L \\ s_L \end{pmatrix}, \quad \begin{pmatrix} t_L \\ b_L \end{pmatrix} \tag{4.2.7}$$

- singlets: $\underline{1}$

$$\begin{matrix} e^-_R & \mu^-_R & \tau^-_R \\ u_R, d_R & c_R, s_R & t_R, b_R \end{matrix} \tag{4.2.8}$$

The usual assumption is that $\nu_{e,R}$, $\nu_{\mu,R}$, $\nu_{\tau,R}$ which means that right-handed neutrinos don't get mass from the Higgs. Thus as fields we could write a *doublet* of the Dirac fields or singlet

$$\begin{aligned} l^i_L &= \begin{pmatrix} \psi_{\nu_e,L} \\ \psi_{e,L} \end{pmatrix}, \quad l_R = \psi_{e,R} \\ \mathcal{H} &= \underset{\text{Poincaré irrep space}}{\mathcal{H}(m^2 = 0, \alpha = -\tfrac{1}{2})} \otimes \underset{SU(2)\text{ doublet space}}{\mathcal{H}_{+\frac{1}{2}}} \end{aligned} \tag{4.2.9}$$

Next we need $U(1)_Y$, writing $\rho_Y = e^{3iqY}$, where the 3 is the conventional *weak hypercharge* $Y = \{0\pm, \pm 1/3, \pm 2/3, \dots\}$. We have

$$\begin{matrix} \begin{pmatrix} \nu_{e,L} \\ e^-_L \end{pmatrix} & \begin{pmatrix} \nu_{e,L} \\ e^-_L \end{pmatrix} & \begin{pmatrix} \nu_{e,L} \\ e^-_L \end{pmatrix} & Y = -1 \\ e^-_R & \mu^-_R & \tau^-_R & Y = -2 \\ \begin{pmatrix} u_L \\ d_L \end{pmatrix} & \begin{pmatrix} c_L \\ s_L \end{pmatrix} & \begin{pmatrix} t_L \\ b_L \end{pmatrix} & Y = 1/3 \\ u_R & c_R & t_R & Y = -4/3 \\ d_R & s_R & b_R & Y = -2/3 \end{matrix} \tag{4.2.10}$$

For each particle we can define the *conjugate anti-particle* by taking the *conjugate representation*. Remember $\overline{\underline{2}} \sim \underline{2}$ for $SU(2)$ and for helicity $L \longleftrightarrow R$ under conjugation, so

$$\begin{matrix} e^+_R & \textit{is conjugate of } e^-_L & (+ \longleftrightarrow - \textit{ for leptons}) \\ \overline{u}_R & \textit{is conjugate of } u_L & (u, \overline{u} \textit{ for neutrinos and quarks}) \end{matrix} \tag{4.2.11}$$

So

$$\begin{array}{cccc}
\begin{pmatrix}\overline{e}^+_R\\ \overline{\nu}_{e,R}\end{pmatrix} & \textit{doublet} & Y=+1 & \textit{...for other left-handed leptons}\\[2ex]
\overline{e}^+_L & \textit{singlet} & Y=+2 & \textit{...for other right-handed leptons}\\[2ex]
\begin{pmatrix}\overline{d}_R\\ \overline{u}_R\end{pmatrix} & \textit{doublet} & Y=-1/3 & \textit{...for other right-handed quarks}\\[2ex]
\begin{array}{c}\overline{d}_R\\ \overline{u}_R\end{array} & \textit{singlets} & \begin{array}{c}Y=-4/3\\ Y=2/3\end{array} & \textit{...for other left-handed quarks}
\end{array} \tag{4.2.12}$$

Finally we also have,

$$SU(3) \quad \textit{colour symmetry} \quad \textit{quantum chromodynamics} \tag{4.2.13}$$

We know that

- trivial representation $\underline{1}$
- defining representation $\underline{3}$
- conjugate of defining representation $\overline{\underline{3}}$.

we have

- all leptons are *singlets*
- all quarks are *triplets*, $\underline{3}$ (antiquarks are $\overline{\underline{3}}$)

so we note that each quark comes in 3 different types

$$u_R = \begin{pmatrix}u_{R,1}\\ u_{R,2}\\ u_{R,3}\end{pmatrix} \begin{array}{cc}\longleftarrow & \textit{red}\\ \longleftarrow & \textit{green}\\ \longleftarrow & \textit{blue}\end{array} \tag{4.2.14}$$

So for example $\begin{pmatrix}u_L\\ d_L\end{pmatrix}$ has a representation space,

$$\begin{array}{ccccccccc}
\mathcal{H} = & \mathcal{H}(m^2=0,\alpha=-1/2) & \otimes & \mathcal{H}_{1/2} & \otimes & \mathcal{H}_{4/3} & \otimes & \mathcal{H}_{\underline{3}}\\
& \uparrow & & \uparrow & & \uparrow & & \uparrow\\
& \textit{Poincaré} & & SU(2) & & U(1) & & SU(3)
\end{array} \tag{4.2.15}$$

$$\begin{array}{l}\psi = \psi_L^{ia} \quad i: \;\; SU(2) \;\; \textit{doublet} \quad a: \;\; SU(3) \;\; \textit{triplet}\\ \textit{field}\end{array} \tag{4.2.16}$$

4.3 Breaking discrete symmetries

Consider an theory of a one-dimensional scalar field, $\phi(x)$ equipped with a $\mathbb{Z}_2$ symmetric potential such that $V(\phi)$ is symmetric under the following transformation $\phi(x) \to -\phi(x)$. Explicitly this can be written as

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi \partial^\mu \phi - V(\phi); \quad V(\phi) = -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4 \tag{4.3.1}$$

This potential has two vacuum states, v, which can be found by differentiation with respect to ϕ

$$\left[\frac{\partial V}{\partial \phi}\right]_{\phi=v} = -\mu^2 v + \lambda v^3 = 0 \tag{4.3.2}$$

from which we find, $v = \pm\mu/\sqrt{\lambda}$. Now let's look at perturbations of the potential about the positive vacuum state.

$$\begin{aligned}
V(v+\phi) &= -\frac{1}{2}\mu^2(v+\phi)^2 + \frac{1}{4}\lambda(v+\phi)^4 \\
&= -\frac{1}{2}\mu^2(v^2 + 2v\phi + \phi^2) + \frac{1}{4}\lambda(v^4 + 4v^3\phi + 6v^2\phi^2 + v\phi^3 + \phi^4) \\
&= -\frac{\mu^4}{2\lambda} - \frac{\mu^3}{\sqrt{\lambda}}\phi - \frac{1}{2}\mu^2\phi^2 + \frac{\mu^4}{4\lambda} + \frac{\mu^3}{\sqrt{\lambda}}\phi + \frac{3\mu^2}{2}\phi^2 + \frac{\mu}{\sqrt{\lambda}}\phi^3 + \frac{\lambda}{4}\phi^4 \\
&= -\frac{\mu^4}{4\lambda} + \mu^2\phi^2 + \sqrt{\lambda}\mu\phi^3 + \frac{\lambda}{4}\phi^4
\end{aligned} \tag{4.3.3}$$

All that for what? Barring that irrelevant constant (doesn't apprear in EOM), we can see that that second term is describing a massive scalar with a mass, $m_\phi = \sqrt{2}\mu$, and a bunch of higher order terms describing interactions. So we started with a theory with two degenerate vacuum states, when we expanded about the vacuum, we found that it resulted in appearance of the massive scalar we know and love. This is the most basic example of spontaneous symmetry breaking. We considered a theory with a discrete symmetry, what about theories with continuous symmetries?

4.4 Spontaneously Broken Continuous Symmetries

Let's look at a simple Abelian example of the $U(1)$ Lagrangian. The Lagrangian is given by,

$$\mathcal{L} = \partial_\mu\phi^*\partial^\mu\phi - V(\phi^*, \phi); \quad V(\phi^*, \phi) = -\mu^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4 \tag{4.4.1}$$

The potential is invariant under a global $U(1)$ transformation such that $\phi \to e^{i\theta}\phi$. Cool. To find the vacua of the theory, we differentiate

$$\frac{\partial V}{\partial \phi^*} = -\mu^2\phi + \lambda|\phi|^2\phi \tag{4.4.2}$$

Which shows that the condition for the minimum is complex numbers with square modulus,

$$|\phi|^2 = \frac{\mu^2}{\lambda} \equiv \frac{v^2}{2}, \tag{4.4.3}$$

were the factor of $\frac{1}{2}$ is just a convenient choice. This means that the vacua themselves occupy a circle,

$$\phi = \frac{v}{\sqrt{2}} e^{i\theta}, \quad \theta \in [0, 2\pi). \tag{4.4.4}$$

In general, a continuous set of minima is called a *vacuum manifold*. As in the discrete case, the different vacua are all identical, and so we choose the one that forces the vacuum to be real, $\phi = \phi_0 = \frac{v}{\sqrt{2}}$ with real v. We choose to then expand around the this minimum, by using the following form of the scalar,

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \varphi(x) + i\xi(x)), \quad \varphi, \xi \in \mathbb{R}. \tag{4.4.5}$$

Focus on expanding the potential about the vacuum of choice,

$$\begin{aligned}
V(\phi^*, \phi) &= -\mu^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4 \\
&= -\frac{1}{2}\mu^2[(v+\varphi)^2 + \xi^2] + \frac{1}{8}\lambda[(v+\varphi)^2 + \xi^2]^2 \\
&= -\frac{1}{2}\mu^2 v^2 - \mu^2 v\varphi - \frac{1}{2}\mu^2(\varphi^2 + \xi^2) + \frac{1}{8}\lambda(v^2 + 2v\varphi + \varphi^2 + \xi^2)^2 \\
&= -\frac{1}{2}\mu^2 v^2 - \mu^2 v\varphi - \frac{1}{2}\mu^2(\varphi^2 + \xi^2) + \frac{1}{8}\lambda(v^2 + 2v\varphi)^2 + \frac{1}{4}\lambda(v^2 + 2v\varphi)(\varphi^2 + \xi^2) + \frac{1}{8}\lambda(\varphi^2 + \xi^2)^2 \\
&= -\frac{1}{2}\mu^2 v^2 - \mu^2 v\varphi + \frac{1}{8}\lambda(v^2 + 2v\varphi)^2 + \frac{1}{2}\lambda v\varphi(\varphi^2 + \xi^2) + \frac{1}{8}\lambda(\varphi^2 + \xi^2)^2 \\
&= -\frac{1}{2}\mu^2 v^2 - \mu^2 v\varphi + \frac{1}{8}\lambda(v^2 + 2v\varphi)^2 + \frac{1}{2}\lambda v\varphi(\varphi^2 + \xi^2) + \frac{1}{8}\lambda(\varphi^2 + \xi^2)^2 \\
&= -\frac{1}{2}\mu^2 v^2 - \mu^2 v\varphi + \frac{1}{8}\lambda v^4 + \mu^2 v\varphi + \frac{1}{2}\lambda v^2\varphi^2 + \frac{1}{2}\lambda v\varphi(\varphi^2 + \xi^2) + \frac{1}{8}\lambda(\varphi^2 + \xi^2)^2 \\
&= -\frac{\mu^4}{2\lambda} + \mu^2\varphi^2 + \mathcal{L}_{\text{int}}
\end{aligned} \tag{4.4.6}$$

where $\mathcal{L}_{\text{int}}$ contains the interactions between φ and ξ. As before, we end up with one massive scalar, φ with mass $m_\varphi = \sqrt{2}\mu$ but the mass term for ξ is not present in the final form of the perturbed Lagrangian. A degree of freedom whose mass term disappears in this way is called a *Goldstone Boson*.

4.5 Scalar Mass Matrix

In general, we can consider an n-component real scalar field $\phi = (\phi_1, ..., \phi_n)^T$ with a Lagrangian,

$$\mathcal{L} = \frac{1}{2}\partial_\mu \phi^T \partial^\mu \phi - V(\phi) \tag{4.5.1}$$

This particular kinetic term requires the symmetry transformation to be *orthogonal*, which we will assume for simplicity.

The vacuum state $\phi = \phi_0$ can be found from the condition,

$$\left.\frac{\partial V}{\partial \phi_i}\right|_{\phi=\phi_0} = 0 \quad \text{for all } i \tag{4.5.2}$$

Taylor expanding around the vacuum using $\phi = \phi_0 + \varphi$, we have,

$$V(\phi) = V(\phi_0 + \varphi) = V(\phi_0) + \varphi_i \left.\frac{\partial V}{\partial \phi_i}\right|_{\phi_0} + \frac{1}{2}\varphi_i\varphi_j \left.\frac{\partial^2 V}{\partial \phi_i \partial \phi_j}\right|_{\phi_0} + ... \tag{4.5.3}$$

which can be written in the following way,

$$V(\phi) = V(\phi_0) + \frac{1}{2}m_{ij}^2\varphi_i\varphi_j + \mathcal{O}(\varphi^3). \tag{4.5.4}$$

Where the linear order term vanishes due to the vacuum being a minimum of the potential. We have also defined the *scalar mass matrix*,

$$m_{ij}^2 \equiv \left.\frac{\partial^2 V}{\partial \phi_i \partial \phi_j}\right|_{\phi_0} \tag{4.5.5}$$

The eigenvectors of this matrix determine the particle species, and their masses are the square roots the eigenvalues of m_{ij}^2.

4.6 Broken Generators and Goldstone's Theorem

Assume now that the theory is invariant under some symmetry group G, under which the field transforms as $\phi \to M\phi$, where $M \in \rho(G)$ where $\rho(G)$ is an appropriate representation of the group G. Consider an infinitesimal transformation,

$$\phi_i \to \phi_i + i\theta^a T_{ij}^a \phi_j. \tag{4.6.1}$$

Similarly, we find that the vev ϕ_0 transforms as $\phi_0 \to \phi_0 + \delta\phi$, where $\delta\phi = i\theta^a T_{ij}^a \phi_{0j}$

As a simple example, we choose $SO(3)$ with a vev of $\phi_0 = (0, 0, v)^T$. In this case, we have,

$$t^1\phi_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ -iv \\ 0 \end{pmatrix} \tag{4.6.2}$$

$$t^2\phi_0 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ v \end{pmatrix} = \begin{pmatrix} iv \\ 0 \\ 0 \end{pmatrix} \tag{4.6.3}$$

$$t^3\phi_0 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{4.6.4}$$

We see that the generators t^1 and t^2 change the direction of the vev, ϕ_0, but the generator t^3, does not. We say that t^1 and t^2 are broken whereas t^3 is an unbroken generator. In general, if T is an unbroken generator, then $T\phi_0 = 0$.

