

# The White Hole Existence Principle (Mathematics, Science and NS\* Version)

## WHEP Framework

Created by Prof. Nils Efverman

|

2025

Project by FÈUE      Contact: [feue.com1@outlook.com](mailto:feue.com1@outlook.com)

DISCLAIMER: THIS IS NOT ONLY FACTS, BUT ALSO A THEORY

FÈUE copyright 2025

Thing that is'ent yet in the document:

10.7 Halogens

10.7.1 Fluorine (F)

10.7.2 Chlorine (Cl)

10.7.3 Bromine (Br)

10.7.4 Iodine (I)

10.7.5 Astatine (At)

10.7.6 Tennessine (Ts)

10.8 Noble Gases

10.8.1 Helium (He)

10.8.2 Neon (Ne)

10.8.3 Argon (Ar)

10.8.4 Krypton (Kr)

10.8.5 Xenon (Xe)

10.8.6 Radon (Rn)

10.8.7 Oganesson (Og)

10.9 Lanthanides

10.9.1 Lanthanum (La)

10.9.2 Cerium (Ce)

10.9.3 Praseodymium (Pr)

10.9.4 Neodymium (Nd)

10.9.5 Promethium (Pm)

Written and created by Prof. Nils Efverman. This project is powered by FÈUE inc. FÈUE office website: feue256.github.io. Learn more about Gaia BH1 on:  
<https://tinyurl.com/3xceeda5>

10.9.6 Samarium (Sm)

10.9.7 Europium (Eu)

10.9.8 Gadolinium (Gd)

10.9.9 Terbium (Tb)

10.9.10 Dysprosium (Dy)

10.9.11 Holmium (Ho)

10.9.12 Erbium (Er)

10.9.13 Thulium (Tm)

10.9.14 Ytterbium (Yb)

10.9.15 Lutetium (Lu)

10.10 Actinides

10.10.1 Actinium (Ac)

10.10.2 Thorium (Th)

10.10.3 Protactinium (Pa)

10.10.4 Uranium (U)

10.10.5 Neptunium (Np)

10.10.6 Plutonium (Pu)

10.10.7 Americium (Am)

10.10.8 Curium (Cm)

10.10.9 Berkelium (Bk)

10.10.10 Californium (Cf)

10.10.11 Einsteinium (Es)

10.10.12 Fermium (Fm)

10.10.13 Mendelevium (Md)

10.10.14 Nobelium (No)

10.10.15 Lawrencium (Lr)

10.11 Superheavy Elements

10.11.1 Rutherfordium (Rf)

10.11.2 Dubnium (Db)

10.11.3 Seaborgium (Sg)

10.11.4 Bohrium (Bh)

- 10.11.5 Hassium (Hs)
- 10.11.6 Meitnerium (Mt)
- 10.11.7 Darmstadtium (Ds)
- 10.11.8 Roentgenium (Rg)
- 10.11.9 Copernicium (Cn)
- 10.11.10 Nihonium (Nh)
- 10.11.11 Flerovium (Fl)
- 10.11.12 Moscovium (Mc)
- 10.11.13 Livermorium (Lv)
- 10.11.14 Tennessine (Ts)
- 10.11.15 Oganesson (Og)

## 11. Molecules

### 11.1 Simple Diatomic Molecules

- 11.1.1 Hydrogen (H<sub>2</sub>)
- 11.1.2 Oxygen (O<sub>2</sub>)
- 11.1.3 Nitrogen (N<sub>2</sub>)
- 11.1.4 Carbon monoxide (CO)
- 11.1.5 Hydrogen chloride (HCl)

### 11.2 Water and Related Compounds

- 11.2.1 Water (H<sub>2</sub>O)
- 11.2.2 Hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>)
- 11.2.3 Hydroxide ion (OH<sup>-</sup>)

### 11.3 Carbon Compounds

- 11.3.1 Methane (CH<sub>4</sub>)
- 11.3.2 Ethane (C<sub>2</sub>H<sub>6</sub>)
- 11.3.3 Ethene (C<sub>2</sub>H<sub>4</sub>)
- 11.3.4 Ethyne (C<sub>2</sub>H<sub>2</sub>)
- 11.3.5 Glucose (C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>)

### 11.4 Salts

- 11.4.1 Sodium chloride (NaCl)
- 11.4.2 Potassium chloride (KCl)

### 11.4.3 Calcium carbonate (CaCO<sub>3</sub>)

## 11.5 Acids and Bases

### 11.5.1 Hydrochloric acid (HCl)

### 11.5.2 Sulfuric acid (H<sub>2</sub>SO<sub>4</sub>)

### 11.5.3 Nitric acid (HNO<sub>3</sub>)

### 11.5.4 Ammonia (NH<sub>3</sub>)

## 12. Computer Sciences

### 12.1 Classical Computers

#### 12.1.1 Hardware Components

##### 12.1.1.1 CPU (Central Processing Unit)

##### 12.1.1.2 GPU (Graphics Processing Unit)

##### 12.1.1.3 RAM (Random Access Memory)

##### 12.1.1.4 Storage Devices (HDD, SSD)

##### 12.1.1.5 Motherboard

##### 12.1.1.6 Input/Output Devices

#### 12.1.2 Software Components

##### 12.1.2.1 Operating Systems

##### 12.1.2.2 Applications

##### 12.1.2.3 Programming Languages

##### 12.1.2.4 Drivers

##### 12.1.2.5 Utilities

#### 12.1.3 Computing Concepts

##### 12.1.3.1 Binary System

##### 12.1.3.2 Algorithms

##### 12.1.3.3 Data Structures

##### 12.1.3.4 Networking Basics

##### 12.1.3.5 Security & Encryption

### 12.2 Quantum Computers

#### 12.2.1 Quantum Hardware

##### 12.2.1.1 Qubits

##### 12.2.1.2 Quantum Gates

- 12.2.1.3 Quantum Circuits
- 12.2.1.4 Quantum Memory
- 12.2.1.5 Quantum Error Correction
- 12.2.2 Quantum Software
  - 12.2.2.1 Quantum Algorithms
  - 12.2.2.2 Quantum Programming Languages
  - 12.2.2.3 Simulators
  - 12.2.2.4 Hybrid Classical-Quantum Systems
- 12.2.3 Quantum Concepts
  - 12.2.3.1 Superposition
  - 12.2.3.2 Entanglement
  - 12.2.3.3 Quantum Decoherence
  - 12.2.3.4 Quantum Cryptography
- 13. BHR and WHR:
  - 13.1 Black Hole Radiation (BHR)
  - 13.2 White Hole Radiation (WHR)
- 14. Black and White Holes content (in atom level):
  - 14.1 Black Hole
    - 14.1.1 The accretion disks
    - 14.1.2 The singularity
  - 14.2 White Hole
    - 14.2.1 The singularity
    - 14.2.2 White hole radiation (WHR)

## 15. Education In NS\*

- 15.1 Physics
  - 15.1.1 Quantum Mechanics
    - 15.1.1.1 Wave Functions
    - 15.1.1.2 Schrödinger Equation

15.1.1.3 Heisenberg Uncertainty Principle

15.1.1.4 Quantum Entanglement

15.1.2 Classical Physics

15.1.2.1 Newtonian Mechanics

15.1.2.2 Electromagnetism

15.1.2.3 Thermodynamics

15.1.2.4 Fluid Dynamics

15.1.3 Universe and Space

15.1.3.1 Cosmology

15.1.3.2 Relativity (Special and General)

15.1.3.3 Astrophysics

15.1.3.4 Particle Physics

15.2 Chemistry

15.2.1 Matter

15.2.1.1 States of Matter

15.2.1.2 Properties of Matter

15.2.2 Atoms

15.2.2.1 Atomic Structure

15.2.2.2 Electron Configuration

15.2.3 Subatomic Particles

15.2.3.1 Protons, Neutrons, Electrons

15.2.3.2 Quarks and Leptons

15.2.4 Periodic Table

15.2.4.1 Groups and Periods

#### 15.2.4.2 Element Properties

#### 15.2.5 Physical Chemistry

##### 15.2.5.1 Thermodynamics

##### 15.2.5.2 Kinetics

##### 15.2.5.3 Quantum Chemistry Basics

#### 15.2.6 Electrochemistry

##### 15.2.6.1 Redox Reactions

##### 15.2.6.2 Electrolysis

#### 15.2.7 Computational Chemistry

##### 15.2.7.1 Molecular Modeling

##### 15.2.7.2 Simulation Methods

#### 15.2.8 Quantum Chemistry

##### 15.2.8.1 Molecular Orbitals

##### 15.2.8.2 Schrödinger Equation Applications

### 15.3 Biology

#### 15.3.1 Bioinformatics

##### 15.3.1.1 Sequence Analysis

##### 15.3.1.2 Genomics and Proteomics

#### 15.3.2 Quantum Biology

##### 15.3.2.1 Photosynthesis Mechanisms

##### 15.3.2.2 Enzyme Dynamics

#### 15.3.3 Molecular Biology

##### 15.3.3.1 DNA/RNA Structure and Function

##### 15.3.3.2 Protein Synthesis

### 15.3.3.3 Cell Signaling

## 15.4 Earth Science

### 15.4.1 Meteorology

#### 15.4.1.1 Weather Systems

#### 15.4.1.2 Climate Dynamics

### 15.4.2 Oceanography

#### 15.4.2.1 Ocean Currents

#### 15.4.2.2 Marine Ecosystems

### 15.4.3 Astronomy

#### 15.4.3.1 Planetary Science

#### 15.4.3.2 Stellar Evolution

### 15.4.4 Environmental Science

#### 15.4.4.1 Pollution Studies

#### 15.4.4.2 Sustainability

### 15.4.5 Geology

#### 15.4.5.1 Rock and Mineral Formation

#### 15.4.5.2 Plate Tectonics

## 16. Education In Math

### 16.1 Pure Math

#### 16.1.1 Algebra

#### 16.1.2 Geometry

#### 16.1.3 Topology

#### 16.1.4 Number Theory

### 16.2 Applied Math

Written and created by Prof. Nils Efverman. This project is powered by FEUE inc. FEUE office website: feue256.github.io. Learn more about Gaia BH1 on: <https://tinyurl.com/3xceeda5>

16.2.1 Differential Equations

16.2.2 Mathematical Modeling

16.2.3 Optimization

16.3 Mathematical Finance

16.3.1 Risk Analysis

16.3.2 Financial Modeling

16.4 Structures (Algebraic/Geometric)

16.4.1 Groups, Rings, Fields

16.4.2 Vector Spaces

16.5 Number Systems

16.5.1 Integers, Rationals, Reals, Complex Numbers

16.6 Computer Science

16.6.1 Algorithms and Data Structures

16.6.2 Computational Complexity

16.7 Foundations of Mathematics

16.7.1 Logic

16.7.2 Set Theory

16.7.3 Proof Techniques

16.8 Complex Analysis

16.8.1 Analytic Functions

16.8.2 Cauchy's Theorem

16.9 Cryptography

16.9.1 Symmetric and Asymmetric Encryption

16.9.2 Hash Functions

Written and created by Prof. Nils Efverman. This project is powered by FEUE inc. FEUE office website: feue256.github.io. Learn more about Gaia BH1 on:  
<https://tinyurl.com/3xceeda5>

## 0. Content:

### 1.0 Introduction and information:

#### 1.1 Introduction:

The nearest black hole is 1560 light years away.

This black hole is believed to have formed around 13.75 billion years ago. The name of this black hole is "Gaia BH1". Under Gaia BH1's lifetime, Gaia BH1 has an average intake of  $3.15 \times 10^{23}$  kg of material. And if any material goes over the event horizon it never comes back. What we know is that when something has crossed the line from the outside it comes as if the object has stopped and never moves, but for the object comes it splits into pieces and a minute or so. When this piece comes to the singularity of Gaia BH1 (or other black hole) no one knows what happens, but that is what theory is for. So, we mostly know how a black hole can in the first place exist. The black hole creates when two supernovae collapse together, this happens when Big Bang exploded, and stars exploded and with the power off the force from Big Bang only then could two supernovae that exploded at the same time and collide. And now we know how a black hole creates we can look up what the black hole eats. A black hole can eat whatever is in the path of the black hole and this also includes the fourth-dimension time. When we know what it eats, we can check out what happens when we come singularly.

#### 1.2 Information:

\* NS = Natural Science

Natural Science has:

- Physics
- Chemistry
- Biology
- Earth Science

**DISCLAIMER: THIS IS NOT ONLY FACTS, BUT ALSO A THEORY**

For understanding WHEP as good as possible you need education in:

- Physics
- Chemistry
- Biology
- Earth Science

Written and created by Prof. Nils Efverman. This project is powered by FEUE inc. FEUE office website: feue256.github.io. Learn more about Gaia BH1 on:  
<https://tinyurl.com/3xceeda5>

- Math

Education about physics: Chapter 15.1

Education about chemistry: Chapter 15.2

Education about biology: Chapter 15.3

Education about earth science: Chapter 15.4

Education about math: Chapter 16.

In physics you need to learn about: all quantum mechanics and physics, classical physics and about the universe and space. In chemistry you must learn about matter, atoms, subatomic particles, the periodic table, physical chemistry, electrochemistry, computational chemistry and quantum chemistry. In biology you must learn about: bioinformatics, quantum biology and molecular biology. In earth science you need to learn: meology, meteorology, oceanography, astronomy and environmental science. In the language of all science; math you need to learn about: pure math, applied math, mathematical finance, structures, number systems, computer science, foundations, complex analysis and cryptography.

## 2. Dimensions and Formulas:

So, we know first a singularity is an object that has broken all laws of physics.

### 2.1 Dimensions:

Singularity is an object that we do not know very much about, but this theory says that singularity is the only object in the universe that has five dimensions:

1. Length (x)
2. Width (y)
3. Height (z)
4. Time (t)
5. Wong (w)

This fifth dimension is a dimension that has broken laws of physics. This dimension is called Wong. Wong is like the deep of the singularity. Wong can you not see with your eyes but feel it directly. Our standard 3D spectrum can only see three of the totals and with tools the fourth. Wong has a curved depth that we can feel but not see. And in a long spectrum you can also feel the fourth dimension. The 5D spectrum is like a rainbow, we can only see the standard color and in the 5D spectrum we can only see standard dimensions. Wong is also impossible to see Wong because it is in the object and that is also why an object cannot only have a Wong dimension. Wong is also circle, but it has deep that breaks the law of physics and how we see our own 3D spectrum forever. Now we can understand better what happened when you meet the singularity in a black hole. When you come to the singularity that second you cannot feel, you cannot see, you do not remember anything. This is because you can only follow the laws of physics (else it comes be catastrophe of the balance in the universe), but an object can break the laws of physics and that is why this happens. When this ten second that feels you has died has ended then you come die and all material of you

be pieces the size of an atom. These  $7 \times 10^{27}$  atoms come travel true a tunnel and the diameter of an atom.

## 2.2 Gaia BH1 Formulas:

Gaia GH1 form can simplest be explained by this:

$$D_n = 176 \cdot (1.125)^n$$

$$C_n = \pi \cdot 176^2 \cdot (1.125)^{2n}$$

$$R_n = \pi^2 \cdot 176^3 \cdot t \cdot (1.125)^{2n+nt}$$

Or for making the fifth-dimension core for the Gaia BH1 singularity (if it was possible):

[feue256.github.io/BH.html](https://feue256.github.io/BH.html)

## 3 Black Hole and White Hole Physics:

### 3.1 Gravitational Field Strength Equations:

And in the end, when the object has a diameter equal to that of an atom, then a white hole comes on the other side of the tunnel. A White Hole is created when the first material from the black hole comes. The creation of a white hole can be so long that it can be a black hole, and then in the middle of the tunnel it comes to create a super nova from the start, and it has a chance of creating infinity of black hole and destroying the world. But when the creation of a white hole successfully comes do the opposite of what a black hole does. A white hole also has singularity, but it has negative dimensions. Gravitational Field Strength Near the Singularity for a black hole is defined as:

$$g_{\text{black}}(r, w) = GM(r^2 + \gamma \sin^2(w))(1 - e^{-w^2})$$

Near the singularity of a black hole, space is curved not only in 3D but also warped by the Wong dimension. The  $\gamma \cdot \sin^2(w)$  acts like a dimensional distortion buffer, preventing division by zero as  $r \rightarrow 0$ , while  $(1 - e^{-w^2})$  models how the Wong dimension intensifies the curvature the deeper you go. Gravity here becomes "ultra-curved", but not infinite due to 5D constraints.

Gravitational Field Strength Near the Singularity for a white hole is defined as:

$$g_{\text{white}}(r, w) = GM(r^2 + \delta \cosh(w))(1 + \tanh(\beta w))$$

The gravitational field strength near a black hole singularity, extended with the Wong dimension, is defined as:

$$g(r, w) = G \cdot M \cdot r^2 \cdot e^w$$

Explanation:

A white hole emits matter and time, reversed from the black hole. The hyperbolic cosine  $\cosh(w)$  exponential dimensional repulsion, increasing rapidly with Wong depth. The  $\tanh(\beta w)$  term causes an energy rebound that stabilizes the repulsion. The negative sign flips the force direction: explosive instead of attractive.

### 3.2 Event Horizon and Singularities:

When you are in the tunnel behind the singularity with the Wong dimension your atoms then come feel that it has stop (if atoms could feel), but if you somehow could see it from outer space and if the tunnel was of glass then you would see the atoms in approximately 953.34 lightyears/secund. If you stand on the singularity, then if you also had a telescope and zoom in on Tellus then it would be the year 465 Anno Domini. Why is this happening? This is because the light has a speed of approximately 300 000 km/s (or 186 411 miles/s), so when we look at an object extremely far away in the universe then we see them as they were when the light left them. On the same set, if something could observe Tellus from Gaia BH1, it comes to see the past. The tunnel between the black hole and the white hole is very stable, but the first ten years after birth of the black hole is the tunnel extremely unstable. The start of the universe is what we all know is Big Bang, but it can be more truthful than what we know about our universe. It can be so that when the creation of a black hole and the unstable tunnel between the black hole and the white go into pieces and then a new black hole creates and go so infinity, then we have infinity of black holes. The first black hole is the biggest and it can eat another black holes. Now when it is only one black hole in the hole universe and from the spectrum of the hole universe, this black hole is extremely small like Big Bang. When so much material is in one place and cannot go to a white hole it becomes a white hole. The black hole comes in one Pico secunds ( $1 \text{ ps} = 10^{-12} \text{ s}$ ) into a white hole. This explosion is exactly like Big Bang, but the first world with the black hole how could that black hole come to the first universe? The answer to that is:

The fourth dimension Time in nothing can create material. This material can create a force like a black hole and when it has had all that material it can create a black hole. That means all material in the whole universe is from time itself.

### 3.3 Types of Black Holes:

Our universe that we have at the latest 13,8 billion years has approximately  $10^{21}$  black holes. That are three types of black holes what we know:

Stellar Black Holes
Intermediate-Mass Black Holes
Supermassive Black Holes

Stellar Black Holes creates when two supernovae collapse. This type is 3-20 times bigger than our sun.

Intermediate-Mass Black Holes are the middle size of black holes. This type of black hole is the rarest of the three normal black holes. This is 100-1000 times the size of our sun.

Supermassive Black Holes is the biggest type of black holes. This black hole is in nearly every galaxy in the universe. These black holes can be millions to billions of times bigger than our sun.

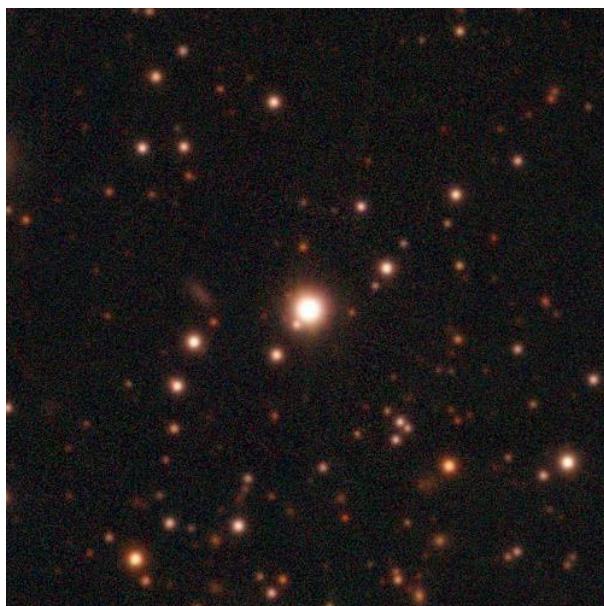
There are also two other types of black hole, but these ones are only theoretical:

Primordial Black Holes is a black hole that can only been created a brief time after Big Bang due to the frequent fluctuations in the early stage of the universe. The size of the Primordial Black Hole can variant all from extremely small to quite big.

Micro Black Holes are very very extremely small black holes. This Black Holes can be created in labs around the world. This would be especially useful för testing different theories.

### 3.4 White Hole Formation:

White Holes are terribly slow when they are spitting out everything from the black hole. This is because Supermassive Black Holes has a Micro White Hole. Every black hole has its opposite size than what the white hole has. The white hole is not a star, even if that is the opposite of what a black hole is. Every white hole is in another multiverse, and in this multiverse there are no black holes. When a one thousand multiverse it will become a mutivhope. Gaia BH1's real name is Gaia DR3 4373465352415301632. Image of Gaia BH1:



### 3.5 Gaia BH1 Case Study:

The tunnel between the black hole and white hole is the first material that the black hole eats at the start of the black hole life. This process can take up years. Where does the black color come from in a black hole? A black hole is black because light cannot escape from within its event horizon. Since no light can be reflected or emitted back, the black hole appears completely dark to any observer. The accretion disk on a black hole can reach temperatures up to ten million Celsius (eighteen million Fahrenheit). When you look at a black hole from the side of the black hole you see the accretion disk bend on the top, but when you look at the black hole above you see the accretion disk in a full circle. This is because the black holes gravity has a so enormous power and then it bends the light.

## 4. Cosmological Material Flow Analogy:

### 4.1 Waterfall Analogy for Black Holes:

The Supermassive Black Hole in the center of our galaxy Milky Way Sagittarius A\* is the size of 4,3 million times our sun. One of the biggest black holes we have ever seen is Ton 618. Ton 618 is sixty-six billion times our sun. Ton 618's dimeter can fit ~40 of our solar systems. If you look at a black hole you will never see anything, go into the black hole. As an example:

Material in space is water. The black hole is a fall. Then would even water go into the black hole and create a waterfall. There the edge of the waterfall is an event horizon. After the water has fallen out from the edge it is faster than the speed of light down.

When you are going into the singularity, gravity increases. When you are far away from the event horizon you come to start the spaghettification. Spaghettification is when long come exceptionally long to you sardar into pieces. This is happening when the gravity with your head is lower than what it is with your feet. In the singularity all time ends. When something crosses the event horizon, there is no way back. Only radiation escapes from the black hole. This radiation consists of old particles from the tunnel between the black hole and the white hole. The radiation is extremely radioactive – up to one thousand times more dangerous than Polonium-210. This radiation is called BHR. We can never see BHR because the BHR come disappears into the accretion disk. When the white hole explodes like Big Bang all stuff come to stay in the white hole in approximately 5 seconds and then put it in another black hole and time starts again and time itself creates the start practice for a new universe starts it is new life. If a black hole were the size as an atom it would weigh like the biggest mountain on Tellus.

## 4.2 Black Hole Mergers and White Hole Events:

We cannot see white holes because the tunnel between black and white holes is going true another universe with only white holes in it. If something goes over the Event Horizon, it cannot come back (if it is not BHR), on same set cannot anything go into a white hole. When two black holes eat another black hole, it can go three separate ways:

1. The big black hole eats the small and it be a large black hole.
2. Two black holes of equal mass are orbiting each other. When a third black hole approaches, the system becomes unstable — eventually leading to the end of their lives through a massive merger called the closing of a black hole system.
3. A small black hole consumes a larger one. This causes a white hole to form for just one picosecond ( $10^{-12}$  seconds), before the system explodes.

## 5. Relativity and Metrics:

As solutions to the Einstein field equations, black holes are defined in general relativity. These metrics detail the shape of spacetime around different type of black holes. Specifically, the following are two instances of systems of these invariants:

The Schwarzschild-metrics for a non-rotating, neutral black hole can be defined by:

$$ds^2 = - \left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$$

The Schwarzschild-metrics for a non-rotating, charged black hole can be defined by:

$$ds^2 = - \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right) dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2}\right)^{-1} dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2$$

The Schwarzschild metric is a solution to Einstein's field equations in general relativity that describes the spacetime around a non-rotating, uncharged (neutral) black hole or a massive spherical object.

## 5.2 Reissner-Nordström Metric:

Reissner–Nordström solution (for charged black holes):

$$r_{\pm} = \frac{GM}{c^2} \pm \sqrt{\left(\frac{GM}{c^2}\right)^2 - \frac{GQ^2}{4\pi\varepsilon_0 c^4}}$$

## 5.3 Kerr Metric:

Kerr solution (for rotating black holes):

$$ds^2 = -\left(1 - \frac{2GMr}{\rho^2 c^2}\right) c^2 dt^2 - \frac{4GMar \sin^2 \theta}{\rho^2 c} dt d\phi + \frac{\rho^2}{\Delta} dr^2 + \rho^2 d\theta^2 + \left(r^2 + a^2 + \frac{2GMa^2 r \sin^2 \theta}{\rho^2 c^2}\right) \sin^2 \theta d\phi^2$$

## 5.4 Hawking Radiation (BHR) and Temperature:

This equation below is Stephen Hawking temperature that a black hole gives out in BHR:

$$T_H = \frac{\hbar c^3}{8\pi G M k_B}$$

$\hbar$  = reduced Planck constant

$k_B$  = Boltzmann's constant

The White holes singularity also has the wong dimension in it. The BH in Gaia BH1 stands for Black Hole and the one in Gaia BH1 stands for the first black holes that is found in the Gaia project by ESA. The European Space Agency (ESA) serves as the European counterpart to NASA, focusing on space exploration, satellite development, and scientific research. ESA's head office is in Paris, France. ESA has twenty-two member countries. At the time of writing this is Sławosz Uznański-Wiśniewski from Poland on ISS (International Space Station). ISS is a space station 400 km (248.55 miles) above Tellus. ISS was starting construction in 1998. ISS has been mannded in nearly 25 years 24 hours a day in 7 day a week. ISS weighs 420 000 kg, and it has a speed of 28 000 km/h (17398.39/h). ISS size is approximately as big as a football field. Participating countries include NASA (USA), Roscosmos (Russia), ESA (Europe), JAXA (Japan), CSA (Canada), and many others.

# 6. Time, Material Creation, and Wong Dimension:

## 6.1 Time Creating Matter:

How can time create from nothing create something?

This is because the time near a white hole changes to the opposite of how time changes when the black hole changes it. The black hole can take the BHR before the BHR has exploded and save it into the event horizon, later the black hole can eat up the old event horizon and it can send in small packages of BHR to the white hole. Then the white hole takes this small packages with a force out from the singularity in the white holes center. Now the time can build a shield over the BHR, and it can build a stone. This process can take up to billions and billions of light years.