In general, depending on the choice of the generators and the vev, ϕ_0, is it possible that even though $T^a\phi_0 \neq 0$ for all T^a, there might be some general linear combination $\hat{T} = c^a T^a$ for which $\hat{T}\phi_0 = 0$. Therefore, we define a $d \times d$ symmetry-breaking matrix:

$$\begin{aligned} S^{ab} &= -(T^a\phi_0)^T T^b \phi_0 \\ &= -\phi_0^T (T^a)^T T^b \phi_0 \\ &= \phi_0^T T^a T^b \phi_0 \\ &= \phi_{0i} T^a_{\ ik} T^b_{\ jk} \phi_{0k} \end{aligned} \tag{4.6.5}$$

This matrix is real and symmetric. Therefore, it's eigenvalues are real and non-negative. In the $SO(3)$ example, we have,

$$S^{ab} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & v^2 & 0 \\ 0 & 0 & v^2 \end{pmatrix} \tag{4.6.6}$$

If $\hat{T}\phi_0 = 0$ for some $\hat{T} = c^a T^a$, then,

$$S^{ab}c^b = \phi_0^T T^a c^b T^b \phi_0 = \phi_0^T T^a \hat{T} \phi_0 = 0 \tag{4.6.7}$$

which means that $c^a T^a \phi_0 = 0$, therefore $\hat{T} = c^a T^a$ is an unbroken generator. Therefore, there is a *one-to-one correspondence* between the **unbroken generators** and the **zero-eigenvalues**. The converse of this statement is also true, any eigenvector with a non-zero eigenvalue corresponds to a broken generator.

Overall, the matrix has d orthonormal eigenvectors c^a and corresponding eigenvalues $\lambda^A \geq 0$ where $A \in \{1, ..., d\}$. Let us order them in such a way that any zero eigenvalues come first, such that $\lambda^A = 0$ for $A \leq d'$, where d' is the number of zero eigenvalues. Correspondingly, $\lambda^A > 0$ for $d' < A \leq d$. Thus we define a new set of generators, $\hat{T} \equiv (c^A)^a T^a$ and vectors $\hat{\phi}^A \equiv i\hat{T}^A\phi_0$. Then,

- $\hat{\phi}^A = 0$ for $A \leq d'$,
- $\hat{\phi}^A \neq 0$ for $d < A \leq d'$.

This means that $\hat{T}^A$ for $A \leq d'$ are the unbroken generators. They generate a Lie group H which is a subgroup of the original group G and under which the vacuum state is symmetric. This is called the **residual symmetry group**. We say that the *symmetry breaking pattern* is $G \to H$.

Unbroken Generator of $SO(3)$

In the example of $SO(3)$, we see that only one generator remains unbroken, t^3. It generates the $U(1)$ or $SO(2)$ subgroup of $SO(3)$. Thus the symmetry breaking pattern is written, $SO(3) \to U(1)$. This can be shown by performing an exponentiation of,

$$t^3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

We need to calculate the following matrix,

$$\exp(i\theta t^3) = \mathbb{I} + i\theta t^3 + \frac{1}{2!}(i\theta)^2 \left(t^3\right)^2 + ... \tag{4.6.8}$$

$$\begin{aligned} t^3 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ (t^3)^2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = I \\ (t^3)^3 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = t^3 \end{aligned}$$

So for odd powers, n of t^3, $(t^3)^n = t^3$ and for even powers, we have $(t^3)^n = I$. We have,

$$\begin{aligned} \exp(i\theta t^3) &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \left(i\theta + \frac{1}{3!}(i\theta)^3 + ...\right) \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \left(1 + \frac{1}{2!}(i\theta)^2 + ...\right) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \left(\theta - \frac{1}{3!}\theta^3 + ...\right) \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \left(1 - \frac{1}{2!}\theta^2 + ...\right) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

which is a rotation in $\mathbb{R}^3$ but about the x and y plane, the definition of an $SO(2)$ transformation. This confirms our initial impressions about the symmetry breaking pattern, the unbroken generator is indeed that of $U(1)$.

4.6.1 Mass matrix for continuous symmetries

Back to the new set of generators $\hat{T}^A$, the transformation law can be written as,

$$\phi_i \to \phi_i + \delta\phi_i; \quad \delta\phi_i = i\hat{\theta}^A \hat{T}^A_{ij}\phi_j; \quad A \in \{1, ..., d\} \tag{4.6.1}$$

because this is a symmetry, the potential does not change under the transformation,

$$\delta V = \delta\phi_i \frac{\partial V}{\partial \phi_i} = i\hat{\theta}^a \hat{T}^A{}_{ij}\phi_j \frac{\partial V}{\partial \phi_i} = 0 \tag{4.6.2}$$

Differentiating with respect to ϕ_k gives,

$$i\hat{\theta}^A \hat{T}^A{}_{ik} \frac{\partial V}{\partial \phi_i} + i\hat{\theta}^A \hat{T}^A{}_{ij}\phi_j \frac{\partial^2 V}{\partial \phi_k \partial \phi_i} = 0 \tag{4.6.3}$$

Now, we evaluate this guy at $\phi = \phi_0$, using,

$$i\hat{\theta}^A \hat{T}^A{}_{ij}\phi_{0j} m^2_{ki} = \hat{\theta}^A m^2_{ki}\hat{\phi}^A_i = 0 \text{ for any coefficients } \hat{\theta}^A \tag{4.6.4}$$

which implies that,

$$m^2_{ij}\hat{\phi}^A_j = 0 \text{ for all } A \in \{1, ..., d\}. \tag{4.6.5}$$

If $\hat{T}^A$ is unbroken, then $\hat{\phi}^A = 0$, and this equation is satisfied trivially. On the other hand, if $\hat{T}^A$ is broken, then $\hat{\phi}^A \neq 0$, and the result shows that $\hat{\phi}^A$ has to be an eigenvector of the mass matrix m^2_{ij} with zero eigenvalue. In other words, it corresponds to a massless Goldstone particle. This can be summarised as **Goldstone's theorem**: *Every broken generator gives rise to a massless Goldstone particle.*

4.7 Higgs Mechanism

We previously considered that the broken symmetry is global, meaning that the symmetries are position-independent, and it was found that it leads to a massless Goldstone boson. If we consider a local symmetry, we find very different results.

4.7.1 Abelian Higgs Model

For simplicity, let us first consider a theory with a complex scalar field and a local $U(1)$ symmetry,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\phi)^*(D^\mu\phi) - V(\phi^*, \phi); \quad V(\phi^*, \phi) = -\mu^2|\phi|^2 + \frac{\lambda}{2}|\phi|^4 \tag{4.7.1}$$

As usual, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $D_\mu = \partial_\mu + ieA_\mu$.

The vacuum states of the theory are given by $A_\mu = 0$, $|\phi| = \mu/\sqrt{\lambda}$ (or any gauge transformation is this state). Therefore, we expand in the similar way, as for the global case, $\phi = (v + \varphi + i\xi)/\sqrt{2}$ and we would end up with the same potential as the global example,

$$V(v + \phi) = -\frac{\mu^4}{2\lambda} + \mu^2\varphi^2 + \frac{1}{2}\lambda v\varphi\left(\varphi^2 + \xi^2\right) + \frac{1}{8}\lambda\left(\varphi^2 + \xi^2\right)^2 \tag{4.7.2}$$

Now, we consider how this perturbation changes the kinetic terms,

$$D_\mu\phi = \partial_\mu\phi + ieA_\mu\phi = \frac{1}{\sqrt{2}}\left(\partial_\mu\varphi + i\partial_\mu\xi + ievA_\mu + ieA_\mu\varphi - eA_\mu\xi\right) \tag{4.7.3}$$

Therefore, the derivative term in the Lagrangian is nothing but,

$$
\begin{aligned}
(D_\mu\phi)^* D^\mu\phi &= \frac{1}{2}(\partial_\mu\varphi - eA_\mu\xi)^2 + \frac{1}{2}\left(\partial_\mu\xi + evA_\mu + eA_\mu\varphi\right)^2 \\
&= \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - e\left(\partial_\mu\varphi\right)A^\mu\xi + \frac{1}{2}e^2\xi^2 A_\mu A^\mu \\
&+ \frac{1}{2}\partial_\mu\xi\partial^\mu\xi + evA^\mu\partial_\mu\xi + eA^\mu\varphi\partial_\mu\xi \\
&+ \frac{1}{2}e^2v^2 A_\mu A^\mu + e^2 v A_\mu A^\mu\varphi + \frac{1}{2}e^2\varphi^2 A_\mu A^\mu
\end{aligned}
\tag{4.7.4}
$$

Up to quadratic order, the Lagrangian is just,

$$
\begin{aligned}
\mathcal{L} = &-\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) + \frac{1}{2}e^2v^2 A_\mu A^\mu \\
&+ \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + \frac{1}{2}\partial_\mu\xi\partial^\mu\xi + evA^\mu\left(\partial_\mu\xi\right) - \mu^2\varphi^2 + \mathcal{L}_{\text{int}}
\end{aligned}
\tag{4.7.5}
$$

Which is where we bump into a problem because of the $evA^\mu\partial_\mu\xi$, the equations of motion for A_μ and ξ do not decouple and therefore the particle spectrum cannot be immediately determined. To deal with the problem, we note that we are free to perform a gauge transformation. In particular, we can choose $\theta(x) = -\arg\phi(x) \approx -\xi/v$. This makes the field ϕ real and corresponds to a particular gauge choice called the *unitary gauge*. In this gauge, $\xi = 0$, so that Lagrangian becomes,

$$
\mathcal{L} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) + \frac{1}{2}e^2v^2 A_\mu A^\mu + \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \mu^2\varphi^2 + \mathcal{L}_{\text{int}} \tag{4.7.6}
$$

This Lagrangian has only two fields now, A_μ and φ, and it is not gauge invariant, but it is physically equivalent to the original gauge-invariant Lagrangian. It has no mixing term between A_μ and φ, and therefore two fields describe two distinct particles.

The φ terms in the Lagrangian are familiar and describe a massive real scalar field with a mass $m_\varphi = \sqrt{2}\mu$. This is known as the Higgs scalar. To analyse the gauge field part of the Lagrangian, the equation of motion is,

$$
\frac{\partial\mathcal{L}}{\partial A_\rho} - \partial_\sigma\left[\frac{\partial\mathcal{L}}{\partial(\partial_\sigma A_\rho)}\right] = e^2v^2A^\rho - \partial_\sigma\left(-F^{\sigma\rho}\right) = 0. \tag{4.7.7}
$$

which gives,

$$
\partial_\mu F^{\mu\nu} + e^2v^2A^\nu = 0. \tag{4.7.8}
$$

To interpret this equation, we differentiate once more,

$$
\partial_\nu\partial_\mu F^{\mu\nu} + e^2v^2\partial_\nu A^\nu = e^2v^2\partial_\nu A^\nu = 0. \tag{4.7.9}
$$

which implies that $\partial_\mu A^\mu = 0$. This looks like the Lorenz gauge condition, but remember that we had already fixed the unitary gauge. Therefore, here it is the actual physical equation of motion which follows from the equations of motion in the unitary gauge. It has the effect of removing one of the four degrees of freedom that a vector field would otherwise carry and therefore brings the number of degrees of freedom down to three.

Let us now write the equation of motion in terms of A_μ,

$$\partial_\mu \left(\partial^\mu A^\nu - \partial^\nu A^\mu\right) + e^2 v^2 A^\nu = \partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu + e^2 v^2 A^\nu = 0. \tag{4.7.10}$$

and since $\partial_\mu A^\mu = 0$ we are left with,

$$\Box A^\mu + e^2 v^2 A^\mu = 0. \tag{4.7.11}$$

which means that each component of the four-vector A_μ satisfies the massive Klein-Gordon equation with mass $m_\gamma = ev$. Therefore A_μ has become a massive vector particle.

In summary, the particle spectrum of the theory consists of a massive vector field A_μ, which consists of three real components and has mass m_γ, and a single real scalar field, with mass $m_\varphi = \sqrt{2}\mu$, so the total number of real degrees of freedom $3+1=4$. There are no massless particles, neither a massless vector now a massless Goldstone scalar. This way of giving a mass to the vector through spontaneous symmetry breaking is called the Higgs mechanism.

In comparison, when the symmetry is not broken, the particle spectrum consists of a massless vector (photon), which consists of two real degrees of freedom, and a complex scalar, which also consists of two real degrees of freedom. The total number of degrees of freedom is therefore $2+2=4$. As we can see, the total number of degrees of freedom is therefore unchanged by the symmetry breaking, but one of the scalar degrees of freedom (more specifically, the Goldstone mode) is *eaten* by the photon, and allows it to become massive.

4.7.2 Non-Abelian Higgs

Let us now generalise the result of the previous sections to the non-Abelian case. Consider now the theory of an n-component real scalar field $\phi = (\phi_1, \ldots, \phi_n)^T$ with some local gauge symmetry group G and the Lagrangian,

$$\mathcal{L} = -\frac{1}{2}\mathrm{Tr} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2}(D_\mu \phi)^*_i (D^\mu \phi)^i - V(\phi), \tag{4.7.1}$$

where $D_\mu = \partial_\mu + igA_\mu$. We assume that that scalar field has a non-zero vev $\phi = \phi_0 \neq 0$ and that we have used the equations above to construct a set of new generators, $\hat{T}^A$ and the corresponding vectors $\hat{\phi}^A = i\hat{T}^A \phi_0$. We now write the gauge field $A_\mu = A^A_\mu \hat{T}^A$ and the scalar field $\phi = \phi_0 + \varphi(x)$. The covariant derivative now becomes,

$$(D_\mu \phi)_i = \partial_\mu \varphi_i + igA^A_\mu \hat{T}^A_{ij} \phi_{0j} + igA^A_\mu \hat{T}^A_{ij} \varphi_j = \partial_\mu \varphi_i + gA^A_\mu \hat{\phi}^A_i + igA^A_\mu \hat{T}^A_{ij} \varphi_j \tag{4.7.2}$$

and the quadratic part of the derivative term in the Lagrangian is

$$\frac{1}{2}(D_\mu \phi)_i (D^\mu \phi)_i = \frac{1}{2}(\partial_\mu \varphi)_i (\partial^\mu \varphi)_i + g(\partial_\mu \varphi)_i A^A_\mu \hat{\phi}^A_i + \frac{1}{2} A^A_\mu A^{\mu B} \hat{\phi}^A_i \hat{\phi}^B_i + \text{(higher order terms)} \tag{4.7.3}$$

4.8 Electroweak Symmetry Breaking and the Higgs

Now let's see how the $SU(2) \times U(1)$ symmetry is broken and the electric charge emerges. We start with the Higgs field.

$$\Phi = \begin{pmatrix} \varphi^+ \\ \varphi_0 \end{pmatrix} \qquad \begin{array}{c} \text{scalar field } SU(2) \text{ doublet} \\ U(1)_Y : \quad Y = 1 \end{array} \tag{4.8.1}$$

Now the lowest energy state for Φ is:

$$\Phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \qquad \textit{vev: vacuum expectation value} \tag{4.8.2}$$

This breaks the symmetry

$$\text{vacuum} \quad \Phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}v \end{pmatrix} \quad \text{is } \textit{not} \text{ invariant under general } SU(2) \times U(1) \tag{4.8.3}$$