## 6.2 Wong Dimension Explanation:

What is the physics or mathematical nature of the Wong dimension, and how can it be experimentally or observationally verified to exist beyond theoretical speculation?

This dimension can only be found in the singularity of a black hole, which is why we cannot experiment with it. If we try, we come die by spaghettification. And the singularity of a black hole cannot exist without breaking the laws of physics, because whatever that is passes the event horizon that's not BHR (and it's new) can go into the singularity of the black hole. As example our sun cannot go into a space that is in this center of the black hole it is most be something behind. And it is with the wong and going into the tunnel to the white hole.

The wong dimension curves because the tunnel to the white hole is always not straight. The wong dimension is time like it always goes to the white hole if the tunnel is stable and if it is not stable it is spacelike. The wong and singularity are not objects it's something we can't understand. Wong is timelike but it's also curved. In the tunnel between a black and white hole all math and physics fail, and all that's not the transaction of atoms fails. When all fail, wong can't fail. When all else fails, Wong does not. The Wong dimension continues no matter what.

## 6.3 Tunnel Between Black and White Hole:

The tunnel between the black hole and the white hole is made of a unique material that naturally forms as element 119 to 139 in the periodic table, called Hongstone. If we ever discover Hongstone, it will officially become the 119 to 139th element, opening new doors to understanding black holes and white holes like never before. This discovery could allow us to build a stable tunnel connecting a micro black hole to a white hole. The tunnel must be made of Hongstone; otherwise, its instability would cause a supernova-like explosion, breaking Tellus (Earth) into tiny pieces. However, we face three major challenges. First, the Hongstone tunnel must have a diameter as small as a single atom. Second, we must create a micro black hole the size of an H<sub>2</sub>O molecule but with the mass of Jupiter. Third, and most difficult, the white hole must be supermassive to perfectly balance with the micro black hole. Solving these problems would help us finally understand what the singularity at the center of a black hole truly is and how these cosmic phenomena work. For it to work it must be in quantum gravity. Quantum gravity isn't possible with the tool we have today. Quantum gravity can be possible with a quantum-computer-chamber on ISS with zero gravity and if it's in a black hole. This project is very dangerous because we don't know how much the black hole comes eat and the force from the white hole, if something goes wrong the whole ISS can go into pieces, (but would that second after because they must be in a black hole). We must have a black hole for creating another black hole.

## 6.4 Hongstone:

Atomic Number: 139

Hypothetical Symbol (WHEP): Hsx

Isotopic Structure (WHEP): one proton, one neutron, 182,983 electrons, 5 HE-bosons

Hongstone is a hypothetical superheavy gaseous element, predicted to exist under extreme WHEP conditions and in the tunnel between the black hole and the white hole. Its minimal nucleus consists of a single proton and one neutron, while the 182,983 electrons form a highly diffuse, high-energy electron cloud. This configuration allows the atom to remain in a gaseous phase, despite its enormous electron count.

The 5 HE-bosons embedded in the nucleus mediate intense hyperelectric interactions, which govern the behavior of both the proton-neutron core and the surrounding electron cloud. In the gas phase, these interactions generate dynamic electron shell fluctuations and HE-quark excitations, producing a highly reactive yet non-condensed atomic state.

As a gas, Hongstone exhibits:

- Extreme hyperelectric conductivity: Electron clouds move freely while interacting with HE-bosons, creating transient fields.
- High quantum fluctuation rates: Valence electrons and HE-core couplings produce rapid spin-dependent excitations.
- Minimal nuclear influence: The single proton and neutron nucleus are dominated by HE-boson effects, enabling clear observation of WHE-quark dynamics.
- Unique gas-phase properties: Low interatomic cohesion allows free expansion, while HE-boson-mediated interactions maintain subtle quantum correlations between atoms.

WHEP predicts that gaseous Hongstone is the ultimate testbed for hyperelectric and nuclear quantum phenomena, where extreme electron populations and HE-boson interactions can be studied without solid-state or liquid-phase interference. Its gaseous state provides direct experimental access to WHE-quark excitations and HE-core energy shifts across an isolated atomic system.

## 6.5 Quantum physics:

Hongstone's strength is due to its structure using quantum bits for stability on the Wong scale. Unlike regular atomic bonds, Hongstone forms connections that are not held together by traditional forces like electromagnetism, but by entangled qubit states that only exist under Wong dimensional curvature. These qubit bonds allow the tunnel walls between the black hole and white hole to self-repair and resist spaghettification, even as the fabric of spacetime collapses around them. Without this quantum-layered strength, any normal matter would disintegrate instantly under the stress of Wong field fluctuations. That's why Hongstone isn't just a material, it's a quantum-stabilized state of existence. Hongstone has come to be the first element with computers and quantum-physics in an element. The Black Holes we can create with hongstone can only last a secund before the explode. Hongstone can also be used as the main material in CPU's for quantum computers. This CPU's can make and find patters in what ever you give then. This is only going to be possible with quantum computers, and they most have this special hongstone CPU's. This computers can solve math problems that we see as impossible, but this computers can solve this problem in seconds. This can also lead to a world of AI that's so smart that it can do things we believe is thing we think of because today we don't understand how to do this thing. If hongstone creates or find's then nobody is allowed to do anything with it because in specific condition can it open a portal to another dimension and that's would be extremely dangerous. We don't know the conditions needed for this to happen and before that it must be illegal to owe it because of this danger that can happen. What we know may be right, but we could still be wrong.

## 7. Pico Doom Virus (PDV) and Cosmic Implications:

### 7.1 PDV Behavior in Black Hole Singularities:

The Black Hole Radiation (BHR) is not only a source of extreme ionizing radiation but also a carrier of a theoretical pathogen referred to as Pico Doom Vires (PDV). PDV is a bio-quantum agent with an effective reproduction number ( $R_0$ ) of infinity. Unlike conventional viruses, PDV does not require cellular machinery to replicate; instead, it propagates through fundamental interactions at the subatomic level, targeting nucleic acid structures (DNA/RNA) directly. Upon exposure—whether airborne, dermal, or via particulate carriers—all known biological systems experience immediate and irreversible molecular collapse. This makes PDV not only the most infectious agent theoretically proposed but also a universal terminator of life, functioning beyond the constraints of biochemistry as we know .

The presence of PDV within BHR emissions suggests that certain black holes may serve not only as gravitational singularities but also as distribution hubs for entropic viral fields, potentially reshaping our understanding of cosmic sterilization events. This raises serious implications for interstellar exploration, planetary quarantine protocols, and the theoretical upper bounds of biological survivability. The Pico Doom Virus (PDV) has a theoretical survival probability of 0% under standard conditions in the observable universe due to its extreme bio-instability outside of specialized quantum environments. However, within the singularity of a black hole, this survival probability approaches 100%.

This anomaly arises due to the black hole's ability to warp quantum physical laws. As spacetime and quantum fields are intensely curved near or within the event horizon, standard probabilities—such as decay or disintegration rates—no longer apply in their classical form. Since PDV is classified as a bio-quantum virus, it is intrinsically entangled with quantum mechanical behavior rather than classical biological constraints.

### 7.2 Implications for Interstellar Exploration:

As quantum mechanics is deformed near the singularity, so too is the statistical framework governing particle and viral decay. This deformation results in a unique condition in which PDV's quantum structure stabilizes rather than deteriorates, effectively reversing its normal survival odds. The PDV is the size of a picometer.

The black hole can grow because the BHR can build a magnetic field and when the PDV particle collides with the magnetic field it can create an intense gravitational field and makes stone that the black hole eats and then grows.

### 7.3 Impact on DNA/RNA Structure:

The interaction of PDV with DNA and RNA in biological systems is one of the most significant aspects of this theoretical virus. Its bio-quantum nature suggests that it could fundamentally alter the structure and function of genetic material in ways we have yet to fully comprehend.

#### 1. Direct Molecular Collapse:

- PDV targets nucleic acids (DNA and RNA) directly at a quantum level, bypassing the need for a host cell. This would result in irreversible molecular collapse upon exposure, as the virus interacts with the subatomic structure of the nucleic acids.

- The viral particles could induce quantum fluctuations in the chemical bonds that hold the nucleotides together, causing them to break apart, resulting in severe genetic instability.

2. Quantum Entanglement with DNA/RNA:

- Due to its quantum nature, PDV could be capable of entangling with DNA or RNA molecules, causing the virus to share the same quantum states as the nucleic acids. This interaction would likely destabilize the double helix structure of DNA, potentially leading to structural distortions and breaks in the strands.
- The molecular collapse could disrupt the genetic code, causing random mutations or deletions of genetic sequences that are essential for cellular function.

3. Altered Replication and Transcription:

- PDV-induced changes in the genetic material could block or distort DNA replication and RNA transcription. This disruption would prevent normal gene expression and protein synthesis, leading to cellular dysfunction.
- The virus may also alter the three-dimensional folding of RNA molecules, which is essential for their proper functioning in protein synthesis.

4. Epigenetic Changes:

- PDV might induce epigenetic changes in DNA that affect how genes are expressed without altering the underlying sequence. For example, it could modify DNA methylation patterns or histone modifications, leading to long-term changes in gene expression.
- Such epigenetic shifts could permanently affect an organism's biology, potentially rendering the infected species nonviable or causing their extinction.

5. Unpredictable Biological Effects:

- Since PDV operates beyond the typical constraints of biochemistry, the exact impact it would have on RNA and DNA is unpredictable. In some cases, it could lead to revolutionary changes in genetic makeup, potentially altering evolutionary processes or creating entirely new forms of life.
- This unpredictability introduces a new level of biological uncertainty when dealing with cosmic pathogens, where traditional biological models would not suffice to explain the virus's behavior.

## 8. Online data:

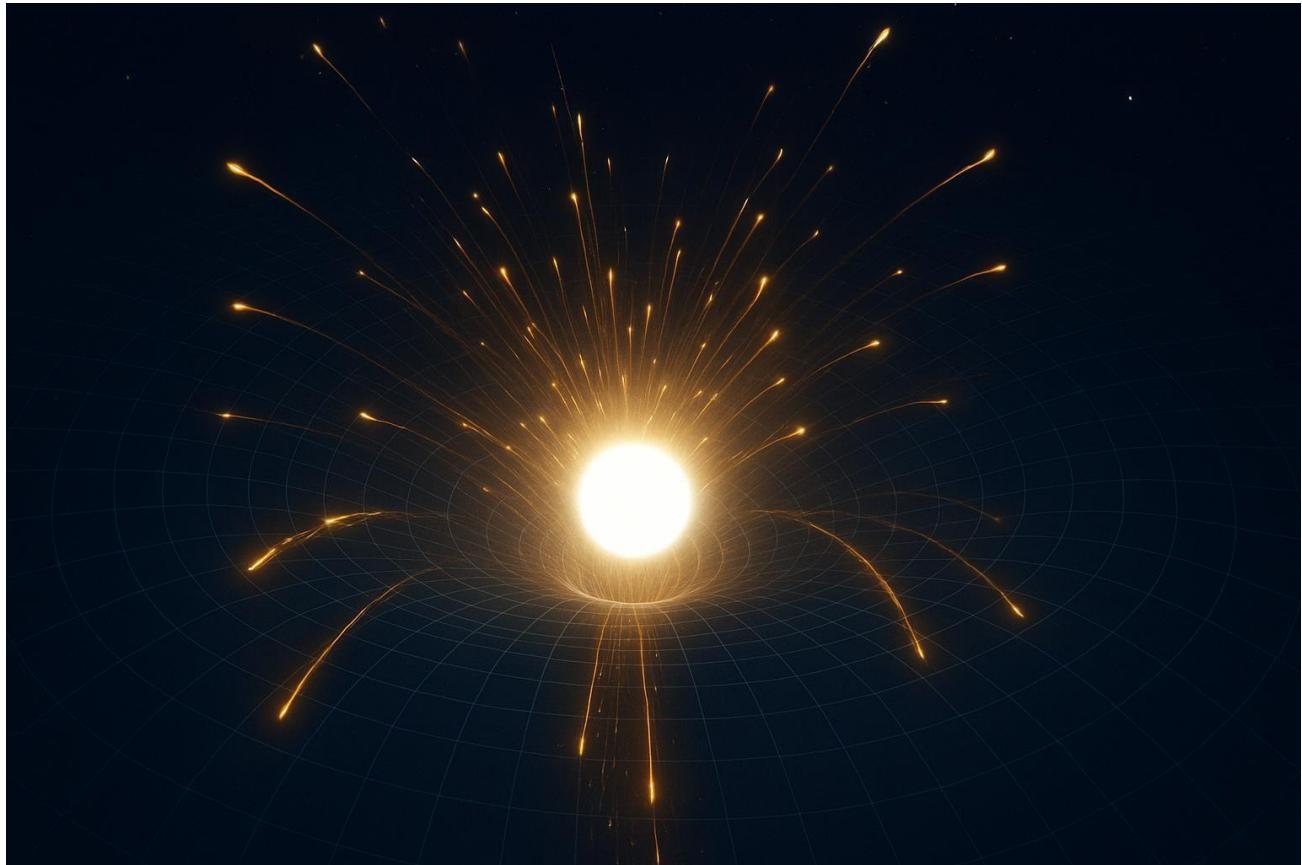
### 8.1 Description of how light bends:

Data Source:

[https://en.wikipedia.org/wiki/Black\\_hole#/media/File:Black\\_Hole\\_Shadow.gif](https://en.wikipedia.org/wiki/Black_hole#/media/File:Black_Hole_Shadow.gif)

The GIF is playing a short video that describes how light bends near a black hole. The yellow is the light, and the black is the singularity.