However, is part of $SU(2) \times U(1)$ preserved? What is the stabalizer group?

$$\rho_{U(1)} \otimes \rho_{SU(2)} \Phi = e^{3i\theta} \begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix} \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}v \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}v \end{pmatrix} \tag{4.8.4}$$

if $b = 0$ and $a = e^{3i\theta}$. The unbroken subgroup is then,

$$H = \left\{ e^{i\theta} \in U(1), \quad \begin{pmatrix} e^{3i\theta} & 0 \\ 0 & e^{-3i\theta} \end{pmatrix} \in SU(2) \right\} \simeq U(1)_{\text{EM}} \tag{4.8.5}$$

Chapter 5

Perturbation Theory in Gravity

5.1 Linearized Gravity

We will consider gravity in the weak-field limit where the variations with time of the metric perturbations are small and we place no restrictions on particle speeds. By *weak* we mean that the metric can be expanded as a linear sum of the Minkowski metric and a small perturbation. Much like the spirit of perturbation analysis in Quantum Mechanics, we introduce the *book-keeping parameter*, ε, which we will assume to be small. We will assume that we can expand our quantities of interest in the following form,

$$Q = {}^{(0)}Q + \varepsilon \, {}^{(1)}Q \tag{5.1.1}$$

Assuming also a flat background spacetime, ${}^{(0)}Q$ will represent the contributions in the limit of special relativity and the first order contribution from gravitational perturbations to Q is ${}^{(1)}Q$. At the end of all calculations, we will set $\varepsilon = 1$. For example, the metric is written perturbatively as,

$$g_{\mu\nu} = \eta_{\mu\nu} + \varepsilon h_{\mu\nu} : \varepsilon \ll 1. \tag{5.1.2}$$

Similarly, the inverse metric is given by,

$$g^{\mu\nu} = \eta^{\mu\nu} - \varepsilon h^{\mu\nu} : \varepsilon \ll 1. \tag{5.1.3}$$

which can be shown by multiplying (5.1.2) and (5.1.3), and ignoring $\mathcal{O}(\varepsilon^2)$ terms, does indeed spits out a Kronecker Delta tensor. This leads to the conclusion that, we use the $\eta^{\mu\nu}$ and $\eta_{\mu\nu}$ to raise and lower indices respectively. Think of this whole limit as a linearized version of GR describing the theory of *symmetric tensor field*, $h_{\mu\nu}$ propagating on a flat background spacetime. This theory is manifestly Lorentz invariant,

$$h_{\mu'\nu'} = \Lambda_{\mu'}{}^{\mu} \, \Lambda_{\nu'}{}^{\nu} h_{\mu\nu}.$$

[Could it be Poincare invariant?] where $\Lambda \in SO^{+}(1,3)$. To note, we could have chosen a curved background, however, this would introduce an additional level of complexity which we should just avoid for now.

We want the equations of motion for this perturbation, $h_{\mu\nu}$. We shall proceed with calculating the important quantities in General Relativity up to leading order contributions in ε. Begin with the Christoffel symbols up to linear order in ε,

$$\begin{aligned}\Gamma^{\rho}_{\mu\nu} &= {}^{(0)}\Gamma^{\rho}_{\mu\nu} + \varepsilon\ {}^{(1)}\Gamma^{\rho}_{\mu\nu} \\ &= \frac{1}{2}\eta^{\rho\sigma}\left(h_{\mu\sigma,\nu} + h_{\nu\sigma,\mu} - h_{\mu\nu,\sigma}\right).\end{aligned} \tag{5.1.4}$$

Naturally, we don't expect a zero-order contribution because flat spacetime has vanishing connection coefficients in inertial coordinates, $(t, \vec{x})$. Next, we calculate the Riemann tensor, and note that the Γ^2 terms can be ignored (because they contain terms quadratic in $\mathcal{O}(\varepsilon^2)$),

$$\begin{aligned}R^{\rho}{}_{\mu\lambda\nu} &= {}^{(0)}R^{\rho}{}_{\mu\lambda\nu} + \varepsilon\ {}^{(1)}R^{\rho}{}_{\mu\lambda\nu} \\ &= \varepsilon\left(\partial_\lambda\ {}^{(1)}\Gamma^{\rho}{}_{\nu\mu} - \partial_\nu\ {}^{(1)}\Gamma^{\rho}{}_{\lambda\mu}\right) + \mathcal{O}(\varepsilon^2) \\ &= 2\partial_{[\lambda}\Gamma^{\rho}{}_{\nu]\mu} \\ R_{\mu\nu\rho\sigma} &= \eta_{\mu\lambda}\partial_\rho\Gamma^{\lambda}{}_{\nu\sigma} - \eta_{\mu\lambda}\partial_\sigma\Gamma^{\lambda}{}_{\nu\rho} \\ &= \frac{1}{2}\left(\partial_\rho\partial_\nu h_{\mu\sigma} + \partial_\sigma\partial_\mu h_{\nu\rho} - \partial_\rho\partial_\mu h_{\nu\sigma} - \partial_\sigma\partial_\nu h_{\mu\rho}\right).\end{aligned} \tag{5.1.5}$$

Again, $\mathcal{O}(\varepsilon)$ are the lowest-order of terms that appear in the expression. Next up is the Ricci tensor,

$$\begin{aligned}R_{\mu\nu} &= {}^{(0)}R_{\mu\nu} + \varepsilon\ {}^{(1)}R_{\mu\nu} \\ &= 2\partial_{[\alpha}\Gamma^{\alpha}{}_{\nu]\mu}, \\ &= \frac{1}{2}\left(\partial_\nu\partial^\sigma h_{\sigma\mu} + \partial_\mu\partial^\sigma h_{\sigma\nu} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h\right),\end{aligned} \tag{5.1.6}$$

where the $h = \eta^{\mu\nu}h_{\mu\nu}$ is the trace of the perturbation metric and $\Box = -\partial_t^2 + \nabla^2$, is the D'Alembertian operator on the flat background spacetime. The Ricci scalar, R, is

$$\begin{aligned}R &= R^{(0)} + \varepsilon R^{(1)} \\ &= 2\eta^{\mu\nu}\partial_{[\alpha}\Gamma^{\alpha}{}_{\nu]\mu}\end{aligned} \tag{5.1.7}$$

From all of these, we can derive the Einstein Tensor,

$$\begin{aligned}G_{\mu\nu} &\equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} \\ &= G^{(0)}_{\mu\nu} + \varepsilon G^{(1)}_{\mu\nu} = 2\partial_{[\alpha}\Gamma^{\alpha}{}_{\nu]\mu} - \eta_{\mu\nu}\eta^{\rho\sigma}\partial_{[\alpha}\Gamma^{\alpha}{}_{\sigma]\rho}. \\ &= \frac{1}{2}\left(\partial_\nu\partial^\sigma h_{\sigma\mu} + \partial_\mu\partial^\sigma h_{\sigma\nu} - \Box h_{\mu\nu} - \partial_\mu\partial_\nu h\right) - \frac{1}{2}\eta_{\mu\nu}\left(\partial_\rho\partial_\sigma h^{\rho\sigma} - \Box h\right) \\ &= \partial^\sigma\partial_{(\mu}h_{\nu)\sigma} - \frac{1}{2}\left(\Box h_{\mu\nu} + \partial_\mu\partial_\nu h + \eta_{\mu\nu}\partial_\rho\partial_\sigma h^{\rho\sigma} - \eta_{\mu\nu}\Box h\right)\end{aligned} \tag{5.1.8}$$

This equation of motion can be derived from the following action,

$$S = \int d^4x \frac{1}{2}\left[(\partial_\mu h^{\mu\nu})(\partial_\nu h) - (\partial_\mu h^{\rho\sigma})(\partial_\rho h^{\mu}{}_{\sigma}) + \frac{1}{2}\eta^{\mu\nu}(\partial_\mu h^{\rho\sigma})(\partial_\nu h_{\rho\sigma}) - \frac{1}{2}\eta^{\mu\nu}(\partial_\mu h)(\partial_\nu h)\right] \tag{5.1.9}$$

This is all interesting, but where does this action come from? Let's chuck in the Ricci scalar into the Einstein-Hilbert action and see what we get. The Einstein-Hilbert action is given by,

$$S_{\mathrm{EH}} = \int d^4x\ \eta^{\mu\nu}\left[\partial_\alpha \Gamma^\alpha{}_{\nu\mu} - \partial_\nu \Gamma^\alpha{}_{\alpha\mu}\right] \tag{5.1.10}$$

In GR, the Einstein tensor is *divergence-free*, $\nabla_\mu G^{\mu\nu} = 0$, which hearkens back to $\partial_\mu F^{\mu\nu}$ from free Electromagnetism? But what would that even look like?

$$\begin{aligned}\partial_\mu G^{\mu\nu} &= \frac{1}{2}\left(\partial_\sigma \Box h^{\nu\sigma} + \partial_\mu \partial_\sigma \partial^\nu h^{\mu\sigma} - \partial_\mu \Box h^{\mu\nu} - \partial^\nu \Box h + \partial^\nu \partial_\rho \partial_\sigma h^{\rho\sigma} + \partial^\nu \Box h\right) \\ &= 0.\end{aligned} \tag{5.1.11}$$

5.2 Classifying the irreducible representations of the proper orthochronous Lorentz Group

A d-dimensional representation of the Lorentz group, where Λ is a Lorentz transformation, which means,

- $\det(D(\Lambda)) > 0$ (*proper*), $[D(\Lambda)]^0{}_0 > 0$ (*orthochronous*),
- $[D(\Lambda_1)][D(\Lambda_2)] = [D(\Lambda_1\Lambda_2)] = [D(\Lambda_3)]$, where $\Lambda_k \in SO^+(1,3)$.
- $[D(\Lambda^{-1})] = [D(\Lambda)]^{-1}$
- $[D(\Lambda)]^T \eta [D(\Lambda)] = \eta$

Choose a matrix representation, what happens to the $h_{\mu\nu}(x)$ tensor field under an infinitesimal Lorentz transformation?

$$[D(\Lambda)]^\mu{}_\nu = \left(e^{-i\theta_{\alpha\beta}\mathcal{M}^{\alpha\beta}}\right)^\mu{}_\nu \tag{5.2.1}$$

If we define the translation operator $\hat{P}_\mu = -i\partial_\mu$, we can write,

$$\Lambda^\mu_{\mu'}\Lambda^\nu_{\nu'} h_{\mu\nu}\left(\Lambda^{-1}x\right) = h_{\mu'\nu'}(x) + \theta_{\alpha\beta}\left(\mathcal{M}^{\alpha\beta}\right)^\rho{}_\sigma x^\sigma \hat{P}_\rho h_{\mu'\nu'} - 2i\theta_{\alpha\beta}\left(\mathcal{M}^{\alpha\beta}\right)^\mu{}_{(\mu'} h_{\nu')\mu} \tag{5.2.2}$$

There is a standard representation of the Lorentz generators give by,

$$\left(\mathcal{M}^{\alpha\beta}\right)^\rho{}_\sigma = i\eta^{\mu\rho}\left(\delta^\alpha_\mu \delta^\beta_\sigma - \delta^\alpha_\sigma \delta^\beta_\mu\right) \tag{5.2.3}$$

Using this representation of the Lorentz algebra, we obtain a new representation of the Lorentz algebra for covariant symmetric 2-tensors,

$$\left(\mathcal{J}^{\alpha\beta}\right)^{\rho\sigma}{}_{\mu\nu} = \left(x^\alpha p^\beta - x^\beta p^\alpha\right)\delta^\rho{}_\mu \delta^\sigma{}_\nu + \left(\mathcal{M}^{\alpha\beta}\right)^\sigma{}_\nu \delta^\rho{}_\mu + \left(\mathcal{M}^{\alpha\beta}\right)^\rho{}_\mu \delta^\sigma{}_\nu. \tag{5.2.4}$$

Does this new representation obey the Lorentz Algebra? Yes, because the first part involving x^α

and p^β is the standard contribution to the Lorentz algebra from fields, and the second part involves the standard Lorentz algebra (**?**), since $\mathfrak{so}^+(3,1)$ is a vector space, then a vector sum of elements of $\mathfrak{so}^+(3,1)$ are also element of the Lorentz algebra, though this forms a higher-dimensional tensor product representation of the Lorentz algebra. $\mathcal{J}^{0i}$ are the three boosts,

$$\left(\mathcal{J}^{0i}\right)^{\rho\sigma}{}_{\mu\nu} = -i\left(t\frac{\partial}{\partial x^i} - x^i\frac{\partial}{\partial t}\right)\delta^\rho{}_\mu\delta^\sigma{}_\nu + \left(\mathcal{K}^i\right)^\sigma{}_\nu\,\delta^\rho{}_\mu + \left(\mathcal{K}^i\right)^\rho{}_\mu\,\delta^\sigma{}_\nu. \tag{5.2.5}$$

and $\mathcal{J}^{ij}$ form representations of the rotation group,

$$\begin{aligned}\frac{1}{2}\epsilon_{ijk}\left(\mathcal{J}^{ij}\right)^{\rho\sigma}{}_{\mu\nu} &= \frac{1}{2}\epsilon_{ijk}\left(x^ip^j - x^jp^i\right)\delta^\rho{}_\mu\delta^\sigma{}_\nu + \left(\frac{1}{2}\epsilon_{ijk}\mathcal{M}^{ij}\right)^\sigma{}_\nu\delta^\rho{}_\mu + \left(\frac{1}{2}\epsilon_{ijk}\mathcal{M}^{ij}\right)^\rho{}_\mu\delta^\sigma{}_\nu \\ &= (\epsilon_{ijk}x^ip^j)\delta^\rho{}_\mu\delta^\sigma{}_\nu + \left(L^k\right)^\sigma{}_\nu\delta^\rho{}_\mu + \left(L^k\right)^\rho{}_\mu\delta^\sigma{}_\nu = \left(\mathcal{L}^k\right)^{\rho\sigma}{}_{\mu\nu}\end{aligned} \tag{5.2.6}$$

Therefore, we can write,

$$\left(\mathcal{L}^i\right)^{\rho\sigma}{}_{\mu\nu} = (-i\mathbf{r}\times\vec{\nabla})^k\delta^\rho{}_\mu\delta^\sigma{}_\nu + \left(L^i\right)^\sigma{}_\nu\delta^\rho{}_\mu + \left(L^i\right)^\rho{}_\mu\delta^\sigma{}_\nu \tag{5.2.7}$$

Let's focus on the intrinsic part of the rotation algebras:

$$\begin{aligned}(\mathcal{K}^i)^{\rho\sigma}{}_{\mu\nu} &= (K^i)^\rho{}_\nu\delta^\sigma{}_\mu + (K^i)^\sigma_\mu\delta^\rho{}_\nu \\ (\mathcal{L}^i)^{\rho\sigma}{}_{\mu\nu} &= (L^i)^\rho{}_\nu\delta^\sigma{}_\mu + (L^i)^\sigma{}_\mu\delta^\rho{}_\nu\end{aligned} \tag{5.2.8}$$

Now, we will complexify the algebra in the following way,

$$\begin{aligned}\mathcal{J}^i_\pm &= \frac{1}{2}\left(\mathcal{L}^i + i\mathcal{K}^i\right) \\ (\mathcal{J}^i_\pm)^{\rho\sigma}{}_{\mu\nu} &= \frac{1}{2}\left((J^i_\pm)^\rho{}_\nu\delta^\sigma{}_\mu + (J^i_\pm)^\sigma{}_\mu\delta^\rho{}_\nu\right)\end{aligned} \tag{5.2.9}$$

where the $J_\pm$ obey the following commutation relations:

$$\begin{aligned}[J^i_\pm, J^j_\pm] &= i\epsilon_{ijk}J^k_\pm \\ [J^i_+, J^j_-] &= 0\end{aligned} \tag{5.2.10}$$

It's easy to see how the $(\mathcal{J}^i_\pm)$ obey similar commutation relations. In this manner, we see the usual isomorphism, $\mathfrak{so}(3,1)_\mathbb{C} \simeq \mathfrak{su}(2)_\mathbb{C} \oplus \mathfrak{su}(2)_\mathbb{C}$. Alongside there is a non-trivial **Abelian** subalgebra.