This tells us how light waves react to a static black hole and how the black holes gravity reacts to the light waves. This GIF is from the wiki article [en.wikipedia.org/wiki/Black\\_hole](https://en.wikipedia.org/wiki/Black_hole) under the label “Physical properties”



### 8.3 Description of quantum particles / the standard model of particle physics:

Data Source :

<https://live-production.wcms.abc-cdn.net.au/2c70e6d5c36be4979a8f1be0e10d3810?src>

This image tells us two separate groups of particles in the Standard Model. The first group is the fermions, and the second group is the bosons. The fermions are the particles that make all matter in the universe. The bosons are the particles that are force carriers or exchange particles. In the fermions there are two subgroups in the fermions, the quarks, and the leptons. The bosons have the same principle; they also have two separate groups as the fermions. They have the gauge bosons and the scale boson. The thing that tells them apart is the spin of the particles. The fermions have a spin of  $\frac{1}{2}$  and the bosons have a spin of one or for the scalar boson is zero.

## 6. The standard model of particle physics:

### 6.1 Fermions:

#### 9.1.1 Quarks:

##### 9.1.1.1 Up quark:

The up quark is one of the most popular quarks. It has noticeably mass at only approximately  $2,3 \text{ MeV}/c^2$ . This makes the up quark very stable. That's why the up quark is extremely popular in the universe. The up quark has a spin of  $\frac{1}{2}$  and a charge of  $\frac{2}{3}$ . The symbol for the up quark is "u", the first letter of up.

##### 9.1.1.2 Down quark:

The down quark is a quark with the symbol "d" because that is the first letter of the quark full name (down). It's the most popular quark on the same place as the up quark. This quark has the same spin as all quarks ( $\frac{1}{2}$ ) and a charge of  $-\frac{1}{3}$ . The mass of the down quark is a little bite heavier than the up quark on approximately  $4.8 \text{ MeV}/c^2$ . It's still stable. The down, strange and the bottom quark always have the same charge, and the up, charm and top quarks also have the same charge.

##### 9.1.1.3 Charm quark:

The charm quark is a heavier quark with the symbol "c", the first letter of charm. It has a spin of  $1/2$  and a charge of  $2/3$ , just like the up quark. The mass of the charm quark is approximately  $1.27 \text{ GeV}/c^2$ , making it much heavier than the up and down quarks. The charm quark is less common in everyday matter but appears in high-energy particle collisions.

##### 9.1.1.4 Strange quark:

The strange quark has the symbol "s", the first letter of strange. It has a spin of  $1/2$  and a charge of  $-1/3$ , the same as the down and bottom quarks. The mass of the strange quark is around  $95 \text{ MeV}/c^2$ . Strange quarks are responsible for forming "strange" particles and appear in cosmic events and particle experiments.

##### 9.1.1.5 Top quark:

The top quark is the heaviest quark and has the symbol "t", from top. Its spin is one-half, and its charge is two-thirds, like the up and charm quarks. The mass of the top quark is about  $173 \text{ GeV}/c^2$ . It is extremely unstable and decays almost immediately after formation, so it is only seen in high-energy accelerators.

##### 9.1.1.6 Bottom quark:

The bottom quark has the symbol "b", from bottom. It has a spin of  $1/2$  and a charge of  $-1/3$ , like the down and strange quarks. Its mass is approximately  $4.18 \text{ GeV}/c^2$ . The bottom quark is heavier than the strange and down quarks but lighter than the top quark. It forms particles called bottom-hadrons, such as B-mesons, and is important in studies of quantum effects like CP-symmetry. In particle physics CP means "Change Parity".

### 9.1.1.7 WHE Quark:

The WHE quark is a hypothetical particle introduced in the context of the White Hole Existence Principle (WHEP). It is theorized to be a key component of the exotic matter associated with white holes and their unique properties. The WHE quark may interact with the newly defined hyperelectromagnetic force (HEF), mediated by the HE-boson.

Properties:

- Charge: The WHE quark could possess a unique type of hyper electric charge that is linked to the hyperelectromagnetic force (HEF). This charge would allow it to interact with the HE-boson, the mediator of the hyperelectromagnetic force.
- Mass: The mass of the WHE quark may vary depending on its interaction with gravitational fields and exotic matter near white holes.
- Spin: Like other quarks, the WHE quark is predicted to have spin one-half.
- Interactions: The WHE quark is theorized to interact with other quarks via the strong nuclear force, mediated by gluons, but it may also participate in hyperelectromagnetic interactions through the HE-boson. Furthermore, it may interact with gravitons, especially in the environment of white holes.

Theoretical Role:

- White Hole Matter: The WHE quark is expected to be a fundamental particle that contributes to the exotic matter found inside white holes. It may help explain the emission of energy and particles from these objects.
- HEF Interaction: Due to its unique hyper electric charge, the WHE quark would play a critical role in mediating interactions within the hyper electromagnetic force (HEF), which is described by the HE-boson. This interaction could explain the formation and dynamics of exotic matter within and around white holes.

Mathematical Representation:

The WHE quark is included in the quantum field theory of WHEP. A possible Lagrangian term for the WHE quark's interactions is as follows:

$$\begin{matrix} L \\ W \\ H \\ E \\ = g \\ W \\ H \\ E \\ \psi \\ - \end{matrix}$$

$$\begin{matrix} W \\ H \\ E \\ \gamma \\ \mu \\ \psi \\ W \\ H \\ E \\ \cdot H \\ E \\ \mu \\ +g \\ W \\ H \\ E \\ -s \\ t \\ r \\ o \\ n \\ g \\ \psi \\ - \end{matrix}$$

$$\begin{matrix} W \\ H \\ E \\ \gamma \\ \mu \\ \psi \\ W \\ H \\ E \\ \cdot G \\ \mu \end{matrix}$$

In this equation:

- $\frac{g}{W}$  is the coupling constant for the interaction between the WHE quark and the  $H_E$  hyperelectromagnetic field.
- $gWHE\text{-strong}$  is the coupling constant for the strong interaction between the WHE quark and gluons.
- $\frac{H_E}{\mu}$  represents the hyperelectromagnetic field.
- $\mu$  is the gluon field, which mediates the strong interaction.

This Lagrangian expresses the coupling between the WHE quark and the HE-boson, as well as its interaction with gluons.

## 6.1.2 Leptons:

### 9.1.2.1 Electron:

The electron is a fundamental particle, classified as a lepton. It has a negative electric charge and is not composed of smaller particles like quarks. The electron has a spin of  $\frac{1}{2}$  and a mass of approximately  $0.511 \text{ MeV}/c^2$ .

### 9.1.2.2 Muon:

The muon is like the electron but has a much greater mass. It is unstable and decays into electrons, neutrinos, and other particles. Its mass is approximately  $105.7 \text{ MeV}/c^2$  and it has a spin of  $\frac{1}{2}$ .

### 9.1.2.3 Tau:

The tau is the heaviest of the charged leptons. It is unstable and decays into lighter particles such as electrons or muons and neutrinos. The tau's mass is  $1.777 \text{ GeV}/c^2$  and it has a spin of  $\frac{1}{2}$ .

### 9.1.2.4 Electron Neutrino:

The electron neutrino is a neutral lepton associated with the electron. It has an exceedingly small mass, which is still under investigation, and does not carry any electric charge. Its spin is  $\frac{1}{2}$ .

### 9.1.2.5 Muon Neutrino:

The muon neutrino is a neutral lepton associated with the muon. It has an exceedingly small mass, and like other neutrinos, interacts weakly with matter. Its spin is  $\frac{1}{2}$ .

### 9.1.2.6 Tau Neutrino:

The tau neutrino is a neutral lepton associated with the tau particle. It has an exceedingly small mass and interacts very weakly with matter. Its spin is  $\frac{1}{2}$ .

## 6.2 Bosons:

Bosons are particles that carry the fundamental forces of nature. They have integer spin (0, 1, 2...) and play a critical role in mediating interactions between other particles.

### 9.2.1 Gauge Bosons:

Gauge bosons are the force-carrier particles that mediate the fundamental forces in the Standard Model of particle physics. These include the gluon, photon, W bosons, and the Z boson.

#### 9.2.1.1 Gluon:

The gluon is the force carrier for the strong nuclear force, which binds quarks together inside protons and neutrons. Gluons are massless and interact only with quarks and other gluons. This interaction is fundamental to the structure of atomic nuclei.

#### 9.2.1.2 Photon:

The photon is the force carrier for electromagnetic interactions. It is responsible for electromagnetic radiation (such as light) and has zero mass and no electric charge. Photons mediate the interaction between electrically charged particles.

#### 9.2.1.3 Z Boson:

The Z boson mediates the weak nuclear force and is electrically neutral. It is involved in processes like neutrino interactions and beta decay. The Z boson, along with the W boson, is crucial in explaining phenomena such as the decay of unstable particles.

#### 9.2.1.4 W Boson:

The W boson mediates the weak nuclear force, which is responsible for particle decay processes like beta decay. There are two types of W bosons:

- $W^+$ : The positively charged W boson.
  - $W^-$ : The negatively charged W boson.
- The W boson plays a key role in interactions that change the type (or flavor) of particles.

#### 9.2.1.5 Gravitons:

The graviton is the one of two new thing in the boson group for the subatomic particle system. The graviton is a force carrier. This new boson has an unlikely spin of two that is specific to the graviton boson. This boson come have the force of gravity. This particle is not really used in the universe, but it has a minor impact on the hongstones atomic core.

#### 9.2.1.6 HE boson:

The HE-boson (Hyperelectromagnetic Boson) is a hypothetical particle introduced to mediate the hyperelectromagnetic force (HEF) within the framework of the White Hole Existence Principle (WHEP). The HE-boson is theorized to play a significant role in interactions involving exotic matter and WHE quarks, particularly in the environment of white holes.

Properties:

- Charge: The HE-boson is expected to be electrically neutral but is involved in the mediation of the hyperelectromagnetic force through its interaction with WHE quarks and other exotic particles.
- Mass: The HE-boson may possess a non-zero mass, which could vary depending on the gravitational conditions within a white hole. The mass is speculated to be relatively small, though it could fluctuate due to interactions with gravitational fields.
- Spin: The HE-boson is theorized to have spin one, classifying it as a gauge boson. This means it is responsible for mediating the hyperelectromagnetic force, much like the photon mediates the electromagnetic force or the gluon mediates the powerful force.
- Interaction: The HE-boson mediates the hyperelectromagnetic force (HEF), which governs the interactions of WHE quarks and other exotic matter. It is a critical component of the theoretical structure of white holes and the exotic matter contained within them.

Theoretical Role:

- Hyperelectromagnetic Force: The HE-boson is hypothesized to be the force carrier for the hyperelectromagnetic force (HEF). This force is an extension of the electromagnetic force, theorized to act on particles with hyper electric charge, such as the WHE quark. The HE-boson would be responsible for transferring energy between these particles.
- White Hole Dynamics: In the context of white holes, the HE-boson could explain the interactions between exotic matter and the energy released by these objects. The HE-boson is theorized to be central to the unique radiation and particle emission processes in white hole environments.
- Potential Gravitational Coupling: Although the HE-boson is primarily associated with the hyperelectromagnetic force, there may also be indirect interactions with gravitational fields, especially in the extreme conditions of a white hole.

Mathematical Representation:

The interaction of the HE-boson with WHE quarks and other exotic particles can be described through the following Lagrangian term:

$$\begin{array}{c}
 L \\
 H \\
 E \\
 = g \\
 H \\
 E \\
 \psi \\
 - \\
 W \\
 H \\
 E \\
 \gamma \\
 \mu \\
 \psi \\
 W \\
 H \\
 E \\
 \cdot H \\
 E \\
 \mu
 \end{array}$$

Where:

$\frac{g}{H}$  is the coupling constant for the interaction between the WHE quark and the HE-boson.  
 $E$

- $\begin{array}{cc} \psi \\ - \\ W \\ H \\ E \end{array}$  and  $\begin{array}{cc} \psi \\ - \\ W \\ H \\ E \end{array}$  are the WHE quark fields.
- $HE\mu HE\mu HE\mu$  represents the hyperelectromagnetic field, mediated by the HE-boson.

This term describes the interaction of the WHE quark with the hyperelectromagnetic force, mediated by the HE-boson.

## 9.2.2 Scalar Boson

A scalar boson is a type of elementary particle that has no directionality of spin, meaning it has a spin of zero. Unlike vector bosons, which mediate forces and have a directional spin, scalar bosons are theorized to interact differently, primarily in the context of mass generation for fundamental particles. Scalar bosons are part of many quantum field theories, and they often play a critical role in the mechanism that gives mass to other particles in the Standard Model.

### 9.2.2.1 Higgs Boson

The Higgs boson is a specific type of scalar boson that is central to the Higgs mechanism in the Standard Model of particle physics. It was proposed in the 1960s by physicist Peter Higgs and later confirmed by the Large Hadron Collider at CERN. The discovery of the Higgs boson in 2012 provided strong experimental support for the existence of the Higgs field, which is thought to give mass to other particles through the Higgs mechanism.

Written and created by Prof. Nils Efverman. This project is powered by FÈUE inc. FÈUE office website: feue256.github.io. Learn more about Gaia BH1 on: <https://tinyurl.com/3xceeda5>

Higgs and others, as part of the mechanism that explains how fundamental particles acquire mass.

Properties of the Higgs Boson:

- Spin: The Higgs boson has a spin of zero, meaning it is a scalar boson.
- Mass: The Higgs boson has a relatively large mass compared to other fundamental particles, approximately  $125 \text{ GeV}/c^2$ .
- Charge: The Higgs boson is electrically neutral.

Role in the Standard Model:

The Higgs boson is the quantum excitation of the Higgs field, a scalar field that permeates all of space. The interaction of particles with the Higgs field gives them mass. This process is essential because, without the Higgs field, particles such as the W and Z bosons (responsible for the weak nuclear force) would be massless, and the weak force would not behave as we observe in nature.