5.3 Dynamical Degrees of Freedom

With this linearised tensor we could immediately fix a gauge and solve the EFEs. However, we can accumulate some additional physics insights by first choosing a fixed inertial coordinate system in the Minkowski background spacetime and decomposing components of the metric perturbation according to their transformation properties under spatial rotations.

5.3.1 Brief aside: Degrees of freedom in EM

The idea of this line of inquiry is to make the analogy to electromagnetism more apparent. Now we will take a brief aside to the covariant formulation of electromagnetism in order to analogise subsequent steps in our analysis of metric degrees of freedom.

Of course, the discovery of electromagnetic fields was made in reverse of what we are trying to do with gravity. Traditionally, we start with the degrees of freedom $\vec{E}$ and $\vec{B}$ and construct the field strength tensor, $F_{\mu\nu}$. $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$, where we define a one-form potential $A_\mu = \left(\phi, \vec{A}\right)$, where ϕ is the *electrostatic potential* and $\vec{A}$ is the *magnetic vector potential*. Combinations of derivatives of components of A_μ produce our physical fields,

$$\begin{aligned} \vec{E} &\equiv -\vec{\nabla}\phi - \frac{\partial \vec{A}}{\partial t}, \\ \vec{B} &\equiv \vec{\nabla} \times \vec{A}. \end{aligned} \tag{5.3.1}$$

A particle of charge q moving through an electromagnetic field with velocity $\vec{v}$ obeys the Lorentz force law,

$$\frac{d\vec{p}}{dt} = q\left(\vec{E} + \vec{v} \times \vec{B}\right). \tag{5.3.2}$$

In a vacuum (source-free Maxwell) the dynamics of the theory are generated purely by the following Lagrangian,

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu},$$

which leads to the following equation of motion,

$$\partial_\rho F^{\rho\mu} = \Box A^\mu - \partial^\mu(\partial_\rho A^\rho) = 0. \tag{5.3.3}$$

But, we make note of the gauge freedom $A'_\mu = A_\mu + \partial_\mu \Lambda(x)$, where $\Lambda(x)$ is some scalar function. Vector potentials that differ by a gradient of a scalar are said to be equivalent, $A'_\mu \sim A_\mu$ and $\sim$ is an equivalence relation. We could go further, however, at this step it is customary to fix a particular *gauge*.

Coulomb Gauge

The **Coulomb gauge** (also known as the transverse gauge) is a popular gauge fixing condition. Coulomb gauge states,

$$\vec{\nabla} \cdot \vec{A} = 0 \tag{5.3.4}$$

which forces the spatial degrees of freedom to have vanishing 3-divergence. It's definitely an interesting choice of gauge, however, the condition itself breaks Lorentz Invariance, so we could care less about it.

Lorenz Gauge

On the other hand, the **Lorenz gauge** is a Lorentz invariant and the condition is written,

$$\partial_\mu A^\mu = 0. \tag{5.3.5}$$

Watch what this does to the gauge freedom equation,

$$\partial_\mu A^{'\mu} = \partial_\mu A^\mu + \Box\Lambda(x) = 0. \tag{5.3.6}$$

But also, we require that $\partial_\mu A^\mu = 0$, therefore, we have an equation of motion for the $\Lambda(x)$ degree of freedom,

$$\Box\Lambda(x) = 0,$$

which is an equation of motion for a massless spin-0 field, called the *residual gauge field*. What about (5.3.3)? We see that the second term contains a $\partial_\rho A^\rho$, which we have chosen to vanish, therefore,

$$\Box A^\mu = 0.$$

This is an interesting equation, because it's four spin-0 fields propagating independently (this is exactly the behaviour of electromagnetic waves which obey the *principle of superposition*). We're almost there, however, we have two too many degrees of freedom. Electromagnetic radiation carries two transverse polarisation vectors (vectors that are perpendicular to the direction of motion). Luckily, the gauge condition can help us here.

Solving equations in Lorenz gauge

Consider the plane wave solution,

$$A^\mu = \epsilon^\mu e^{-ik\cdot x}, \tag{5.3.7}$$

then, from $\Box A^\mu$, we have $k^2 = 0$, a null wavevector. Further, from the Lorenz gauge, we have $\partial_\mu A^\mu \to \epsilon \cdot k = 0$, which means that the polarisation vector is normal to the propagation direction. From arguments of gauge choice, we have been able to accurately describe the nature of electromagnetic radiation. Now we choose a particular frame in which light is propagating along the z-direction. The natural choice from the $k^2 = 0$ condition for k^μ is,

$$k^\mu = (k\ 0\ 0\ k)^T. \tag{5.3.8}$$

The polarisation vector ϵ^μ can be ascertained as,

$$\epsilon^\mu = (\alpha\ e^1\ e^2\ \alpha)^T \tag{5.3.9}$$

and we see $\epsilon \cdot k = 0$. Thus we interpret α as the residual gauge mode, $\Lambda = e^{-ik\cdot x}$

$$A^\mu = \partial^\mu \Lambda = \partial^\mu e^{-ik\cdot x} == -ik^\mu e^{-ik\cdot x} = -ik \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} e^{-ik\cdot x} \tag{5.3.10}$$

where $\alpha = -ik$. So gauge choices result in constraints being places on the polarisation degrees of freedom.

In order to write the general classical solution we introduce the 4 orthonormal **real** polarisation vectors $\epsilon^\mu_{(\alpha)}(\vec{k})$ that provide a basis spannign Minkowski space. We choose these basis vectors so that $\epsilon^\mu_{(0)}(\vec{k})$ is time-directed and $\vec{\epsilon}_{(3)}(\vec{k}) \propto \vec{k}$ id longitudinal. By orthonormality in the Minkowski context, we mean,

$$\eta_{\alpha\beta} = \epsilon^\mu_{(\alpha)}(\vec{k})\epsilon_{\mu(\beta)}(\vec{k}) \tag{5.3.11}$$

and to be a complete basis we require,

$$\eta^{\mu\nu} = \sum_\alpha \frac{\epsilon^\mu_{(\alpha)}(\vec{k})\epsilon^\nu(\alpha)(\vec{k})}{\epsilon_{(\alpha)}(\vec{k})\cdot\epsilon_{(\alpha)}(\vec{k})} \tag{5.3.12}$$

5.4 Back to linearized gravity

Subsequent to beating this dead horse of an analogy, we move back to the case at hand of linearized gravity. The metric perturbations $h_{\mu\nu}$ is a symmetric tensor, as opposed to the antisymmetric $F_{\mu\nu}$. We can decompose this tensor into its constituent degrees of freedom. In particular, the h_{00} component is a scalar, the h_{0i} components form a vector in $\mathbb{R}^3$ and the spatial h_{ij} components form a two index symmetric spatial tensor. The tensor is further reducible into *trace* and *traceless* parts. In mathematical form, we usually choose,

$$\begin{aligned} h_{00} &= -2\Phi \\ h_{0i} &= w_i \\ h_{ij} &= 2\left(s_{ij} - \Psi\delta_{ij}\right) \end{aligned} \tag{5.4.1}$$

where Ψ contains information about the trace of h_{ij} and s_{ij} is traceless. The trace of h_{ij} is the contraction of h_{ij} with the metric on $\mathbb{R}^3$ which is δ^{ij},

$$\begin{aligned} \delta^{ij}h_{ij} &= 2\delta^{ij}s_{ij} - 2\Psi\delta^{ij}\delta_{ij} = -6\Psi \\ \Psi &= -\frac{1}{6}\delta^{ij}h_{ij} \end{aligned}$$

which leads to the form of s_{ij},

$$s_{ij} = \frac{1}{2}\left[h_{ij} - \frac{1}{3}\delta_{ij}\delta^{kl}h_{kl}\right]. \tag{5.4.2}$$

The full metric $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu$ is,

$$\mathrm{d}s^2 = -(1+2\Phi)\mathrm{d}t^2 + w_i(\mathrm{d}t\mathrm{d}x^i + \mathrm{d}x^i\mathrm{d}t) + [(1-2\Psi)\delta_{ij} + 2s_{ij}]\mathrm{d}x^i\mathrm{d}x^j.$$

As of yet, this is only convenient window-dressing. We have chosen a representation of $h_{\mu\nu}$. This choice will make apparent s_{ij} containing information about *gravitational radiation*. But we have fixed an inertial frame for convenience, to express $p^\mu = dx^\mu/d\lambda$ ($\lambda = \tau/m$ if the particle is massive). As usual,

$$p^0 = \frac{dt}{d\lambda} = E \quad \text{and} \quad p^i = Ev^i$$

Consider the geodesic equation,

$$\frac{dp^\mu}{d\lambda} + \Gamma^\mu{}_{\rho\sigma} p^\rho p^\sigma = 0.$$

We can write this in the following form,

$$\frac{dp^\mu}{dt} = -\Gamma^\mu{}_{\rho\sigma} \frac{p^\rho p^\sigma}{E}$$

The $\mu = 0$ term describes the rate of change in energy over time (power),

$$\frac{dE}{dt} = -E\left[\partial_0 \Phi + 2(\partial_k \Phi)v^k - \left(\partial_{(j} w_{k)} - \frac{1}{2}\partial_0 h_{jk}\right) v^j v^k\right]$$

The spatial component $\mu = i$ of the geodesic equation becomes,

$$\frac{dp^i}{dt} = -E\left[\partial_i \Phi + \partial_0 w_i + 2\left(\partial_{[i} w_{j]} + \partial_0 h_{ij}\right) v^j + \left(\partial_{(j} h_{k)i} - \frac{1}{2}\partial_i h_{jk}\right) v^j v^k\right].$$

To interpret this physically, it is convenient to define the *gravitoelectric* and *gravitomagnetic* vector fields, respectively,

$$\begin{aligned} \vec{G} &= -\nabla \Phi - \frac{\partial \vec{w}}{\partial t}, \\ \vec{H} &= \nabla \times \vec{w} \end{aligned} \tag{5.4.3}$$

These field definitions parallel the equations for the fields in electromagnetism (5.3.1). The net force acting on a particle is given by,

$$\frac{dp^i}{dt} = E\left(G^i + (v \times H)^i - 2(\partial_0 h_{ij})v^j + \left(\partial_{(j} h_{k)i} - \frac{1}{2}\partial_i h_{jk}\right) v^j v^k\right) \tag{5.4.4}$$

This is analogous to the Lorentz Force Law (5.3.2) in EM. We should also examine the evolution of the linearized Einstein Field Equations. The Riemann tensor components are,

$$\begin{aligned} R_{0j0l} &= \partial_j \partial_l \Phi + \partial_0 \partial_{(j} w_{l)} - \frac{1}{2}\partial_0 \partial_0 h_{jl} \\ R_{0jkl} &= \partial_j \partial_{[k} w_{l]} - \partial_0 \partial_{[k} h_{l]j} \\ R_{ijkl} &= \partial_j \partial_{[k} h_{l]i} - \partial_i \partial_{[k} h_{l]j}. \end{aligned}$$

The Ricci tensor components are,

$$\begin{aligned} R_{00} &= \nabla^2\Phi + \partial_0\partial_k w^k + 3\partial_0^2\Psi \\ R_{0j} &= -\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j\partial_k w^k + 2\partial_0\partial_k\Psi + \partial_0\partial_l s_j^k \\ R_{ij} &= -\partial_i\partial_j(\Phi-\Psi) - \partial_0\partial_{(i}w_{j)} + (\Box\Psi)\delta_{ij} + \Box s_{ij} \\ &\quad +2\partial_k\partial_{(i}s_{j)}^k \end{aligned}$$

Finally, we can calculate Einstein tensor,

$$\begin{aligned} G_{00} &= 2\nabla^2\Psi + \partial_k\partial_l s^{kl} \\ G_{0j} &= -\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j\partial_k w^k + 2\partial_0\partial_j\Psi + \partial_0\partial_k s_j^k \\ G_{ij} &= (\delta_{ij}\nabla^2 - \partial_i\partial_j)(\Phi-\Psi) + \delta_{ij}\partial_0\partial_k w^k - \partial_0\partial_{(i}w_{j)} \\ &\quad + 2\delta_{ij}\partial_0^2\Psi - \Box s_{ij} + 2\partial_k\partial_{(i}{s_{j)}}^k - \delta_{ij}\partial_k\partial_l s^{kl}. \end{aligned} \tag{5.4.5}$$

As the EFEs are give by $G_{\mu\nu} = 8\pi G T_{\mu\nu}$, we can write the Einstein tensor in terms of matter content, the 00 equation is,

$$\nabla^2\Psi = 4\pi G T_{00} - \frac{1}{2}\partial_i\partial_j s^{ij}$$

which is an equation for which there are no time derivatives of Ψ, T_{00} and s_{ij} are sufficient to express Ψ up to boundary condition. As it completely depends on other variables, Ψ is not a propagating degree of freedom. Next, the $0i$ contributions,

$$\left(\delta_{jk}\nabla^2 - \partial_j\partial_k\right) w^k = -16\pi G T_{0j} + 4\partial_0\partial_j\Psi + 2\partial_0\partial_k {s_j}^k$$

where there are also no time derivatives of w^i which means the components of the three vector are also not physical degrees of freedom.

Finally, the ij component equation gives,

$$\begin{aligned} \left(\delta_{ij}\nabla^2 - \partial_i\partial_j\right)\Phi &= 8\pi G T_{ij} + (\delta_{ij}\nabla^2 - \partial_i\partial_j - 2\delta_{ij}\partial_0^2)\Psi \\ &\quad -\delta_{ij}\partial_0\partial_k w^k + \partial_0\partial_{(i}w_{j)} + \Box s_{ij} - 2\partial_k\partial_{(i}{s_{j)}}^k \\ &\quad -\delta_{ij}\partial_k\partial_l s^{kl}, \end{aligned}$$

where once again, we have no time derivative of Φ, therefore, it is not a propagating degree of freedom. Propagating tensor fields under quantisation give rise to particles of different spins. Despite the fact that some of the fields in this analysis were not physical degrees of freedom, in some alternative theories of gravity, it is possible for these **residual gauge modes** to become physical. In this case, Ψ and Φ would have spin-0, w^i would have spin-1 and s_{ij} would have spin-2. The only physical degree of freedom in regular GR is the **strain tensor**, s_{ij} transforms under a spin-2 representation. The gauge transformation on the perturbation is described as the following,

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \xi_\nu + \partial_\nu \xi_\mu$$

The infinitesimal change in $h_{\mu\nu}$ can be written in terms of the Lie derivative of the Minkowski metric,

$$h^{(\varepsilon)}_{\mu\nu} = h_{\mu\nu} + \varepsilon \mathcal{L}_\xi \eta_{\mu\nu} \tag{5.4.6}$$

We will set $\varepsilon = 1$ and think of ξ itself as small. Under this transformation, the fields of the perturbation will transform as,

$$\Phi \quad \to \quad \Phi + \partial_0 \xi^0 \tag{5.4.7}$$
$$w_i \quad \to \quad w_i + \partial_0 \xi_i - \partial_i \xi^0 \tag{5.4.8}$$
$$\Psi \quad \to \quad \Psi - \frac{1}{3}\partial_i \xi^i \tag{5.4.9}$$
$$s_{ij} \quad \to \quad s_{ij} + \partial_{(i}\xi_{j)} - \frac{1}{3}\partial_k \xi^k \delta_{ij} \tag{5.4.10}$$

Now we shall discuss some gauge choices.

5.4.1 Choices of gauge

Transverse gauge (Coulomb-like gauge)

As in the Electromagnetism analogy, we choose a gauge which sets the divergence of the strain tensor to zero (analogous to the Coulomb gauge),

$$\partial_i s^{ij} = 0.$$

This can be imposed from (5.4.10) by choosing,

$$\nabla^2 \xi^j + \frac{1}{3}\partial_j \partial_i \xi^i = -2\partial_i s^{ij}$$

But the value of ξ^0 is still undetermined, but we can impose a condition on it by choosing,

$$\partial_i w^i = 0.$$

Similarly, we can use (5.4.8) to allow this,

$$\nabla^2 \xi^0 = \partial_i w^i + \partial_0 \partial_i \xi^i$$

It is called the transverse gauge because, in momentum space, the *polarisations* of these degrees of freedom are orthogonal to the propagation direction. If we expand $\partial_i w^i$ in terms of Fourier modes, we have

$$\partial_i w^i = i \int \frac{d^3k}{(2\pi)^3} [k_i \gamma^i(k)] e^{ikx} = 0$$

which leads to the condition that $\vec{k} \cdot \vec{\gamma} = 0$, where $\vec{\gamma}$ is the set of *polarisation* degrees of freedom of $\vec{w}$. Similarly, $k_i \sigma^{ij}(k) = 0$, where σ^{ij} is the polarisation mode tensor for s_{ij}.

Under this gauge choice, the Einstein Equations become,

$$\begin{aligned}
8\pi G T_{00} &= 2\nabla^2 \Psi \\
8\pi G T_{0j} &= \frac{1}{2}\left[4\partial_0 \partial_j \Psi - \nabla^2 w_j\right] \\
8\pi G T_{ij} &= \left[\delta_{ij}\nabla^2 - \partial_i \partial_j\right](\Phi - \Psi) - \partial_0 \partial_{(i} w_{j)} + 2\delta_{ij}\partial_0^2 \Psi - \Box s_{ij}.
\end{aligned}$$

which are considerably simpler than the equations before.

Synchronous Gauge

Another choice of gauge is the *synchronous gauge*, it can be thought of the gravitational analogue of the temporal gauge in Electromagnetism (vanishing electrostatic potential $A^0 = 0$), since this removes the non-spatial components of the perturbation metric. We begin by setting $\Phi = 0$, which can be achieved by setting,

$$\partial_0 \xi^0 = -\Phi. \tag{5.4.1}$$

Furthermore, we are free to choose ξ^i by setting the vector degree of freedom to zero, $w^i = 0$, this is achieved by making the choice,

$$\partial_0 \xi^i = -w^i + \partial_i \xi^0. \tag{5.4.2}$$

The metric in this synchronous gauge takes on a simple and attractive form,

$$ds^2 = -dt^2 + (\delta_{ij} + h_{ij})dx^i dx^j$$

The Einstein Equations in the synchronous gauge are,

$$\begin{aligned}
8\pi G T_{00} &= \partial_i \partial_j s^{ij} + 2\nabla^2 \Psi \\
8\pi G T_{0j} &= 2\partial_0 \partial_j \Psi + \partial_0 \partial_k s_j{}^k \\
8\pi G T_{ij} &= 2\partial_k \partial_{(i} s_{j)}{}^k + \delta_{ij}\partial_k \partial_l s^{kl} - \Box s_{ij} \\
&\quad + \left(2\delta_{ij}\partial_0^2 + \partial_i \partial_j - \delta_{ij}\nabla^2\right)\Psi.
\end{aligned}$$

These are equations of the spatial parts of the metric perturbation h_{ij}.

5.5 Newtonian Fields and Photon Trajectories

We will extend the definition of the Newtonian limit. Relativistic particles respond to spatial components of the metric as well. We can model static gravitating sources by dust, a perfect fluid for which the pressure vanishes. Most of the matter in the universe is well approximated by dust, stars, planets, galaxies and dust. We work in the rest frame of the dust, where the energy-momentum tensor takes the form, $T_{\mu\nu} = \rho U_\mu U_\nu$.

Since the background is Minkowski space, it is straightforward to accomodate moving sources by performing a Lorentz transformation into their rest frame. We are unable to deal with multiple sources moving at large relativistic velocities. Turning to the Einstein equations in the transverse gauge and static sources (time derivatives vanishing), we have,

$$\begin{aligned} \nabla^2\Psi &= 4\pi G\rho \\ \nabla^2 w_j &= 0 \\ \nabla^2 s_{ij} - (\delta_{ij}\nabla^2 - \partial_i\partial_j)(\Phi - \Psi) &= 0 \end{aligned} \tag{5.5.1}$$

Since we are looking for both non-singular and well-behaved solutions at infinity, only the fields that are sourced at the right-hand side of (5.5.1) will be non-vanishing. For example, immediately by taking the trace of the third equation, we obtain,

$$2\nabla^2(\Phi - \Psi) = 0, \tag{5.5.2}$$

which enforces the equality between the scalar modes, $\Phi = \Psi$. If this is the case, we can see that the scalar potential obeys the Poisson equation, as is the case for the usual Newtonian limit. We can further simplify the third equation,

$$\nabla^2 s_{ij} = 0, \tag{5.5.3}$$

which implies that $s_{ij} = 0$ because there is no source. The perturbed metric in the Newtonian limit is simply,

$$ds^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)(dx^2 + dy^2 + dz^2). \tag{5.5.4}$$

Now we consider the path of a photon through this geometry. Essentially, we are solving the null geodesic equations for the *perturbed metric* which is denoted by $x^\mu(\lambda)$. The geodesics can be decomposed into the geodesic of the flat background plus a perturbation,

$$x^\mu(\lambda) = x^{(0)\mu} + \varepsilon x^{(1)\mu}, \tag{5.5.5}$$

where, $x^{(0)\mu}$ are the null geodesics on Minkowski space and again we employ the bookkeeping parameter ε and expand all quantities up to first order in ε followed by setting ε to unity. We then evaluate all quantities along the background metric to solve for $x^{(1)\mu}$. For this, we need to assume Φ does not change a lot along background geodesics; this amounts to requiring that $x^{(1)i}\partial_i\Phi \ll \Phi$. If we consider only very short paths, the deviations will necessarily be small, and our approximation is still valid. But we can assemble larger paths out of small segments. As a result we derive true equations, but the paths we integrate will be the actual path x^μ, rather than $x^{(0)\mu}(\lambda)$. As long as this is understood, out results will be valid for any trajectories in the perturbed spacetime.

For convenience, we denote the wave vector of the background path as k^μ, and the derivative of the deviation vector as l^μ,

$$k^\mu = \frac{dx^{(0)\mu}}{d\lambda} \quad l^\mu = \frac{dx^{(1)\mu}}{d\lambda}. \tag{5.5.6}$$

The condition that a path is null,

$$g_{\mu\nu}\frac{dx^\mu}{d\lambda}\frac{dx^\nu}{d\lambda} = 0, \tag{5.5.7}$$

which must be solved order-by-order. At zeroth order, we have the wave vector dispersion, $\eta_{\mu\nu}k^\mu k^\nu = 0$, or

$$(k^0)^2 = |\vec{k}| = k^2 \tag{5.5.8}$$

where $\vec{k}$ is the spatial wave vector. At first order,

$$\begin{aligned}(\eta_{\mu\nu} + \varepsilon h_{\mu\nu})(k^\mu + \varepsilon l^\mu)(k^\nu + \varepsilon l^\nu) &= 0\\ -kl^0 + 2\vec{k}\cdot\vec{l} &= 2\Phi k^2\end{aligned} \tag{5.5.9}$$

Chapter 6

Alternative Theories of Gravity

6.1 Scalar-Tensor Theories

Despite the continual success of General Relativity in describing gravitational interactions, it is always possible that the next experiment we perform could show considerable deviations and reveal effects that we cannot describe using it. In principle, there is an infinite range of modifications that could be made to GR, however, there are a few that we direct special attention to. For later comparisons and a reminder the usual Einstein-Hilbert action in $(1+3)$-dimensional Lorentzian spacetime, $\mathcal{M}$ is,

$$S_R = \frac{1}{16\pi G}\int_{\mathcal{M}} d^4x\sqrt{-g}R \tag{6.1.1}$$

where, R is the Ricci scalar, G is the Gravitational Constant and $g = \det(g_{\mu\nu})$ is the metric determinant of our spacetime. This leads to the usual Einstein Field Equations in the presence of matter,

$$G_{\mu\nu} = 8\pi G T_{\mu\nu} \tag{6.1.2}$$

A popular set of theories is known as **scalar-tensor theories** since they contain a metric tensor, $g_{\mu\nu}$, and a scalar field φ [**?**]. The action of these theories can be written as the following sum,

$$S = S_{fR} + S_\varphi + S_M, \tag{6.1.3}$$

where,

$$\begin{aligned} S_{fR} &= \int d^4x\sqrt{-g}f(\varphi)R, & (6.1.4)\\ S_\varphi &= \int d^4x\sqrt{-g}\left[-\frac{1}{2}\mathfrak{h}(\varphi)g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi - U(\varphi)\right], & (6.1.5)\end{aligned}$$

and

$$S_M = \int d^4x\sqrt{-g}\hat{\mathcal{L}}_M(g_{\mu\nu},\psi_i). \tag{6.1.6}$$

The functions $f(\varphi)$, $\mathfrak{h}(\varphi)$ and $U(\varphi)$ define the theory we are interested in and the matter Lagrangian $\hat{\mathcal{L}}_M$ is a function of the metric and the matter fields ψ_i but not of φ. By the usual variations of the inverse metric $g^{\mu\nu} \to g^{\mu\nu} + \delta g^{\mu\nu}$, we can obtain the Field Equations of this system. Variations with respect to the inverse metric can be used to obtain the variations with respect to the metric using the following conversion,

$$\begin{aligned} \delta(g^{\mu\sigma} g_{\nu\sigma}) &= \delta g^{\mu\sigma} g_{\nu\sigma} + g^{\mu\sigma}\delta g_{\nu\sigma} = 0 \\ \delta g_{\mu\nu} &= -g_{\mu\rho} g_{\nu\sigma}\delta g^{\rho\sigma} \end{aligned} \tag{6.1.7}$$

Focusing on the gravitational contributions for the moment, from the Leibnitz product rule, we have,

$$\delta S_{fR} = \delta S_1 + \delta S_2 + \delta S_3 \tag{6.1.8}$$

where,

$$\begin{aligned} \delta S_1 &= \int d^4x\sqrt{-g} f(\varphi) g^{\mu\nu}\delta R_{\mu\nu} \\ \delta S_2 &= \int d^4x\sqrt{-g} f(\varphi) R_{\mu\nu}\delta g^{\mu\nu} \\ \delta S_3 &= \int d^4x\sqrt{-g} f(\varphi) R\delta\sqrt{-g} \end{aligned}$$

We see that δS_2 is in the form that we require, however, δS_1 and δS_3 require some extra care. Starting with δS_1, we use the Riemann tensor,

$$R^{\rho}{}_{\mu\lambda\nu} = \partial_\lambda \Gamma^{\rho}{}_{\nu\mu} + \Gamma^{\rho}{}_{\lambda\sigma}\Gamma^{\sigma}{}_{\nu\mu} - \partial_\nu \Gamma^{\rho}{}_{\lambda\mu} - \Gamma^{\rho}{}_{\nu\sigma}\Gamma^{\sigma}{}_{\lambda\mu} \tag{6.1.9}$$

Following steps... The variations of the metric lead to a gnarly integral,

$$\delta S_{fR} = \int d^4x\sqrt{-g} f(\varphi)\left[\left(R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}\right)\delta g^{\mu\nu} + \nabla_\sigma\nabla^\sigma(g_{\mu\nu}\delta g^{\mu\nu}) - \nabla_\mu\nabla_\nu(\delta g^{\mu\nu})\right] \tag{6.1.10}$$

For the usual Einstein-Hilbert action, $f(\varphi)$ is a constant which means the last two terms are total derivatives, which we can integrate by parts. This will furnish the action with terms involving derivatives of $f(\varphi)$,

$$\delta S_{fR} = \int d^4x\sqrt{-g}\left[G_{\mu\nu} f(\varphi) + g_{\mu\nu}\Box f(\varphi) - \nabla_\mu\nabla_\nu f(\varphi)\right]\delta g^{\mu\nu} \tag{6.1.11}$$

with the usual $G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}$. The field equations are then given by,

$$G_{\mu\nu} = \frac{1}{f(\varphi)}\left(\frac{1}{2} T^{(M)}_{\mu\nu} + \frac{1}{2} T^{(\varphi)}_{\mu\nu} + \nabla_\mu\nabla_\nu f - g_{\mu\nu}\Box f\right) \tag{6.1.12}$$

wherein the energy-momentum tensors are defined in the standard sense in GR,

$$T^{(I)}_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta S_I}{\delta g^{\mu\nu}}, \tag{6.1.13}$$

where the index I denotes the matter or scalar parts of the actions. In particular,

$$T^{(\varphi)}_{\mu\nu} = \mathfrak{h}(\varphi)\nabla_\mu\varphi\nabla_\nu\varphi - g_{\mu\nu}\left[\frac{1}{2}\mathfrak{h}(\varphi)g^{\rho\sigma}\nabla_\rho\varphi\nabla_\sigma\varphi + U(\varphi)\right] \tag{6.1.14}$$

Looking closely at (6.1.12) and (6.1.2), we can identify $f(\varphi) = 1/16\pi G$, in the Einstein-Hilbert case. This identification gives us the usual GR and so is nothing new. However, this identifies a gravity-like theory, but with a strength that varies with the value of the scalar φ at different spacetime positions. This places a bound on how large the scalar field amplitude can be on cosmological scale and the scale of the Solar System (because it would be obviously observable... and we can't observe it). We can determine the equations of motion for φ,

$$\mathfrak{h}(\varphi)\Box\varphi + \frac{1}{2}\mathfrak{h}'(\varphi)g^{\mu\nu}(\nabla_\mu\varphi)(\nabla_\nu\varphi) - U'(\varphi) + f'(\varphi)R = 0 \tag{6.1.15}$$

where the prime $'$ denotes differentiation with respect to φ. If $\mathfrak{h}(\varphi) = 1$ the conventional scalar field equations are obtained,

$$\Box\varphi + f'(\varphi)R = U'(\varphi) \tag{6.1.16}$$

but with this curious Ricci scalar coupling. From the previous discussion, we can see that with little variations in $f(\varphi)$ this new term can be dropped to obtain the usual dynamics of φ. Another mechanism to lessen this effect is to choose a potential with a minimum and ensuring that φ cannot deviate from this minimum without a sufficient energy input. In other words, a very large mass for φ. Take a potential with $U(0) = U'(0) = 0$,

$$U(\varphi) \;\approx\; \frac{1}{2}m_\varphi^2\varphi^2 + \dots \tag{6.1.17}$$

where we identify $m_\varphi^2 = U''(0) > 0$ and $m_\varphi^2 \gg 1$. Or we could choose f and $\mathfrak{h}$ so that large changes in φ give rise to relatively small changes in the effective changes in G.