The Higgs mechanism explains the mass of fundamental particles by interacting with them. The stronger a particle interacts with the Higgs field, the more massive it becomes. For example, the top quark, which has a large mass, interacts very strongly with the Higgs field, while the photon (the force carrier for electromagnetism) does not interact with the Higgs field at all, and remains massless.

Discovery:

The discovery of the Higgs boson was one of the most significant milestones in modern physics. It was detected in July 2012 by scientists at the Large Hadron Collider (LHC) at CERN, confirming the existence of the Higgs field and validating the Standard Model. The discovery was awarded the 2013 Nobel Prize in Physics to Peter Higgs and François Englert.

Implications:

The detection of the Higgs boson not only confirmed the existence of the Higgs field but also solidified our understanding of the mass generation mechanism in the universe. It has profound implications for further research in particle physics, particularly in areas such as:

- Quantum Field Theory (QFT): The Higgs boson provides crucial evidence for the existence of scalar fields in quantum field theory.
- New Physics: While the discovery of the Higgs boson completed the Standard Model, it has also raised questions about what lies beyond it, such as potential connections to dark matter, supersymmetry, and other exotic theories.

Conclusion:

The Higgs boson is a pivotal component of the Standard Model and is critical for understanding the mass of elementary particles. Its discovery has profound implications not only for particle physics but also for the fundamental understanding of the universe's structure and the forces at play within it.<sup>10</sup>

## 10. Atoms

Atoms are what everything is made of, on a bigger level them quarks. All atoms are sorted in the periodic table of elements. Picture of the periodic table of elements:

**Periodic Table of the Elements**

1 IA																		18 VIIIA														
1 H Hydrogen 1.008	2 Be Boron 9.0122	3 Li Lithium 6.94	4 Be Beryllium 9.0122	5 VIB	6 VIIIB	7 VIIIB	8 VIIIB	9 VIIIB	10 VIIIB	11 IB	12 IIB	13 IIIA	14 IVA	15 VA	16 VIA	17 VIIA	18 VIIIA															
11 Na Sodium 22.990	12 Mg Magnesium 24.305	13 IIIB	14 IVB	15 VB	16 VIB	17 VIIIB	18 VIIIB	19 VIIIB	20 VIIIB	21 IB	22 IIB	23 Sc Scandium 44.956	24 Ti Titanium 46.942	25 V Vanadium 50.942	26 Cr Chromium 51.996	27 Mn Manganese 54.938	28 Fe Iron 55.845	29 Co Cobalt 58.937	30 Ni Nickel 58.934	31 Cu Copper 63.544	32 Zn Zinc 65.438	33 Ga Gallium 69.723	34 Ge Germanium 72.630	35 As Arsenic 74.922	36 Se Selenium 78.911	37 Br Bromine 79.904	38 Kr Krypton 83.798					
19 K Potassium 39.098	20 Ca Calcium 40.078	21 Ca Calcium 40.078	22 Sc Scandium 44.956	23 Ti Titanium 46.942	24 V Vanadium 50.942	25 Cr Chromium 51.996	26 Mn Manganese 54.938	27 Fe Iron 55.845	28 Co Cobalt 58.937	29 Ni Nickel 58.934	30 Cu Copper 63.544	31 Zn Zinc 65.438	32 Ga Gallium 69.723	33 Ge Germanium 72.630	34 As Arsenic 74.922	35 Se Selenium 78.911	36 Br Bromine 79.904	37 Kr Krypton 83.798														
37 Rb Rubidium 85.448	38 Sr Strontium 87.621	39 Y Yttrium 88.905	40 Zr Zirconium 91.224	41 Nb Niobium 92.906	42 Mo Molybdenum 95.941	43 Tc Technetium 95.941	44 Ru Ruthenium 96.941	45 Rh Rhodium 98.941	46 Pd Palladium 98.62	47 Ag Silver 101.87	48 Cd Cadmium 101.47	49 In Indium 113.40	50 Sn Tin 118.71	51 Sb Antimony 121.77	52 Te Tellurium 127.60	53 I Iodine 126.90	54 Xe Xenon 131.29	55 Cs Cesium 132.91	56 Ba Barium 137.33													
55 Cs Cesium 132.91	56 Ba Barium 137.33	57-71 Lanthanides	72 Hf Hafnium 178.49	73 Ta Tantalum 180.95	74 W Tungsten 183.84	75 Re Rhenium 184.21	76 Os Osmium 190.23	77 Ir Iridium 192.22	78 Pt Platinum 191.08	79 Au Gold 196.97	80 Hg Mercury 200.59	81 Tl Thallium 204.38	82 Pb Lead 207.2	83 Bi Bismuth 208.98	84 Po Polonium (209)	85 At Astatine (210)	86 Rn Radon (222)	87 Fr Francium (223)	88 Ra Radium (226)													
87 Fr Francium (223)	88 Ra Radium (226)	89-103 Actinides	104 Rf Rutherfordium (260)	105 Db Dubnium (261)	106 Sg Seaborgium (262)	107 Bh Bohrium (263)	108 Hs Meitnerium (265)	109 Mt Mendelevium (268)	110 Ds Darmstadtium (280)	111 Rg Roentgenium (281)	112 Cn Copernicium (285)	113 Nh Nihonium (286)	114 Fl Florium (289)	115 Mc Meissium (290)	116 Lv Livermorium (292)	117 Ts Tennessine (293)	118 Og Oganesson (294)	57 La Lanthanum (139)	58 Ce Cerium (140)	59 Pr Praseodymium (141)	60 Nd Neodymium (142)	61 Pm Promethium (143)	62 Sm Samarium (145)	63 Eu Europium (152)	64 Gd Gadolinium (154)	65 Tb Terbium (158)	66 Dy Dysprosium (160)	67 Ho Holmium (164)	68 Er Erbium (167)	69 Tm Thulium (169)	70 Yb Ytterbium (173)	71 Lu Lutetium (175)
89 Ac Actinium (227)	90 Th Thorium (232)	91 Pa Protactinium (231)	92 U Uranium (238)	93 Np Neptunium (237)	94 Pu Plutonium (239)	95 Am Americium (243)	96 Cm Curium (247)	97 Bk Berkelium (247)	98 Cf Californium (251)	99 Es Escoinium (251)	100 Fm Fermium (252)	101 Md Mendelevium (253)	102 No Nobelium (254)	103 Lr Lawrencium (254)	104 Hs Hassium (254)	105 Mt Meitnerium (254)	106 Ts Tennessine (254)	107 Og Oganesson (254)														

### 10.1 Alkali Metals:

#### 10.1.1 Hydrogen (H):

Hydrogen is the simplest and most fundamental element in the universe. It has atomic number 1 and in its most common isotope (protium) it consists of a single proton and one electron. The proton is composed of two up quarks (u) and one down quark (d) (uud), bound together by gluons through the strong nuclear force. Unlike many other nuclei, hydrogen-1 contains no neutron, which makes it the most elementary example of baryonic matter. Its electron, belonging to the lepton family, is not made of quarks but is an indivisible elementary particle.

Chemically, hydrogen is highly reactive and often forms compounds with electronegative elements such as oxygen, nitrogen, and halogens. In stars, hydrogen undergoes nuclear fusion, where four protons fuse under extreme conditions to form helium nuclei. This fusion process is the foundation of stellar energy production and the origin of heavier elements in the universe.

From the perspective of the White Hole Existence Principle (WHEP), hydrogen plays a key role as the most abundant element in the cosmos. White holes are predicted to generate hyperelectric fields mediated by the HE-boson, which may interact with hydrogen nuclei at the quantum level. The proton's quark structure could temporarily shift under such conditions, possibly involving hypothetical WHE-quarks, leading to exotic emissions of both electromagnetic radiation and novel particle streams. This makes hydrogen not only the

basis of ordinary cosmic matter but also a critical testing ground for how matter behaves near white holes and under extreme hyperelectric influence.

In summary, hydrogen is both the building block of stars and galaxies and a cornerstone in WHEP, bridging ordinary matter and exotic quantum phenomena in white hole environments.

### 10.1.2 Lithium (Li)

Lithium, with atomic number 3, is the lightest of the alkali metals and one of the most essential elements for both chemistry and modern technology. Its nucleus in the most stable isotope, Lithium-7, consists of three protons and four neutrons. Each proton is composed of two up-quarks and one down-quark (uud), while each neutron consists of two down-quarks and one up-quark (udd). In total, the lithium-7 nucleus contains seventeen up-quarks and sixteen down-quarks, bound together by gluons via the strong interaction. The three electrons surrounding the nucleus are elementary leptons and are not made of quarks.

Chemically, lithium is highly reactive, especially with water, where it produces hydrogen gas and lithium hydroxide. Due to its low density, lithium can even float on water. Its position as the first true alkali metal gives it properties like sodium and potassium, but it is less reactive compared to its heavier counterparts. Lithium compounds are widely used in rechargeable batteries, ceramics, and in psychiatric medicine as a mood stabilizer.

From the WHEP perspective, lithium has a unique role. Being both light and stable, it is considered one of the "transition" elements between hydrogen/helium (formed in the Big Bang) and heavier elements created in stars. In white hole environments, lithium nuclei are hypothesized to undergo transformations under the influence of hyperelectric fields mediated by the HE-boson. The delicate balance between protons and neutrons in lithium makes it a possible candidate for quark-level resonance states, where the inclusion of a WHE-quark could momentarily stabilize exotic isotopes. Such states might explain anomalous radiation signatures near white holes that cannot be accounted for by hydrogen alone.

In summary, lithium bridges the gap between the simplicity of hydrogen and the complexity of heavier elements. In WHEP, it may act as an experimental cornerstone to study how ordinary baryonic matter transitions into exotic matter under extreme quantum-gravitational conditions.

### 10.1.3 Sodium (Na)

Sodium, with atomic number 11, is one of the most well-known alkali metals. In its most common isotope, Sodium-23, the nucleus contains eleven protons and twelve neutrons. Each proton is built of two up-quarks and one down-quark (uud), while each neutron is composed of two down-quarks and one up-quark (udd). This means the sodium-23 nucleus contains a total of sixty-seven up-quarks and sixty-eight down-quarks, all bound by gluons in the strong nuclear force. Surrounding the nucleus are eleven electrons, arranged in the configuration [Ne]3s<sup>1</sup>, with a single valence electron in the outermost shell. This lone electron is what makes sodium highly reactive.

Chemically, sodium reacts violently with water, producing hydrogen gas and sodium hydroxide, often with enough heat to ignite the hydrogen. Its salts, such as sodium chloride (NaCl), are essential to life and are the foundation of biological ionic balance. Sodium ions play a critical role in nerve impulse transmission and cellular processes, making sodium indispensable for biology. Despite its intense reactivity as a pure element, in compounds it is both stable and life supporting.

In the WHEP framework, sodium plays a significant role as a mid-level element between the exceptionally light alkali metals and the much heavier ones. The relatively balanced number of protons and neutrons in sodium-23 provides a testing ground for studying HE-boson field interactions. The hyperelectric influence of white holes is theorized to distort the proton-neutron ratio in certain isotopes, possibly leading to transient states where WHE-quarks integrate into nuclear matter. These exotic sodium isotopes could contribute to unexplained emission lines observed near white hole candidates.

From an astrophysical perspective, sodium is notable for its strong yellow doublet emission (the sodium D-lines), which is often used in spectroscopy to identify its presence in stars and nebulae. Under WHEP, these emissions may be further modified by white hole fields, producing broadened or shifted spectral lines that serve as indirect evidence of hyperelectric interactions.

In summary, sodium is both a cornerstone of chemistry and biology, while in WHEP it serves as a probe into the intermediate complexity of baryonic matter. It illustrates how the quark-level structure of nuclei may respond differently in extreme white hole environments compared to lighter or heavier alkali metals.

#### 10.1.4 Potassium (K)

Potassium, atomic number 19, is a soft, silvery alkali metal notable for its high reactivity. Its most common isotope, Potassium-39, contains nineteen protons and twenty neutrons. Each proton consists of two up-quarks and one down-quark (uud), while each neutron has two down-quarks and one up-quark (udd). The nucleus therefore contains fifty-seven up-quarks and sixty down-quarks, bound together by gluons through the strong nuclear force. Potassium has nineteen electrons, with a single valence electron in the 4s orbital, giving it typical alkali metal chemical behavior.

Chemically, potassium reacts vigorously with water, producing hydrogen gas and potassium hydroxide. This reactivity increases with atomic number in the alkali group. Potassium ions are crucial in biological systems, playing a significant role in nerve function, heart rhythm, and cellular fluid balance. In industry, potassium compounds such as potassium nitrate and potassium carbonate are used in fertilizers, glass production, and explosives.

From the WHEP perspective, potassium's larger nucleus and increased number of neutrons provide an ideal test case for studying nuclear interactions under hyperelectric fields generated by white holes. WHEP predicts that the HE-boson may induce subtle energy shifts in the nucleus, and under extreme conditions, transient WHE-quark states could appear. These effects could influence potassium's spectral lines or isotopic stability near white holes, providing observable markers for WHEP phenomena.

In astrophysics, potassium is detected through its characteristic spectral lines in stars and interstellar gas. In WHEP, these lines might exhibit shifts or broadenings when potassium nuclei are subjected to white hole hyperelectric forces, potentially serving as experimental evidence for exotic quark interactions.

In summary, potassium exemplifies a heavier alkali metal whose quark composition and chemical properties make it not only essential in life and industry but also a valuable probe for WHEP research, bridging ordinary matter and exotic white hole interactions.