6.2 Brans-Dicke Theory

A famous example of a scalar-tensor theory is **Brans-Dicke theory**, and corresponds to the following choices,

$$f(\varphi) = \frac{\varphi}{16\pi}, \qquad \mathfrak{h}(\varphi) = \frac{\lambda}{8\pi\varphi}, \qquad U(\varphi) = 0, \tag{6.2.1}$$

where ξ is a coupling constant. The scalar-tensor action is written as,

$$S_{\rm BD} = \int d^4x\frac{\sqrt{-g}}{16\pi}\left[R\varphi - \lambda g^{\mu\nu}\frac{(\partial_\mu\varphi)(\partial_\nu\varphi)}{\varphi}\right]. \tag{6.2.2}$$

Brans-Dicke theory chooses a massless scalar. But in the limit $\lambda \to \infty$ the field becomes non-dynamical and ordinary GR is recovered. Current tests in the solar system imply $\lambda > 500$, of if there is such a scalar field it must couple only weakly to R.

6.2.1 Using Conformal Transformations

One way to deal with scalar-tensor theories is to perform a conformal transformation to make the connection to gravity more apparent. We choose the conformal metric, $\tilde{g}_{\mu\nu}$,

$$\tilde{g}_{\mu\nu} = 16\pi\tilde{G}f(\varphi)g_{\mu\nu} \tag{6.2.1}$$

where $\tilde{G}$ is the Gravitational constant in the conformal frame. The connection coefficients in this new frame ($\tilde{\Gamma}^{\mu}{}_{\rho\sigma}$) are,

$$\tilde{\Gamma}^{\rho}{}_{\mu\nu} = \frac{1}{2}\tilde{g}^{\rho\sigma}\left(\tilde{g}_{\mu\sigma,\nu} + \tilde{g}_{\nu\sigma,\mu} - \tilde{g}_{\mu\nu,\sigma}\right) \tag{6.2.2}$$

Substituting the conformal metric (6.2.1) in terms of the usual metric give,

$$\begin{aligned}
\tilde{\Gamma}^{\rho}{}_{\mu\nu} &= \Gamma^{\rho}{}_{\mu\nu} + \frac{1}{2}f(\varphi)^{-1}\big((\partial_\nu f)\delta^\rho_\mu + (\partial_\mu f)\delta^\rho_\nu - g^{\rho\sigma}g_{\mu\nu}(\partial_\sigma f)\big) && (6.2.3)\\
&= \Gamma^{\rho}{}_{\mu\nu} + C^{\rho}{}_{\mu\nu}, && (6.2.4)
\end{aligned}$$

where $C^{\rho}{}_{\mu\nu}$ is the difference between connections and so is a tensor (see Appendix C). Using the chain rule, $\partial_\mu f(\varphi) = (\partial f/\partial\varphi)\partial_\mu\varphi$, we can write,

$$\tilde{\Gamma}^{\rho}{}_{\mu\nu} = \Gamma^{\rho}{}_{\mu\nu} + \frac{1}{2}\frac{\partial \ln[f(\varphi)]}{\partial\varphi}\left[(\partial_\nu\varphi)\delta^\rho_\mu + (\partial_\mu\varphi)\delta^\rho_\nu - g^{\rho\sigma}g_{\mu\nu}(\partial_\sigma\varphi)\right] \tag{6.2.5}$$

This connection defines a new covariant derivative on a vector field, $V = V^\mu(x)\partial_\mu$,

$$\begin{aligned}
\tilde{\nabla}_\mu V^\rho &= \partial_\mu V^\rho + \tilde{\Gamma}^{\rho}{}_{\mu\nu}V^\nu,\\
&= \partial_\mu V^\rho + \Gamma^{\rho}{}_{\mu\nu}V^\nu + \frac{1}{2}\frac{\partial \ln[f(\varphi)]}{\partial\varphi}\left[(V^\nu\partial_\nu\varphi)\delta^\rho_\mu + (\partial_\mu\varphi)V^\rho - g^{\rho\sigma}(\partial_\sigma\varphi)V_\mu\right]
\end{aligned}$$

and on a covector field, $\omega = \omega_\mu(x)dx^\mu$,

$$\begin{aligned}
\tilde{\nabla}_\mu \omega_\nu &= \partial_\mu\omega_\nu - \tilde{\Gamma}^{\rho}{}_{\mu\nu}\omega_\rho,\\
&= \partial_\mu\omega_\nu - \Gamma^{\rho}{}_{\mu\nu}\omega_\rho - \frac{1}{2}\frac{\partial \ln[f(\varphi)]}{\partial\varphi}\left[(\partial_\nu\varphi)\omega_\mu + (\partial_\mu\varphi)\omega_\nu - g_{\mu\nu}(\partial_\sigma\varphi)\omega^\sigma\right]
\end{aligned}$$

which generalises to higher-order tensors in the usual manner Continuing with the general calculations of the Riemann and Ricci tensors, we obtain the following action in terms of the conformal Ricci Scalar,

$$S_{fR} = \int d^4x\ \frac{\sqrt{-g}}{16\pi\tilde{G}}\left[\tilde{R} + \frac{3}{2}\tilde{g}^{\rho\sigma}f^{-2}\left(\frac{df}{d\varphi}\right)^2(\tilde{\nabla}_\rho\varphi)(\tilde{\nabla}_\sigma\varphi)\right] \tag{6.2.6}$$

where we have integrated by parts and discarded surface terms. In the conformal frame, therefore, the curvature scalar appears by itself, which looks a lot like GR. For this reason, this frame is called the **Einstein frame**, since the Einstein equations for the conformal metric $\tilde{g}_{\mu\nu}$ take on their conventional form. The original frame with metric $g_{\mu\nu}$ is called the **Jordan frame**.

If we make the choice,

$$f(\varphi) = e^{\varphi/\sqrt{3}}, \quad \mathfrak{h}(\varphi) = U(\varphi) = 0 \tag{6.2.7}$$

which is a specific choice of $f(\varphi)$ and turns off the S_φ contribution from the pure scalar. With these choices, in the Einstein frame, we see a conventional kinetic term for the scalar field,

$$S = \int d^4x \sqrt{-\tilde{g}} (16\pi\tilde{G})^{-1} \left[\tilde{R} - \frac{1}{2} \tilde{\nabla}_\rho \varphi \tilde{\nabla}^\rho \varphi \right] + S_M. \tag{6.2.8}$$

Even without an explicit kinetic term in the Jordan frame, we see a kinetic term generated in the Einstein frame.

6.3 Linearised Scalar Tensor Theories

But we need to prepare (6.1.12) expressing the perturbative expansion of the metric in our analysis. For simplicity and generality, we will write $T^{(\varphi)}_{\mu\nu} + T^{(M)}_{\mu\nu} = T_{\mu\nu}$, giving,

$$G_{\mu\nu} = \frac{1}{f(\varphi)} \left(\frac{1}{2} T_{\mu\nu} + \nabla_\mu \nabla_\nu f - g_{\mu\nu} \Box f \right) \tag{6.3.1}$$

Note the following expressions:

- $\nabla_\mu \nabla_\nu f = \partial_\mu \partial_\nu f - \Gamma^\alpha_{\mu\nu} \partial_\alpha f$, where ${\Gamma^\rho}_{\mu\nu}$ refers to (**??**),
- $g_{\mu\nu} \Box f = \eta_{\mu\nu} \Box f + \eta_{\mu\nu} h^{\rho\sigma} \partial_\rho \partial_\sigma f + h_{\mu\nu} \Box f$, where on the RHS, $\Box = \eta^{\alpha\beta} \partial_\alpha \partial_\beta$,

Let's put it together,

$$\begin{aligned} G_{\mu\nu} &= \partial^\sigma \partial_{(\mu} h_{\nu)\sigma} - \frac{1}{2} \left(\Box h_{\mu\nu} + \partial_\mu \partial_\nu h + \eta_{\mu\nu} \partial_\rho \partial_\sigma h^{\rho\sigma} + \eta_{\mu\nu} \Box h \right) \\ &= \frac{1}{f(\varphi)} \left(\frac{1}{2} T_{\mu\nu} + \partial_\mu \partial_\nu f - \Gamma^\alpha_{\mu\nu} \partial_\alpha f - \eta_{\mu\nu} \Box f - \eta_{\mu\nu} h^{\rho\sigma} \partial_\rho \partial_\sigma f - h_{\mu\nu} \Box f \right), \end{aligned} \tag{6.3.2}$$

which is ghastly to be honest. That Christoffel symbol term doesn't look so good. Let's substitute in the definition (**??**),

$$\begin{aligned} G_{\mu\nu} &= \partial^\sigma \partial_{(\mu} h_{\nu)\sigma} - \frac{1}{2} \left(\Box h_{\mu\nu} + \partial_\mu \partial_\nu h + \eta_{\mu\nu} \partial_\rho \partial_\sigma h^{\rho\sigma} + \eta_{\mu\nu} \Box h \right) \\ &= \frac{1}{f(\varphi)} \left(\frac{1}{2} T_{\mu\nu} + \partial_\mu \partial_\nu f - \partial_{(\mu} h_{\nu)\beta} \partial^\beta f + \frac{1}{2} h_{\mu\nu,\beta} \partial^\beta f - \eta_{\mu\nu} \Box f - \eta_{\mu\nu} h^{\rho\sigma} \partial_\rho \partial_\sigma f - h_{\mu\nu} \Box f \right). \end{aligned} \tag{6.3.3}$$

We have reduced all derived quantities in terms of the metric perturbations $h_{\mu\nu}$. Let's now decompose $G_{\mu\nu}$ in terms of the degrees of freedom expressed in (5.4.1),

$$
\begin{aligned}
G_{00} &= \frac{1}{f(\varphi)}\left[\frac{1}{2}T^{I}_{00} + (1+2\Phi+2\Psi)\nabla^2 f - 4\Phi\partial_0^2 f - 2(\partial_0\Phi)(\partial_0 f) + \partial_0 w^i\partial_i f - w^i\partial_i\partial_0 f + 2s^{kl}\partial_k\partial_l f\right]\\
G_{0i} &= \frac{1}{f(\varphi)}\left[\frac{1}{2}T^{I}_{0i} + \partial_0\partial_i f + \left(\frac{1}{2}\partial_0 w^i - \partial_i\Phi\right)\partial_0 f - \left(\partial_0 s_{ij} - \frac{1}{2}\partial_i w^j\right)\partial_j f + (\partial_0\Psi)\partial_i f - w^i\Box f\right]\\
G_{ij} &= \frac{1}{f(\varphi)}\left[\frac{1}{2}T^{I}_{ij} + \partial_i\partial_j f + \partial_{(i}w_{j)} - 2\left(\partial_{(i}s_{j)k} - \delta_{k(j}\partial_{i)}\Psi\right)\partial_k f\right.\\
&\qquad \left. -\delta_{ij}\Box f - \delta_{ij}h^{\rho\sigma}\partial_\rho\partial_\sigma f - 2\left(s_{ij} + \Psi\delta_{ij}\right)\Box f\right],
\end{aligned}
\tag{6.3.4}
$$

where the LHS of the equation is equal to the corresponding terms in (5.4.5). Note, that when $f(\varphi) = (16\pi G)^{-1}$ we obtain the same perturbed equations for ordinary GR as before. Now, this is getting a bit hairy, but we are ready to fix a gauge.

We'll bring the equations back down here for a sec,

$$
\begin{aligned}
G_{00} &= 2\nabla^2\Psi + \partial_k\partial_l s^{kl}\\
G_{0j} &= -\frac{1}{2}\nabla^2 w_j + \frac{1}{2}\partial_j\partial_k w^k + 2\partial_0\partial_j\Psi + \partial_0\partial_k s^k_j\\
G_{ij} &= (\delta_{ij}\nabla^2 - \partial_i\partial_j)(\Phi - \Psi) + \delta_{ij}\partial_0\partial_k w^k - \partial_0\partial_{(i}w_{j)}\\
&\qquad + 2\delta_{ij}\partial_0^2\Psi - \Box s_{ij} + 2\partial_k\partial_{(i}s_{j)}^{\ k} - \delta_{ij}\partial_k\partial_l s^{kl}.
\end{aligned}
\tag{6.3.5}
$$

We choose the **synchronous frame** ($\Phi = w^i = 0$), as was the choice in ordinary GR,

$$
\begin{aligned}
G_{00} &= \frac{1}{f(\varphi)}\left[\frac{1}{2}T_{00} + (1+2\Psi)\nabla^2 f + 2s^{kl}\partial_k\partial_l f\right] = 2\nabla^2\Psi + \partial_k\partial_l s^{kl}\\
G_{0i} &= \frac{1}{f(\varphi)}\left[\frac{1}{2}T_{0i} + \partial_0\partial_i f - \partial_0 s_{ij}\partial_j f\right] = 2\partial_0\partial_j\Psi + \partial_0\partial_k s^k_j\\
G_{ij} &= \frac{1}{f(\varphi)}\left[\frac{1}{2}T_{ij} + \partial_i\partial_j f - 2\left(\partial_{(i}s_{j)k} - \delta_{k(j}\partial_{i)}\Psi\right)\partial_k f - \delta_{ij}\Box f\right.\\
&\qquad \left. - \delta_{ij}h^{\rho\sigma}\partial_\rho\partial_\sigma f - 2\left(s_{ij} + \Psi\delta_{ij}\right)\Box f\right]\\
&= (\delta_{ij}\nabla^2 - \partial_i\partial_j)(\Phi - \Psi) + 2\delta_{ij}\partial_0^2\Psi - \Box s_{ij} + 2\partial_k\partial_{(i}s_{j)}^{\ k} - \delta_{ij}\partial_k\partial_l s^{kl}.
\end{aligned}
\tag{6.3.6}
$$

Metric in this gauge will be,

$$
\mathrm{d}s^2 = -dt^2 + (\delta_{ij} + h_{ij})\,\mathrm{d}x^i\,\mathrm{d}x^j
\tag{6.3.7}
$$

Appendices

Appendix A

Shortcuts to calculate objects in General Relativity

This chapter will address calculations in General Relativity. GR contains many important expressions that are required to be calculated but the sheer number of indices can often be a barrier in detangling the values of these expressions. We work through some examples and explain the importance of developing a technique in calculations.