### 10.1.5 Rubidium (Rb)

Rubidium, atomic number 37, is a soft, highly reactive alkali metal. Its most stable isotope, Rubidium-85, has thirty-seven protons and forty-eight neutrons. Protons (uud) and neutrons (udd) form a nucleus with 111 up-quarks and 132 down-quarks, held together by the formidable force via gluons. Rubidium's thirty-seven electrons include a single valence electron in the 5s orbital, responsible for its typical alkali reactivity.

Chemically, rubidium reacts explosively with water, forming rubidium hydroxide and hydrogen gas. Rubidium ions are used in atomic clocks and research on quantum phenomena.

In WHEP, rubidium is of interest because its heavier nucleus may interact more strongly with HE-boson fields, potentially creating temporary WHE-quark states in the nucleus. Such interactions could influence emission spectra near white holes, making rubidium a candidate for observational tests of WHEP predictions.

---

### 10.1.6 Cesium (Cs)

Cesium, atomic number 55, is a soft, gold-colored alkali metal with extreme reactivity. The common isotope, Cesium-133, contains fifty-five protons and seventy-eight neutrons, totaling 169 up-quarks and 156 down-quarks in its nucleus. The single 6s valence electron makes cesium highly reactive.

Cesium is widely known for its use in atomic clocks, which define the standard for time measurement. Chemically, it reacts violently with water and is stored under inert liquids to prevent explosions.

Within WHEP, cesium's heavy nucleus provides a platform to study how hyperelectric fields influence large nuclei. Cesium could temporarily form exotic isotopes with WHE-quark contributions, potentially detectable through altered spectral lines or particle emissions in white hole environments.

---

### 10.1.7 Francium (Fr)

Francium, atomic number 87, is extremely rare and highly radioactive. Its most stable isotope, Francium-223, has eighty-seven protons and 136 neutrons, giving a nucleus composed of 311 up-quarks and 259 down-quarks. Like other alkali metals, it has one valence electron in the 7s orbital, making it chemically reactive but highly unstable.

Francium's natural scarcity and short half-life limit its practical applications, though it has been studied for nuclear research.

In WHEP, francium is particularly interesting because its large, neutron-rich nucleus could exhibit pronounced effects under HE-boson fields, possibly producing temporary WHE-quark excitations. Observations of francium near white hole analogues could reveal insights into quark-level nuclear behavior under extreme hyperelectric influences.

---

## 10.2 Alkaline Earth Metals:

### 10.2.1 Beryllium (Be)

Beryllium, atomic number 4, is a lightweight alkaline earth metal with a high melting point. Its most abundant isotope, Beryllium-9, contains four protons and five neutrons, giving a total of thirteen up-quarks and fourteen down-quarks in the nucleus. Electrons are arranged as  $1s^22s^2$ , with two valence electrons in the 2s orbital.

Chemically, beryllium is relatively stable compared to other light metals but forms toxic compounds such as beryllium oxide. It is used in aerospace components, X-ray windows, and nuclear reactors.

In WHEP, beryllium's small but stable nucleus makes it ideal for studying HE-boson effects on light nuclei. Its quark structure allows potential formation of WHE-quark excitations, which could slightly alter nuclear energy levels under hyperelectric influence near white holes.

---

### 10.2.2 Magnesium (Mg)

Magnesium, atomic number 12, is an essential alkaline earth metal. The common isotope Magnesium-24 has twelve protons and twelve neutrons, totaling thirty-six up-quarks and thirty-six down-quarks. Electrons occupy  $[Ne]3s^2$ , giving two valence electrons that drive its chemical behavior.

Magnesium reacts moderately with water and readily with acids. It is biologically crucial for photosynthesis, enzyme function, and bone structure, and is used industrially in alloys and fireworks.

In WHEP, magnesium's nucleus offers a mid-sized testing ground for HE-boson interactions. Hyperelectric fields may induce temporary WHE-quark states, affecting nuclear stability or energy emission, providing measurable markers for white hole-induced exotic matter effects.

---

### 10.2.3 Calcium (Ca)

Calcium, atomic number 20, is a biologically essential alkaline earth metal. Its main isotope, Calcium-40, has twenty protons and twenty neutrons, yielding sixty up-quarks and sixty down-quarks. Electrons occupy  $[Ar]4s^2$ , with two valence electrons influencing chemical reactivity.

Calcium is critical for bone and teeth formation, muscle function, and cellular signaling. Industrially, calcium compounds are used in cement, metallurgy, and chemical processes.

From a WHEP perspective, calcium's nucleus could interact more noticeably with HE-boson fields in white hole environments. The nucleus may briefly enter states with WHE-quark participation, which might subtly alter isotopic energy levels and emission spectra, serving as observational indicators for hyperelectric nuclear effects.

### 10.2.4 Strontium (Sr)

Strontium, atomic number 38, is a soft, silvery alkaline earth metal. Its most stable isotope, Strontium-88, has thirty-eight protons and fifty neutrons, yielding a nucleus composed of 114

up-quarks and 138 down-quarks, bound by gluons. Electrons occupy [Kr]5s<sup>2</sup>, with two valence electrons contributing to its reactivity.

Chemically, strontium reacts moderately with water and oxygen, forming oxides and hydroxides. Its salts are widely used in fireworks to produce bright red colors, and in ferrite magnets. Biologically, strontium can replace calcium in bones in trace amounts, and radioactive isotopes (like Sr-90) are significant in nuclear science.

In WHEP, strontium's larger nucleus provides a medium-heavy system to study HE-boson effects and potential WHE-quark states. Hyperelectric fields from white holes could induce temporary distortions in the proton-neutron arrangement, influencing isotopic stability and producing exotic emission patterns detectable via spectroscopy.

---

### 10.2.5 Barium (Ba)

Barium, atomic number 56, is a soft, reactive alkaline earth metal. Its dominant isotope, Barium-138, has fifty-six protons and eighty-two neutrons, totaling 170 up-quarks and 164 down-quarks in the nucleus. Electrons occupy [Xe]6s<sup>2</sup>, with two outer electrons responsible for its chemical reactivity.

Barium reacts with water and oxygen, forming oxides and hydroxides. Its compounds are widely used in drilling fluids, fireworks, and medical imaging (barium sulfate). Biologically, barium is toxic in soluble forms but inert in barium sulfate.

From the WHEP perspective, barium's heavy nucleus makes it a prime candidate for studying how HE-boson hyperelectric fields influence multi-quark nuclear systems. Temporary formation of WHE-quark excitations may affect nuclear energy levels, isotopic stability, and potential particle emission near white hole environments.

---

### 10.2.6 Radium (Ra)

Radium, atomic number 88, is a highly radioactive alkaline earth metal. Its most studied isotope, Radium-226, has eighty-eight protons and 138 neutrons, giving a nucleus with 262 up-quarks and 226 down-quarks. Electrons are arranged as [Rn]7s<sup>2</sup>, with two valence electrons dictating chemical behavior, though its extreme radioactivity limits direct handling.

Chemically, radium behaves similarly to barium, forming oxides and hydroxides. Its radioactivity historically led to applications in luminescent paints and cancer treatment.

In WHEP, radium's large, neutron-rich nucleus is ideal for investigating hyperelectric influences on heavy nuclei. HE-boson fields may induce transient WHE-quark states, temporarily modifying nuclear configurations, decay pathways, or emission signatures. Observations of radium in white hole analogues could provide experimental evidence of exotic quark dynamics and nuclear responses under extreme fields.

## 10.3 Transition Metals:

### 10.3.1 Scandium (Sc)

Scandium, atomic number 21, is the lightest transition metal and exhibits characteristics of both group 3 elements and early transition metals. Its most stable isotope, Scandium-45, contains twenty-one protons and twenty-four neutrons, resulting in a nucleus with sixty-three up-quarks and sixty-six down-quarks, bound via the strong interaction mediated by gluons. This composition gives scandium a moderately small, compact nucleus that is overly sensitive to perturbations in hyperelectric fields in WHEP environments.

Electronically, scandium's configuration is [Ar]3d<sup>1</sup>4s<sup>2</sup>. The single electron in the 3d orbital combined with two 4s electrons provides scandium with three valence electrons, which are involved in metallic bonding and chemical reactions. Scandium forms a variety of oxides and halides, often displaying +3 oxidation state. In industrial applications, scandium is incorporated in aerospace alloys, where lesser amounts improve strength and corrosion resistance.

Within WHEP, scandium nuclei are used to probe HE-boson field effects on lighter transition metals. The HE-boson core can induce transient WHE-quark excitations, altering local energy levels of the nucleus. This interaction can also modify effective electromagnetic coupling for the valence electrons, leading to measurable deviations in spectra when exposed to strong hyperelectric fields. The combination of a small nucleus, accessible valence electrons, and sensitivity to HE-core perturbations makes scandium a valuable model for studying exotic quantum effects in WHEP, including spin-dependent phenomena and potential emission of exotic particles under HE-boson influence.

In summary, scandium provides a compact, experimentally relevant platform within WHEP to explore quark-level responses to HE-boson cores, changes in electromagnetic interactions, and the emergent effects on chemical reactivity and nuclear stability under extreme hyperelectric conditions.

### 10.3.2 Titanium (Ti)

Titanium, atomic number 22, isotope Ti-48, has twenty-two protons and twenty-six neutrons, for a total of sixty-six up-quarks and seventy down-quarks. Electron configuration [Ar]3d<sup>2</sup>4s<sup>2</sup> gives four valence electrons, contributing to its strong metallic bonds. Titanium is corrosion-resistant, biologically inert, and widely used in aerospace, medical implants, and high-strength alloys.

In WHEP, titanium nuclei interact with HE-boson cores such that spin-dependent hyperelectric forces may slightly modify nuclear energy levels. These interactions can alter the electromagnetic coupling for its valence electrons, producing subtle spectral shifts that are measurable with WHEP-specific instrumentation. Titanium's mid-sized nucleus and multiple valence electrons make it ideal for exploring quark-spin dynamics under HE-boson influence, as well as the onset of emergent phenomena in heavy-element transitions.

### 10.3.3 Vanadium (V)

Vanadium, atomic number 23, isotope V-51, has twenty-three protons and twenty-eight neutrons, totaling sixty-nine up-quarks and seventy-four down-quarks. Electron configuration

$[Ar]3d^34s^2$  yields five valence electrons, facilitating strong metallic bonding and versatile oxidation states (+2, +3, +4, +5). Vanadium is used in steel alloys, catalysts, and chemical reagents.

WHEP predicts that vanadium's nucleus, under HE-boson core influence, may exhibit temporary WHE-quark excitations. These can affect local nuclear energy levels and slightly modify electromagnetic interactions of the valence electrons. Such effects can create unique spectral lines, providing experimental signatures of hyperelectric field interactions.

Vanadium's nucleus is large enough to amplify HE-boson effects, yet small enough to maintain calculable quantum dynamics, making it a key system in WHEP studies.

### 10.3.4 Chromium (Cr)

Chromium, atomic number 24, isotope Cr-52, contains twenty-four protons and twenty-eight neutrons, totaling seventy-two up-quarks and seventy-six down-quarks. Electron configuration  $[Ar]3d^54s^1$  gives six valence electrons, with a half-filled 3d shell that contributes to stability and strong metallic bonding. Chromium forms oxides and alloys are highly corrosion-resistant and are used in stainless steel and plating industries.

Within WHEP, chromium nuclei are sensitive to HE-boson core hyperelectric interactions. The half-filled d-shell allows measurable spin alignment shifts in the WHE-quarks, producing subtle changes in nuclear energy levels. Chromium's electromagnetic coupling in HE-core environments may deviate slightly from classical predictions, offering insights into quark-level interactions and exotic particle emissions mediated by HE-bosons.

### 10.3.5 Manganese (Mn)

Manganese, atomic number 25, isotope Mn-55, has twenty-five protons and thirty neutrons, with seventy-five up-quarks and eighty-one down-quarks. Its electron configuration  $[Ar]3d^54s^2$  results in seven valence electrons, enabling a wide range of oxidation states (+2 to +7). Manganese is used in steel alloys, batteries, and biological enzymes.

In WHEP, manganese nuclei are predicted to exhibit enhanced HE-boson coupling, with quark spins interacting strongly with hyperelectric fields. These interactions can subtly shift nuclear energy levels, modify valence electron electromagnetic behavior, and generate observable WHEP-specific spectral deviations. The combination of a moderately large nucleus and multiple valence electrons makes manganese a prime candidate for studying spin-dependent HE-boson effects.

### 10.3.6 Iron (Fe)

Iron, atomic number 26, isotope Fe-56, contains twenty-six protons and thirty neutrons, with seventy-eight up-quarks and eighty-one down-quarks. Electron configuration  $[Ar]3d^64s^2$  provides eight valence electrons, allowing complex magnetic and chemical behaviors. Iron is fundamental in planetary cores, biological systems (hemoglobin), and industrial alloys.

In WHEP, iron's mid-sized nucleus experiences significant interactions with the HE-boson core, affecting both quark spins and hyperelectric coupling. These interactions can modify electromagnetic properties of valence electrons and induce subtle energy shifts in the nuclear spectrum. Iron serves as a central system for studying WHE-quark spin dynamics,

nuclear stability, and emergent phenomena under HE-boson influence, especially given its natural magnetic moment and multiple oxidation states.

### 10.3.7 Cobalt (Co)

Cobalt, atomic number 27, isotope Co-fifty-nine, contains twenty-seven protons and thirty-two neutrons, giving eighty-one up-quarks and eighty-seven down-quarks in the nucleus. Its electron configuration  $[Ar]3d^74s^2$  results in nine valence electrons. Cobalt is ferromagnetic and widely used in superalloys, batteries, and catalysts.

Within WHEP, cobalt nuclei exhibit pronounced HE-boson-induced hyperelectric interactions. The partially filled 3d shell allows for measurable quark spin alignments that can alter nuclear energy states. These interactions modify the effective electromagnetic coupling of the valence electrons, producing observable deviations in spectra and spin-dependent behaviors. Cobalt thus serves as a key system for studying magnetism under HE-boson influence and potential exotic particle emissions.