We are almost always given a metric ansatz $g_{\mu\nu}(x)$. From this initial metric, we can work calculate basically every piece of information about the dynamics on our manifold. The following cascade will help to understand,

$$ \tag{A.0.1}$$

Following this cascade, will help us to determine the curvature side of the Einstein Field Equations, $G_{\mu\nu} = 8\pi G T_{\mu\nu}$

A.1 Christoffel Symbols

The most basic quantity we can calculate, given a metric, are the Christoffel symbols. At the face of it, $\Gamma^{\alpha}{}_{\mu\nu}$ has three indices each of which can range from $\{0, 1, 2, 3\}$. So, in a d-dimensional spacetime (1 timelike dimension and $(d-1)$ spacelike dimensions) that's at most, d^3 ($= 64$ in $d = 4$-spacetime) non-zero components to calculate. However, we note that that Christoffel symbols are symmetric under the exchange of the two covector indices, $\Gamma^{\rho}{}_{\mu\nu} = \Gamma^{\rho}{}_{(\mu\nu)}$, which means we can eliminate some of these possibilities. Taking this into consideration, we now have at most $2d(d+1)$ non-zero terms to calculate, which is a little bit better ($d^3 \geq 2d(d+1)$).

A.1.1 Calculating the Christoffel Symbols

We can imagine the Christoffel symbols as a d-paged booklet, with each page labelled by $\rho = \{0, 1, 2, 3\}$. Each page contains a $(d \times d)$-matrix.

This is more instructive through examples.

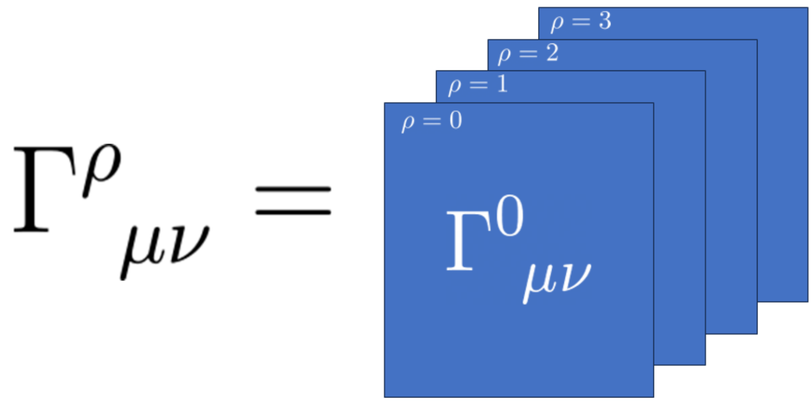


Figure A.1.1: Each page of this book is a matrix labelled by the value of ρ.

Example: Metric on a S^2

We will initially work on calculating the Christoffel symbols on a unit two-sphere, S^2 $(d = 2)$. The metric is given by,

$$\mathrm{d}s^2 = \mathrm{d}\theta^2 + \sin^2(\theta)\,\mathrm{d}\phi^2, \tag{A.1.1}$$

and so we can already note that, $g_{\theta\theta} = 1$ and $g_{\phi\phi} = \sin^2(\theta)$ (note: θ and ϕ are not Lorentz indices, but are labels for the corresponding coordinates in the metric). There are only two *pages* in this set of Christoffel symbols, $\Gamma^{\theta}{}_{ij}$ and $\Gamma^{\phi}{}_{ij}$

$$\begin{aligned}
\Gamma^{\theta}{}_{ij} &= \frac{1}{2} g^{\theta k}\left(g_{ik,j} + g_{jk,i} - g_{ij,k}\right) \\
\Gamma^{\phi}{}_{ij} &= \frac{1}{2} g^{\phi k}\left(g_{ik,j} + g_{jk,i} - g_{ij,k}\right)
\end{aligned} \tag{A.1.2}$$

An additional note on the structure of these expressions, the metric is purely diagonal, therefore there is an additional constraint placed on the index k, in that it has to match the other index on the $g^{\mu\nu}$ factor[1]. Also we can drop terms containing derivatives of ϕ because it is an *ignorable coordinate* (g is not an explicit function of ϕ),

$$\begin{aligned}
\Gamma^{\theta}{}_{ij} &= \frac{1}{2}\left(g_{i\theta,j} + g_{j\theta,i} - g_{ij,\theta}\right) \\
\Gamma^{\phi}{}_{ij} &= \frac{1}{2\sin^2\theta}\left(g_{i\phi,j} + g_{j\phi,i}\right).
\end{aligned} \tag{A.1.3}$$

This is still completely general for all symbols in this example. Now we can choose all possible combinations of the coordinates, which is easy in this example because there's only two coordinates.

[1] This will almost always be the case as we will deal with only diagonal metrics.

Terms proportional to any derivatives of $g_{\theta\theta}$ will vanish,

$$\begin{aligned}
\Gamma^{\theta}{}_{\theta\theta} &= \Gamma^{\theta}{}_{\theta\phi} = 0, \quad \Gamma^{\theta}{}_{\phi\phi} = -\frac{1}{2}g_{\phi\phi,\theta} = -\frac{1}{2}\frac{\partial}{\partial\theta}(\sin^2\theta) = -\sin\theta\cos\theta \\
\Gamma^{\phi}{}_{\theta\theta} &= \Gamma^{\phi}{}_{\phi\phi} = 0, \quad \Gamma^{\phi}{}_{\theta\phi} = \frac{1}{2\sin^2\theta}\frac{\partial}{\partial\theta}g_{\phi\phi} = \frac{1}{2\sin^2\theta}(2\sin\theta\cos\theta) = \cot\theta.
\end{aligned}$$

In matrix form, we can write the following,

$$\Gamma^{\theta}{}_{ij} = \begin{pmatrix} 0 & 0 \\ 0 & -\sin\theta\cos\theta \end{pmatrix}_{ij} \qquad \Gamma^{\phi}{}_{ij} = \begin{pmatrix} 0 & \cot\theta \\ \cot\theta & 0 \end{pmatrix}_{ij}. \tag{A.1.4}$$

Although this was a simple example, using this method becomes more efficient for keeping track of indices in higher-dimensional examples and facilitates a neater calculation.

Metric on a surface of constant proper time embedded in $\mathbb{M}^{1,3}$

Consider the Minkowski spacetime in inertial coordinates, x^{μ}. Consider the three-dimensional surface of points that can be reached in proper time R, where $R \geq 0$. The surface is given by the equation,

$$R^2 = t^2 - (x^2 + y^2 + z^2) \tag{A.1.5}$$

This surface can be parametrised by the following coordinate transformation,

$$t = R\cosh\xi, \quad z = R\sinh\xi\cos\theta, \quad y = R\sinh\xi\sin\theta\sin\phi, \quad x = R\sinh\xi\sin\theta\cos\phi, \tag{A.1.6}$$

with $\xi \in (0, +\infty]$, $\theta \in [0, \pi)$ and $\phi \in (0, 2\pi]$. By change of coordinates, the metric can be found to be. Under a general coordinate transformation, we know that the metric components transform as,

$$g_{\mu'\nu'}(x') = \frac{\partial x^{\mu}}{\partial x'^{\mu'}}\frac{\partial x^{\nu}}{\partial x'^{\nu'}}g_{\mu\nu}(x), \tag{A.1.7}$$

where $x^{\mu} = (t, x, y, z)$ and $x'^{\mu'} = (\xi, \theta, \phi)$ (we have one less coordinate in the new frame because the dependence (A.1.5) removes one of the degrees of freedom). If we recall that the metric is a diagonal matrix, the only surviving terms on the right-hand side have $\mu = \nu$. So we are then free to choose the μ' and ν' indices. Look at the $g_{\xi\xi}$ term and notice that each component of x^{μ} depends on ξ,

$$\begin{aligned}
g_{\xi\xi} &= \frac{\partial x^{\mu}}{\partial\xi}\frac{\partial x^{\nu}}{\partial\xi}g_{\mu\nu}(x) \\
&= -\left(\frac{\partial t}{\partial\xi}\right)^2 + \left(\frac{\partial x}{\partial\xi}\right)^2 + \left(\frac{\partial y}{\partial\xi}\right)^2 + \left(\frac{\partial z}{\partial\xi}\right)^2 \\
&= R^2\left[-(\sinh\xi)^2 + (\cosh\xi\cos\theta)^2 + (\cosh\xi\sin\theta\sin\phi)^2 + (\cosh\xi\sin\theta\cos\phi)^2\right] \\
&= R^2.
\end{aligned} \tag{A.1.8}$$

Following along the same lines, we find the metric,

$$\mathrm{d}s^2 = R^2\left[\mathrm{d}\xi^2 + \sinh^2\xi(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right] \tag{A.1.9}$$

This is the line element on a three-dimensional hyperbolic space, $\mathbb{H}^3$. Now, the Christoffel symbols of this metric (note that ϕ is ignorable, so any derivatives with respect to ϕ will vanish),

$$\begin{aligned}
{\Gamma^{\xi}}_{ij} &= \frac{1}{2R^2}\left(g_{i\xi,j} + g_{j\xi,i} - g_{ij,\xi}\right) \\[2ex]
{\Gamma^{\xi}}_{\xi\xi} &= {\Gamma^{\xi}}_{\xi\theta} = {\Gamma^{\xi}}_{\xi\phi} = {\Gamma^{\xi}}_{\theta\phi} = 0, \\
{\Gamma^{\xi}}_{\theta\theta} &= -\sinh\xi\cosh\xi, \\
{\Gamma^{\xi}}_{\phi\phi} &= -\sinh\xi\cosh\xi\sin^2\theta
\end{aligned} \tag{A.1.10}$$

$$\begin{aligned}
{\Gamma^{\phi}}_{ij} &= \frac{1}{2R^2\sinh^2\xi}\left(g_{i\theta,j} + g_{j\theta,i} - g_{ij,\theta}\right) \\[2ex]
{\Gamma^{\theta}}_{\xi\xi} &= {\Gamma^{\theta}}_{\xi\phi} = {\Gamma^{\theta}}_{\theta\theta} = {\Gamma^{\theta}}_{\theta\phi} = 0 \\
{\Gamma^{\theta}}_{\xi\theta} &= -\sin^2\theta\coth\xi \\
{\Gamma^{\theta}}_{\phi\phi} &= -\sin\theta\cos\theta.
\end{aligned} \tag{A.1.11}$$

$$\begin{aligned}
{\Gamma^{\phi}}_{ij} &= \frac{1}{2R^2\sinh^2\xi\sin^2\theta}\left(g_{i\phi,j} + g_{j\phi,i}\right) \\[2ex]
{\Gamma^{\phi}}_{\xi\xi} &= {\Gamma^{\phi}}_{\xi\theta} = {\Gamma^{\phi}}_{\theta\theta} = {\Gamma^{\phi}}_{\phi\phi} = 0 \\
{\Gamma^{\phi}}_{\theta\phi} &= \cot\theta \\
{\Gamma^{\phi}}_{\xi\phi} &= \coth\xi
\end{aligned} \tag{A.1.12}$$

We can write these in the following matrix form,

$$\begin{aligned}
{\Gamma^{\xi}}_{ij} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\sinh\xi\cosh\xi & 0 \\ 0 & 0 & -\sinh\xi\cosh\xi\sin^2\theta \end{pmatrix}_{ij} \\
{\Gamma^{\theta}}_{ij} &= \begin{pmatrix} 0 & -\sin^2\theta\coth\xi & 0 \\ -\sin^2\theta\coth\xi & 0 & 0 \\ 0 & 0 & -\sin\theta\cos\theta \end{pmatrix}_{ij} \\
{\Gamma^{\phi}}_{ij} &= \begin{pmatrix} 0 & 0 & \cot\theta \\ 0 & 0 & \coth\xi \\ \cot\theta & \coth\xi & 0 \end{pmatrix}_{ij}.
\end{aligned} \tag{A.1.13}$$

Christoffel symbols of the Schwarzschild Metric

Consider the metric of static, two-sphere of mass M, with Schwarzschild radius, $r_S = 2GM$,

$$\mathrm{d}s^2 = -(1 - r_S/r)\,\mathrm{d}t^2 + (1 - r_S/r)^{-1}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2 \tag{A.1.14}$$

where $\mathrm{d}\Omega^2$ is the metric on a two-sphere as in (A.1.1). The Christoffel symbols for a static solution are,

$$\Gamma^t{}_{ij} = -\frac{r}{2(r - r_S)}\left(g_{it,j} + g_{jt,i}\right) \tag{A.1.15}$$

But the only non-vanishing component of this matrix is,

$$\Gamma^t{}_{tr} = \frac{r_S}{2r(r - r_S)} \tag{A.1.16}$$

Now we move on to the other components

$$\begin{aligned}\Gamma^r{}_{ij} &= \frac{r - r_S}{2r}\left(g_{ir,j} + g_{jr,i} - g_{ij,r}\right)\\ &= \begin{pmatrix} -\frac{r_S(r-r_S)}{2r^3} & 0 & 0 & 0\\ 0 & -\frac{r_S}{2r(r-r_S)} & 0 & 0\\ 0 & 0 & (r_S - r) & 0\\ 0 & 0 & 0 & (r_S - r)\sin^2\theta \end{pmatrix}\end{aligned} \tag{A.1.17}$$

Not the θ matrix,

$$\Gamma^\theta{}_{ij} = \frac{1}{2r^2}\left(g_{i\theta,j} + g_{j\theta,i} - g_{ij,\theta}\right) = \begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & r^{-1} & 0\\ 0 & r^{-1} & 0 & 0\\ 0 & 0 & 0 & -\sin\theta\cos\theta \end{pmatrix} \tag{A.1.18}$$