---

### 10.3.8 Nickel (Ni)

Nickel, atomic number 28, isotope Ni-58, has twenty-eight protons and thirty neutrons, totaling eighty-four up-quarks and ninety down-quarks. Electron configuration  $[Ar]3d^84s^2$  yields ten valence electrons. Nickel is corrosion-resistant, used in alloys, coins, and batteries, and is mildly magnetic.

In WHEP, nickel nuclei under HE-boson-core influence experience quark-spin interactions that can subtly modify nuclear energy levels. The valence electron cloud also shows minor shifts in electromagnetic behavior due to HE-boson hyperelectric coupling. These effects provide measurable signals for testing spin-dependent HE-core dynamics and investigating WHE-quark excitations in medium-sized transition metals.

---

### 10.3.9 Copper (Cu)

Copper, atomic number 29, isotope Cu-63, contains twenty-nine protons and thirty-four neutrons, resulting in eighty-seven up-quarks and ninety-five down-quarks. Its electron configuration  $[Ar]3d^{10}4s^1$  produces one valence electron for high electrical conductivity. Copper is widely used in electronics, plumbing, and alloys.

Within WHEP, copper's nearly full 3d shell and single 4s valence electron makes it an ideal probe of HE-boson-mediated hyperelectric effects. Quark spins in the nucleus can transiently be coupled to the HE-boson field, slightly modifying the electromagnetic interactions of the valence electron. Copper thus demonstrates how HE-core influence on quanta can produce subtle, measurable effects on conductivity, spectral lines, and nuclear stability.

### 10.3.10 Zinc (Zn)

Zinc, atomic number 30, isotope Zn-64, contains thirty protons and thirty-four neutrons, totaling ninety up-quarks and ninety-four down-quarks. Its electron configuration  $[Ar]3d^{10}4s^2$  provides two valence electrons. Zinc is chemically versatile, forming salts, oxides, and alloys, and is essential in biological enzymes.

In WHEP, zinc nuclei are sensitive to HE-boson hyperelectric fields, though the fully filled 3d shell reduces spin-dependent effects compared to partially filled d-shell metals. The valence electrons may still experience subtle shifts in effective electromagnetic coupling under strong HE-boson influence. Zinc serves as a baseline system for studying WHE-core interactions in relatively stable, closed-shell nuclei, providing reference data for exotic particle excitations and spin-alignment dynamics.

---

### 10.3.11 Yttrium (Y)

Yttrium, atomic number 39, isotope Y-89, has thirty-nine protons and fifty neutrons, resulting in 117 up-quarks and 131 down-quarks. Electron configuration  $[Kr]4d^15s^2$  yields three valence electrons. Yttrium is used in superconductors, LEDs, and advanced alloys.

Within WHEP, yttrium's moderately heavy nucleus allows enhanced HE-boson core interactions. The 4d electron and three valence electrons experience modifications in electromagnetic coupling due to hyperelectric spin-dependent forces. Yttrium nuclei can also exhibit measurable WHE-quark excitations, with energy shifts detectable in WHEP spectroscopic experiments. Its role in superconducting and magnetic systems makes it ideal for studying collective HE-boson effects on electron behavior.

---

### 10.3.12 Zirconium (Zr)

Zirconium, atomic number 40, isotope Zr-90, contains forty protons and fifty neutrons, totaling 120 up-quarks and 130 down-quarks. Electron configuration  $[Kr]4d^25s^2$  gives four valence electrons. Zirconium is corrosion-resistant, used in nuclear reactors, alloys, and ceramics.

In WHEP, zirconium's nucleus is a prime example of medium-heavy transition metals interacting with HE-boson cores. Quark spins in the nucleus couple to the hyperelectric field, slightly modifying energy levels and effective EM interactions for valence electrons. Zirconium can thus serve as a model system for studying spin-dependent HE-boson effects, nuclear stability, and the emergence of WHE-quark excitations under strong hyperelectric influence.

### 10.3.13 Niobium (Nb)

Niobium, atomic number 41, isotope Nb-93, contains forty-one protons and fifty-two neutrons, totaling 123 up-quarks and 134 down-quarks. Electron configuration  $[Kr]4d^45s^1$  gives five valence electrons. Niobium is used in superconducting materials, aerospace alloys, and electronics.

Within WHEP, niobium nuclei interact with the HE-boson core, leading to subtle modifications of nuclear energy levels. The partially filled 4d shell enhances spin-dependent quark effects, which slightly influence the electromagnetic behavior of valence electrons. Niobium provides a robust system for studying WHE-quark excitations and the effect of hyperelectric fields on medium-heavy nuclei.

---

### 10.3.14 Molybdenum (Mo)

Molybdenum, atomic number 42, isotope Mo-98, has forty-two protons and fifty-six neutrons, giving 126 up-quarks and 140 down-quarks. Electron configuration  $[Kr]4d^55s^1$  yields six valence electrons. Molybdenum is essential in steel alloys, catalysts, and enzymes.

In WHEP, molybdenum's half-filled 4d shell makes it particularly sensitive to HE-boson hyperelectric interactions. Quark spins can temporarily be coupled to HE-core fields, producing measurable energy shifts in both the nucleus and valence electron orbitals. These effects provide experimental access to spin-dependent phenomena and exotic particle emission under WHE-boson influence.

---

### 10.3.15 Technetium (Tc)

Technetium, atomic number 43, isotope Tc-98, contains forty-three protons and fifty-five neutrons, totaling 129 up-quarks and 140 down-quarks. Electron configuration  $[Kr]4d^55s^2$  gives seven valence electrons. Technetium is radioactive, with no stable isotopes, used in medical imaging and nuclear research.

Within WHEP, technetium nuclei are highly influenced by HE-boson cores due to their instability and unpaired quark spins. Hyperelectric interactions can transiently modify nuclear energy levels, valence electron electromagnetic coupling, and may trigger WHE-quark excitations, producing detectable spectral deviations. Technetium provides a unique system to study radiative decay under HE-boson-induced spin effects, combining nuclear instability with hyperelectric coupling phenomena.

### 10.3.16 Ruthenium (Ru)

Ruthenium, atomic number 44, isotope Ru-102, contains forty-four protons and fifty-eight neutrons, resulting in 132 up-quarks and 140 down-quarks. Electron configuration  $[Kr]4d^75s^1$  yields eight valence electrons. Ruthenium is used in electronics, catalysts, and platinum alloys.

In WHEP, ruthenium nuclei interact with the HE-boson core, producing subtle energy level shifts through spin-dependent quark interactions. Hyperelectric fields can slightly alter electromagnetic behavior of valence electrons, enabling WHEP-specific observations of WHE-quark excitations. Ruthenium's nucleus is heavy enough to amplify HE-boson effects while remaining experimentally tractable for mid-range transition metal studies.

---

### 10.3.17 Rhodium (Rh)

Rhodium, atomic number 45, isotope Rh-103, has forty-five protons and fifty-eight neutrons, totaling 135 up-quarks and 140 down-quarks. Electron configuration  $[Kr]4d^85s^1$  provides nine valence electrons. Rhodium is highly corrosion-resistant, used in catalytic converters and specialty alloys.

Within WHEP, rhodium's nearly full 4d shell allows precision probing of HE-boson hyperelectric interactions. Quark spins in the nucleus can be coupled with HE-core fields, producing slight shifts in energy levels and affecting the valence electron cloud. Rhodium's

stability and electron configuration make it ideal for studying spin-mediated nuclear-EM coupling and exotic WHE-quark dynamics in heavy nuclei.

---

### 10.3.18 Palladium (Pd)

Palladium, atomic number 46, isotope Pd-106, contains forty-six protons and sixty neutrons, giving 138 up-quarks and 150 down-quarks. Electron configuration  $[Kr]4d^{10}5s^0$  yields ten fully occupied valence electrons in the 4d shell. Palladium is widely used in hydrogen storage, catalysis, and electronics.

In WHEP, palladium's filled 4d shell reduces quark-spin sensitivity compared to partially filled d-shell metals, but HE-boson interactions still produce measurable effects on nuclear energy levels. The valence electrons may experience subtle hyperelectric coupling modifications, providing reference data for studies of HE-core induced spin alignment and emergent particle phenomena in closed-shell heavy nuclei.

### 10.3.19 Silver (Ag)

Silver, atomic number 47, isotope Ag-107, contains forty-seven protons and sixty neutrons, totaling 141 up-quarks and 150 down-quarks. Electron configuration  $[Kr]4d^{10}5s^1$  yields one valence electron, giving silver exceptional electrical conductivity and chemical reactivity in ionic form.

Within WHEP, silver nuclei interact subtly with HE-boson cores, producing minor shifts in nuclear energy levels through quark-spin alignment. The single valence electron is particularly sensitive to hyperelectric field modifications, making silver ideal for experiments measuring WHE-quark excitation effects on electron behavior. Its combination of a heavy nucleus and simple valence shell allows precise monitoring of HE-boson-induced phenomena.

---

### 10.3.20 Cadmium (Cd)

Cadmium, atomic number 48, isotope Cd-112, has forty-eight protons and sixty-four neutrons, giving 144 up-quarks and 160 down-quarks. Electron configuration  $[Kr]4d^{10}5s^2$  provides two valence electrons. Cadmium is used in batteries, coatings, and pigments.

In WHEP, cadmium's fully filled 4d shell reduces nuclear quark-spin sensitivity, but the valence electrons remain weakly affected by HE-boson hyperelectric fields. Nuclear energy levels experience slight shifts under HE-core influence, allowing cadmium to serve as a stable reference system for WHEP studies of spin-alignment dynamics and subtle electromagnetic deviations.

---

### 10.3.21 Hafnium (Hf)

Hafnium, atomic number 72, isotope Hf-178, contains seventy-two protons and 106 neutrons, totaling 216 up-quarks and 212 down-quarks. Electron configuration  $[Xe]4f^{14}5d^26s^2$  gives four valence electrons. Hafnium is highly corrosion-resistant and used in nuclear reactors, superalloys, and electronics.

Within WHEP, hafnium's heavy nucleus strongly interacts with HE-boson cores, producing spin-dependent quark excitations and subtle energy shifts in the valence electrons. These hyperelectric interactions can slightly modify electromagnetic coupling, providing experimental insights into HE-core effects in heavy transition metals and the emergence of WHE-quark dynamics in large nuclei. Hafnium's unique combination of high nuclear mass and valence electrons makes it a prime candidate for advanced WHEP investigations.

### 10.3.22 Tantalum (Ta)

Tantalum, atomic number 73, isotope Ta-181, contains seventy-three protons and 108 neutrons, totaling 219 up-quarks and 216 down-quarks. Electron configuration  $[Xe]4f^15d^36s^2$  yields five valence electrons. Tantalum is highly corrosion-resistant and widely used in electronics, surgical implants, and superalloys.

Within WHEP, tantalum's heavy nucleus exhibits strong HE-boson core interactions. Quark spins couple with hyperelectric fields, slightly modifying nuclear energy levels. Valence electrons also experience measurable adjustments in electromagnetic coupling. Tantalum's stability and mid-heavy nucleus make it a suitable candidate for exploring HE-boson-induced WHE-quark excitations and spin-dependent nuclear phenomena in heavy elements.

---

### 10.3.23 Tungsten (W)

Tungsten, atomic number 74, isotope W-184, contains seventy-four protons and 110 neutrons, resulting in 222 up-quarks and 220 down-quarks. Electron configuration  $[Xe]4f^15d^66s^2$  gives six valence electrons. Tungsten is known for its high melting point, density, and use in alloys, filaments, and industrial tools.

In WHEP, tungsten's heavy nucleus is overly sensitive to HE-boson hyperelectric effects, producing spin-aligned quark excitations that subtly influence valence electron energy levels. These interactions provide experimental observations of nuclear-HE is coupling and emergent phenomena in heavy transition metals. Tungsten is ideal for WHEP studies requiring both nuclear stability and strong HE-boson responses.

---

### 10.3.24 Rhenium (Re)

Rhenium, atomic number 75, isotope Re-187, contains seventy-five protons and 112 neutrons, totaling 225 up-quarks and 224 down-quarks. Electron configuration  $[Xe]4f^15d^56s^2$  provides seven valence electrons. Rhenium is used in superalloys, jet engines, and electrical contacts.

Within WHEP, rhenium's heavy nucleus interacts with HE-boson cores, inducing minor shifts in nuclear energy levels via quark-spin coupling. Valence electrons experience slight electromagnetic deviations due to hyperelectric field effects. Rhenium's combination of high nuclear mass and multiple valence electrons makes it a prime system for studying WHE-quark dynamics, HE-core effects, and emergent spin-dependent phenomena in heavy transition metals.

---

### 10.3.25 Osmium (Os)

Osmium, atomic number 76, isotope Os-192, contains seventy-six protons and 116 neutrons, totaling 228 up-quarks and 232 down-quarks. Electron configuration  $[Xe]4f^{14}5d^66s^2$  yields eight valence electrons. Osmium is extremely dense, hard, and used in specialized alloys and electrical contacts.

Within WHEP, osmium's heavy nucleus is strongly influenced by HE-boson core hyperelectric fields. Quark spins in the nucleus can be coupled transiently with HE-core interactions, subtly shifting nuclear energy levels. Valence electrons experience minor electromagnetic modifications, making osmium a prime candidate for observing spin-dependent WHE-quark excitations and testing HE-core effects in ultra-dense transition metals.