Which leaves ϕ,

$$\Gamma^\phi{}_{ij} = \frac{1}{2r^2\sin^2\theta}\left(g_{i\phi,j} + g_{j\phi,i} - g_{ij,\phi}\right) = \begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & r^{-1}\\ 0 & 0 & 0 & \cot\theta\\ 0 & r^{-1} & \cot\theta & 0 \end{pmatrix} \tag{A.1.19}$$

Appendix B

Conformal Transformations

A conformal transformation is a transformation between metric spaces which we choose an arbitrary symmetric conformal metric, $\tilde{g}_{\mu\nu}$,

$$\tilde{g}_{\mu\nu} = 16\pi\tilde{G}f(\varphi)g_{\mu\nu} \tag{B.0.1}$$

where $\tilde{G}$ is the Gravitational constant in the conformal frame. The connection coefficients in this new frame (${\tilde{\Gamma}^{\mu}}_{\rho\sigma}$) are,

$${\tilde{\Gamma}^{\rho}}_{\mu\nu} = \frac{1}{2}\tilde{g}^{\rho\sigma}\left(\tilde{g}_{\mu\sigma,\nu} + \tilde{g}_{\nu\sigma,\mu} - \tilde{g}_{\mu\nu,\sigma}\right) \tag{B.0.2}$$

Substituting the conformal metric (B.0.1) in terms of the usual metric give,

$$\begin{aligned}
{\tilde{\Gamma}^{\rho}}_{\mu\nu} &= \frac{1}{2}f(\varphi)^{-1}g^{\rho\sigma}\left(\partial_\nu(fg_{\mu\sigma}) + \partial_\mu(fg_{\nu\sigma}) - \partial_\sigma(fg_{\mu\nu})\right) && \text{(B.0.3)}\\
&= {\Gamma^{\rho}}_{\mu\nu} + \frac{1}{2}f(\varphi)^{-1}g^{\rho\sigma}\left((\partial_\nu f)g_{\mu\sigma} + (\partial_\mu f)g_{\nu\sigma} - (\partial_\sigma f)g_{\mu\nu}\right) && \text{(B.0.4)}\\
&= {\Gamma^{\rho}}_{\mu\nu} + \frac{1}{2}f(\varphi)^{-1}\left((\partial_\nu f)\delta^\rho_\mu + (\partial_\mu f)\delta^\rho_\nu - g^{\rho\sigma}g_{\mu\nu}(\partial_\sigma f)\right) && \text{(B.0.5)}
\end{aligned}$$

Using the chain rule, $\partial_\mu f(\varphi) = (\partial f/\partial\varphi)\partial_\mu\varphi$, we can write,

$${\tilde{\Gamma}^{\rho}}_{\mu\nu} = {\Gamma^{\rho}}_{\mu\nu} + \frac{1}{2}\frac{\partial\ln[f(\varphi)]}{\partial\varphi}\left[(\partial_\nu\varphi)\delta^\rho_\mu + (\partial_\mu\varphi)\delta^\rho_\nu - g^{\rho\sigma}g_{\mu\nu}(\partial_\sigma\varphi)\right] \tag{B.0.6}$$

This connection defines a new covariant derivative on a vector field, $A = A^\mu(x)\partial_\mu$,

$$\begin{aligned}
\tilde{\nabla}_\mu A^\rho &= \partial_\mu A^\rho + {\tilde{\Gamma}^{\rho}}_{\mu\nu}A^\nu,\\
&= \partial_\mu A^\rho + {\Gamma^{\rho}}_{\mu\nu}A^\nu + \frac{1}{2}\frac{\partial\ln[f(\varphi)]}{\partial\varphi}\left[(A^\nu\partial_\nu\varphi)\delta^\rho_\mu + (\partial_\mu\varphi)A^\rho - g^{\rho\sigma}A_\mu(\partial_\sigma\varphi)\right]
\end{aligned}$$

and on a covector field, $\omega = \omega_\mu(x)dx^\mu$,

$$\begin{aligned}
\tilde{\nabla}_\mu\omega_\nu &= \partial_\mu\omega_\nu - {\tilde{\Gamma}^{\rho}}_{\mu\nu}\omega_\rho,\\
&= \partial_\mu\omega_\nu - {\Gamma^{\rho}}_{\mu\nu}\omega_\rho - \frac{1}{2}\frac{\partial\ln[f(\varphi)]}{\partial\varphi}\left[(\partial_\nu\varphi)\omega_\mu + (\partial_\mu\varphi)\omega_\nu - g_{\mu\nu}(\partial_\sigma\varphi)\omega^\sigma\right]
\end{aligned}$$

which generalises to higher-order tensors.

Appendix C

Fifth forces in alternative theories of gravity

C.1 Dynamics in conformal metrics

Consider a metric in the Jordan Frame,

$$g_{\mu\nu} = A(\varphi)^2 \tilde{g}_{\mu\nu} \tag{C.1.1}$$

in a scalar-tensor theory $(\tilde{g}_{\mu\nu}, \varphi)$. $A(\varphi)$ is usually give by,

$$A(\varphi) = 1 + \frac{\varphi^2}{2M^2} + \mathcal{O}\left(\frac{\varphi^4}{M^4}\right), \tag{C.1.2}$$

which is an effective field theory [**?**]. The geodesics of this theory are those of the Jordan frame metric. A massive particle will move on some affinely-parametrised geodesic,

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu{}_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} = 0, \tag{C.1.3}$$

where $\Gamma^\mu{}_{\rho\sigma}$ defines a Levi-Civita connection of $g_{\mu\nu}$, where $\frac{dx^\mu}{d\tau}$ are tangent vectors of the geodesics in the Jordan Frame. It is important to note that the analogous equation for the Einstein frame metric, $\tilde{g}_{\mu\nu}$, deviates from the geodesics

$$a^\mu_E = \frac{d^2 x^\mu}{d\tau^2} + \tilde{\Gamma}^\mu{}_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} \neq 0. \tag{C.1.4}$$

where a^μ_E is an acceleration of a massive particle in the Einstein frame. This is because the tangent vectors generate the geodesics in the Jordan frame and so are not necessarily geodesics in the Einstein frame (in the Einstein Frame, this corresponds to geodesic deviation). Our goal is to calculate the Γ-coefficients of Jordan Frame in terms of $\tilde{g}_{\mu\nu}$ given the above considerations.

C.2 Conformal transformations on Connections

We begin by substituting the Jordan frame metric into the standard definition of the Christoffel Symbols.

$$\Gamma^{\mu}{}_{\rho\sigma} = \frac{1}{2} A^{-2} \tilde{g}^{\mu\nu} \left[(A^2 \tilde{g}_{\rho\nu}),_{\sigma} + (A^2 \tilde{g}_{\nu\sigma}),_{\rho} - (A^2 \tilde{g}_{\rho\sigma}),_{\nu} \right] \tag{C.2.1}$$

Expanding this out and noticing that $\tilde{\Gamma} \sim \tilde{g}^{-1}\partial\tilde{g}$, we can write,

$$\Gamma^{\mu}{}_{\rho\sigma} = \tilde{\Gamma}^{\mu}{}_{\rho\sigma} + A^{-1} \left[\delta^{\mu}_{\rho}(\partial_{\sigma} A) + \delta^{\mu}_{\sigma}(\partial_{\rho} A) - \tilde{g}_{\rho\sigma}\tilde{g}^{\mu\nu}(\partial_{\nu} A) \right]. \tag{C.2.2}$$

Since the second term is the difference between two connections, we can identify it as $(2,1)$ tensor, $C^{\mu}{}_{\rho\sigma}$. As $A(\varphi)$ is only explicitly a function of φ, the chain rule can be used to convert this expression to the following expression for $C^{\mu}{}_{\rho\sigma}$,

$$\begin{aligned} C^{\mu}{}_{\rho\sigma} &= A^{-1}\frac{\partial A}{\partial\varphi} \left[\delta^{\mu}_{\rho}(\partial_{\sigma}\varphi) + \delta^{\mu}_{\sigma}(\partial_{\rho}\varphi) - \tilde{g}_{\rho\sigma}\tilde{g}^{\mu\nu}(\partial_{\nu}\varphi) \right] \\ &= \frac{\partial(\ln A(\varphi))}{\partial\varphi} \left[\delta^{\mu}_{\rho}(\partial_{\sigma}\varphi) + \delta^{\mu}_{\sigma}(\partial_{\rho}\varphi) - \tilde{g}_{\rho\sigma}\tilde{g}^{\mu\nu}(\partial_{\nu}\varphi) \right] \end{aligned} \tag{C.2.3}$$

This is a tensor because we can be expressed completely covariantly,

$$C^{\mu}{}_{\rho\sigma} = A^{-1}\frac{\partial A}{\partial\varphi} \left[\delta^{\mu}_{\rho}(\nabla_{\sigma}\varphi) + \delta^{\mu}_{\sigma}(\nabla_{\rho}\varphi) - \tilde{g}_{\rho\sigma}\tilde{g}^{\mu\nu}(\nabla_{\nu}\varphi) \right], \tag{C.2.4}$$

because $\nabla_{\mu}\varphi \equiv \partial_{\mu}\varphi$.

This can be proven in terms of general coordinate transformations:

The difference between two Levi-Civita connections ${C^\mu}_{\rho\sigma} = {\Gamma^\mu}_{\rho\sigma} - {\tilde{\Gamma}^\mu}_{\rho\sigma}$ is a tensor under general coordinate transformations.

Proof: Consider a matrix coordinate transformation x and x',

$$M : x^\mu \to {x'}^{\mu'}.$$

A general Levi-Civita connection will transform as follows,

$${\Gamma^\mu}_{\rho\sigma} \to {\Gamma^{\mu'}}_{\rho'\sigma'} = \frac{1}{2} g^{\mu'\nu'} \left(g_{\rho'\nu',\sigma'} + g_{\nu'\sigma',\rho'} - g_{\rho'\sigma',\nu'} \right) \tag{C.2.5}$$

The component terms of the above connection only differ in index placement, a particular term will transform as follows:

$$\begin{aligned} \partial_\mu g_{\rho\sigma} \to \partial_{\mu'} g_{\rho'\sigma'} &= \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial}{\partial x^\mu} \left(\frac{\partial x^\rho}{\partial x^{\rho'}} \frac{\partial x^\nu}{\partial x^{\nu'}} g_{\rho\nu} \right) \\ &= \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial x^\nu}{\partial x^{\nu'}} \frac{\partial}{\partial x^\mu} \left(\frac{\partial x^\rho}{\partial x^{\rho'}} \right) g_{\rho\nu} + \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial x^\rho}{\partial x^{\rho'}} \frac{\partial}{\partial x^\mu} \left(\frac{\partial x^\nu}{\partial x^{\nu'}} \right) g_{\rho\nu} \\ &\quad + \frac{\partial x^{\mu'}}{\partial x^\mu} \frac{\partial x^\rho}{\partial x^{\rho'}} \frac{\partial x^\nu}{\partial x^{\nu'}} \partial_\mu g_{\rho\nu} \end{aligned}$$

The geodesic equation is now,

$$\frac{d^2 x^\mu}{d\tau^2} + {\tilde{\Gamma}^\mu}_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} + {C^\mu}_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} = 0 \tag{C.2.6}$$

For which we identify the first two terms as a^μ_E,

$$a^\mu_E = -{C^\mu}_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} = F^\mu \tag{C.2.7}$$

where the equality is the covariant force vector for a unit-mass particle. Writing $u^\mu = \frac{dx^\mu}{d\tau}$ we get the following force on this test particle is given by,

$$\begin{aligned} F^\mu &= A^{-1} \frac{\partial A}{\partial \varphi} \left[\delta^\mu_\rho (\nabla_\sigma \varphi) + \delta^\mu_\sigma (\nabla_\rho \varphi) - \tilde{g}_{\rho\sigma} \tilde{g}^{\mu\nu} (\nabla_\nu \varphi) \right] u^\rho u^\sigma && \text{(C.2.8)} \\ &= A^{-1} \frac{\partial A}{\partial \varphi} \left[2u^\mu (u^\alpha \nabla_\alpha \varphi) - \tilde{g}^{\mu\nu} (\nabla_\nu \varphi) u^\alpha u_\alpha \right] && \text{(C.2.9)} \end{aligned}$$

C.3 Considering forces in Local Inertial Frames

Consider the expression for the force on the particle in a local inertial frame. This amounts to appropriate substitutions to use this relations, $\lim_{\text{LIF}} \tilde{g}_{\mu\nu} \to \eta_{\mu\nu}$ and $\nabla_\lambda \to \partial_\lambda$,

$$F^\mu = -\frac{\partial(\ln A(\varphi))}{\partial \varphi} \left[2u^\mu (u^\alpha \partial_\alpha \varphi) - \eta^{\mu\nu} (\partial_\nu \varphi) u^\sigma u_\sigma \right]. \tag{C.3.1}$$

where $u_\alpha = \eta_{\alpha\beta}u^\beta$ now. For a massive particle u^μ is a time-like vector, in the mostly-plus Lorentzian signature, $u^\mu u_\mu < 0$ we can perform a Lorentz transformation into the inertial frame of the test particle, which amounts to identifying, $u^\mu = (1, \vec{0})$,

$$F^\mu = \frac{\partial(\ln A(\varphi))}{\partial\varphi}\left[2u^\mu\dot{\varphi} + \eta^{\mu\nu}(\partial_\nu\varphi)\right]. \tag{C.3.2}$$

Choosing now to separate these into equations, for spacelike and timelike components $(\Pi, \vec{F}_5)$ now gives us,

$$\vec{F}_5 = -\frac{\partial(\ln A)}{\partial\varphi}\vec{\nabla}\varphi = -\vec{\nabla}(\ln A) \tag{C.3.3}$$

$$\Pi = \frac{\partial(\ln A)}{\partial\varphi}\dot{\varphi} \tag{C.3.4}$$

where Π is the power generated by the test particle accelerated by $\vec{F}_5$, which could be of interest. This is a derivation of the fifth force, $\vec{F}_5$ generated in the Einstein Frame.

C.4 Calculating the Fifth Force

We see that we can calculate the non-relativistic force on the test particle, all that needs to be assigned is the form of $A(\varphi)$. In this particular theory we use,

$$A(\varphi) = 1 + \frac{\varphi^2}{2M^2} \tag{C.4.1}$$

with higher order terms ignored. We see that $\ln(A) = \ln(1 + \varphi^2/2M^2)$, where for $\frac{\varphi}{M} \ll 1$, we can expand in powers of $\frac{\varphi}{M}$

$$\ln A(\varphi) \approx \frac{\varphi^2}{2M^2} + ... \tag{C.4.2}$$

which results in the following force,

$$\vec{F}_5 \approx -\frac{\varphi\vec{\nabla}\varphi}{M^2}. \tag{C.4.3}$$