---

### 10.3.26 Iridium (Ir)

Iridium, atomic number 77, isotope Ir-193, contains seventy-seven protons and 116 neutrons, giving 231 up-quarks and 232 down-quarks. Electron configuration  $[Xe]4f^{14}5d^76s^2$  provides nine valence electrons. Iridium is highly corrosion-resistant, dense, and used in crucibles, electrical contacts, and aerospace alloys.

In WHEP, iridium nuclei interact significantly with HE-boson cores, producing measurable spin-dependent shifts in nuclear energy levels. Valence electrons are slightly influenced by hyperelectric fields, altering electromagnetic coupling. Iridium serves as a model system for nuclear spin alignment studies and the detection of WHE-quark excitations in dense heavy elements.

---

### 10.3.27 Platinum (Pt)

Platinum, atomic number 78, isotope Pt-195, contains seventy-eight protons and 117 neutrons, totaling 234 up-quarks and 234 down-quarks. Electron configuration  $[Xe]4f^{14}5d^96s^1$  yields ten valence electrons. Platinum is corrosion-resistant, catalytically active, and widely used in jewelry, electronics, and chemical catalysts.

Within WHEP, platinum nuclei experience HE-boson hyperelectric effects that subtly modify nuclear energy levels via quark-spin interactions. Valence electron behavior is affected by altered electromagnetic coupling, enabling experimental observation of WHE-quark excitations and spin-dependent HE-core phenomena. Platinum's combination of high nuclear mass and filled d-shell valence electrons makes it ideal for advanced WHEP studies.

### 10.3.28 Gold (Au)

Gold, atomic number 79, isotope Au-197, contains seventy-nine protons and 118 neutrons, totaling 237 up-quarks and 236 down-quarks. Electron configuration  $[Xe]4f^{14}5d^{10}6s^1$  provides one valence electron. Gold is highly conductive, corrosion-resistant, and widely used in electronics, jewelry, and monetary systems.

Within WHEP, gold nuclei experience HE-boson core interactions, producing minor quark-spin-induced shifts in nuclear energy levels. The single valence electron is particularly sensitive to hyperelectric field modifications, making gold ideal for observing WHE-quark excitations and spin-dependent electromagnetic phenomena in heavy, stable nuclei.

### 10.3.29 Mercury (Hg)

Mercury, atomic number 80, isotope Hg-202, contains eighty protons and 122 neutrons, resulting in 240 up-quarks and 242 down-quarks. Electron configuration  $[Xe]4f^15d^16s^2$  yields two valence electrons. Mercury is liquid at room temperature, highly dense, and used in thermometers, switches, and scientific instruments.

In WHEP, mercury's filled d-shell and liquid-phase behavior create unique HE-boson hyperelectric interactions. Nuclear quark spins can transiently couple to the HE-core, producing slight shifts in energy levels. Valence electrons exhibit minor electromagnetic deviations, offering experimental opportunities to study WHE-quark dynamics in liquid heavy metals under HE-core influence.

---

### 10.3.30 Rutherfordium (Rf)

Rutherfordium, atomic number 104, isotope Rf-267, contains 104 protons and 163 neutrons, totaling 312 up-quarks and 326 down-quarks. Electron configuration  $[Rn]5f^16d^27s^2$  provides four valence electrons. Rutherfordium is synthetic, radioactive, and primarily studied in nuclear physics research.

Within WHEP, rutherfordium's very heavy nucleus strongly interacts with HE-boson cores, producing measurable spin-dependent quark excitations and hyperelectric field effects. Valence electrons experience subtle electromagnetic coupling shifts. Rutherfordium serves as an extreme system for observing WHE-core phenomena, exotic particle emissions, and energy-level modifications in superheavy elements.

---

### 10.3.31 Dubnium (Db)

Dubnium, atomic number 105, isotope Db-268, contains 105 protons and 163 neutrons, totaling 315 up-quarks and 326 down-quarks. Electron configuration  $[Rn]5f^16d^37s^2$  gives five valence electrons. Dubnium is synthetic, highly unstable, and primarily produced for nuclear research.

In WHEP, dubnium's superheavy nucleus exhibits strong HE-boson core interactions, resulting in pronounced quark-spin alignment effects. Nuclear energy levels are slightly shifted by hyperelectric interactions, while valence electrons experience modified electromagnetic coupling. Dubnium provides a rare system to study WHE-quark excitations and extreme HE-core phenomena in superheavy transition metals.

---

### 10.3.32 Seaborgium (Sg)

Seaborgium, atomic number 106, isotope Sg-271, contains 106 protons and 165 neutrons, totaling 318 up-quarks and 330 down-quarks. Electron configuration  $[Rn]5f^16d^47s^2$  yields six valence electrons. Seaborgium is synthetic and highly radioactive, observed only in trace quantities.

Within WHEP, seaborgium nuclei interact strongly with HE-boson cores, inducing measurable spin-dependent nuclear shifts. Valence electrons are subtly affected by hyperelectric coupling. Seaborgium serves as a model for WHEP studies of extreme nuclear mass and HE-core interactions, demonstrating how superheavy elements manifest quark-spin alignment and exotic particle phenomena.

---

### 10.3.33 Bohrium (Bh)

Bohrium, atomic number 107, isotope Bh-270, contains 107 protons and 163 neutrons, giving 321 up-quarks and 326 down-quarks. Electron configuration  $[Rn]5f^{14}6d^57s^2$  provides seven valence electrons. Bohrium is highly unstable, synthetic, and used for experimental nuclear physics studies.

In WHEP, bohrium nuclei exhibit strong HE-boson core coupling, producing spin-aligned quark excitations and minor nuclear energy shifts. Hyperelectric interactions affect valence electrons' electromagnetic behavior, enabling studies of WHE-quark excitations in extreme nuclear environments. Bohrium represents one of the heaviest systems where HE-core phenomena are significant and experimentally relevant.

---

### 10.3.34 Hassium (Hs)

Hassium, atomic number 108, isotope Hs-277, contains 108 protons and 169 neutrons, totaling 324 up-quarks and 338 down-quarks. Electron configuration  $[Rn]5f^{14}6d^67s^2$  yields eight valence electrons. Hassium is synthetic, highly radioactive, and studied for nuclear stability and chemical properties.

Within WHEP, hassium's superheavy nucleus strongly interacts with HE-boson cores, resulting in significant spin-dependent quark alignment. Nuclear energy levels are slightly modified, and valence electrons experience measurable hyperelectric coupling effects. Hassium serves as a benchmark for studying WHE-core dynamics in superheavy transition metals and the emergent behavior of quark interactions under extreme conditions.

---

## 10.4.1 Aluminum (Al)

Aluminum (chemical symbol Al) is a lightweight, silvery-white metal that belongs to the post-transition metals in group 13 of the periodic table. It is the most abundant metal in the Earth's crust, making up about 8% by weight, although it is never found in its pure form naturally. Instead, it is commonly extracted from the ore bauxite through the Bayer process and then refined using the Hall–Héroult process, which requires large amounts of electricity.

One of aluminum's most important properties is its low density combined with a relatively high strength, especially when alloyed with other elements such as copper, magnesium, and zinc. This makes it an essential material in industries like aerospace, construction, packaging, and transportation. For example, airplanes rely heav-

ily on aluminum alloys because of their excellent strength-to-weight ratio. Aluminum is also widely used in beverage cans, foil, and kitchen utensils due to its resistance to corrosion. The protective oxide layer that naturally forms on its surface prevents further oxidation, which makes it long-lasting in everyday applications.

Another major advantage of aluminum is that it is highly recyclable. Nearly 75% of all aluminum ever produced is still in use today, since recycling requires only a fraction of the energy needed to produce new aluminum from ore. Environmentally, this makes it one of the most sustainable industrial metals.

However, aluminum also has limitations. Pure aluminum is relatively soft and not as strong as steel. It also has a relatively low melting point (660 °C), which limits its use in high-temperature applications. Despite these drawbacks, its unique combination of abundance, lightness, corrosion resistance, and recyclability ensures that aluminum remains one of the most important metals in modern technology and industry.

#### 10.4.2 Gallium (Ga)

Gallium (chemical symbol Ga) is a soft, silvery post-transition metal found in group 13 of the periodic table. Unlike many metals, gallium is not found in its pure elemental form in nature. Instead, it is extracted as a by-product from the processing of aluminum and zinc ores. Its abundance in the Earth's crust is relatively low, but still sufficient for industrial use.

One of gallium's most fascinating properties is its low melting point of 29.7 °C (85 °F), which means it can melt in the palm of your hand. At the same time, it has a very high boiling point (over 2200 °C), giving it an extremely wide liquid temperature range. This unusual property makes it valuable for use in high-temperature thermometers and specialized cooling applications.

Gallium is best known for its role in modern electronics. It is a critical component in gallium arsenide (GaAs) and gallium nitride (GaN) semiconductors, which are used in light-emitting diodes (LEDs), laser diodes, solar cells, and high-frequency electronics. For example, LEDs in smartphones, traffic lights, and DVD players all rely on gallium-based compounds. GaN is especially important in power electronics and 5G communication technology due to its efficiency at handling high voltages and frequencies.

In its pure metallic form, gallium does not corrode in air or water, and it can wet glass and many other surfaces. While not considered highly toxic, gallium compounds must still be handled carefully in laboratory and industrial settings. Overall, gallium's unusual physical properties and critical role in electronics make it one of the most important "hidden" metals in modern technology.

### 10.4.3 Indium (In)

Indium (chemical symbol In) is a rare, soft, silvery-white metal in group 13 of the periodic table. It was discovered in 1863 and named after the indigo-blue spectral line observed during its identification. Indium is usually obtained as a by-product of zinc ore processing, and it is much rarer in Earth's crust compared to aluminum or tin.

Indium's most important property is its excellent ability to form transparent conductive coatings. The compound indium tin oxide (ITO) is widely used in touch screens, flat-panel displays, and solar cells, where it allows electricity to flow while remaining optically transparent. Without indium, modern devices such as smartphones, tablets, and LCD monitors would not function as they do today.

Another application of indium is in low-melting alloys. Indium can be combined with other metals to create alloys that melt at very low temperatures, which are useful in specialized solders, fuses, and thermal interface materials. Indium also adheres well to glass and other surfaces, making it useful in sealing and coating technologies.

Indium is relatively soft and malleable, and it does not oxidize easily in air, which increases its durability. It has a melting point of 156 °C, much higher than gallium but still lower than most metals. While indium itself is not highly toxic, some of its compounds must be handled with caution in industrial settings.

Due to its rarity, indium is considered a critical raw material. Demand is growing because of the increasing use of touch screens and renewable energy technologies, while supply is limited by zinc production. This makes indium one of the key “technology metals” of the 21st century.

### 10.4.4 Thallium (Tl)

Thallium (chemical symbol Tl) is a soft, heavy, bluish-gray metal belonging to group 13. It was discovered in 1861 by spectroscopy and named after the Greek word *thallos*, meaning “green shoot,” due to its bright green spectral line. Thallium occurs in small amounts in ores of copper, lead, and zinc, and is usually obtained as a by-product of their refining.

Unlike aluminum or indium, thallium is extremely toxic, and safety concerns have limited its use over time. In the past, thallium sulfate was used in rat poison and insecticides, but this has largely been banned due to environmental and health risks.

Today, thallium is mainly used in specialized applications. It is important in electronics, especially in semiconductors and photoelectric cells, where its compounds are sensitive to infrared radiation. Thallium bromide-iodide crystals are also used in infrared detectors, night vision systems, and specialized optical lenses. In medicine, small

amounts of radioactive thallium isotopes are used in nuclear imaging to detect heart disease.

Physically, thallium is soft enough to be cut with a knife, and it tarnishes quickly when exposed to air, forming a bluish-gray oxide layer. Its melting point is 304 °C, and it shares many chemical similarities with lead, which is directly below it in the periodic table.

Although its toxicity limits large-scale use, thallium remains scientifically significant due to its unique optical, electrical, and medical properties. Researchers continue to explore safe ways to use this rare and dangerous element in high-tech industries.

#### 10.4.5 Tin (Sn)

Tin (chemical symbol Sn) is a silvery-white metal in group 14, known since antiquity. It was historically important in the creation of bronze, an alloy of copper and tin that marked the beginning of the Bronze Age around 3000 BCE. Tin is relatively abundant in the Earth's crust and is primarily obtained from the ore cassiterite ( $\text{SnO}_2$ ).

One of tin's most useful properties is its resistance to corrosion, especially from water. This makes it valuable as a protective coating for other metals, most famously in tin-plated steel cans used for food storage. Tin is also widely used in solder, an alloy that joins metal components in electronics and plumbing. In fact, modern electronics would not function without tin-based solder connections.

Tin is relatively soft, malleable, and has a low melting point of 232 °C. It exists in different structural forms, the most notable being white tin (metallic) and gray tin (powdery, nonmetallic). At very low temperatures, white tin can transform into gray tin in a process called "tin pest," which can damage tin objects in cold climates.

In modern technology, tin compounds are also used in glass production, ceramics, and some chemical catalysts. While not considered highly toxic compared to lead or thallium, certain organotin compounds can be harmful and require regulation.

Because of its historical role in alloys and its continuing importance in electronics, tin remains a vital post-transition metal that bridges ancient metallurgy and modern technology.

#### 10.4.6 Lead (Pb)

Lead (chemical symbol Pb) is a heavy, dense, bluish-gray metal in group 14. It has been used by humans for thousands of years, dating back to ancient Rome, where it was employed in pipes, paints, and cosmetics. Lead is relatively easy to extract from

its main ore, galena (PbS), and has a low melting point of 327 °C, making it one of the most accessible metals in history.

Lead's most notable property is its density and ability to block radiation. This makes it invaluable for shielding against X-rays and nuclear radiation in medical and industrial settings. It is also used in lead-acid batteries, which are still common in cars and backup power systems.

However, lead is also highly toxic. Exposure to lead can cause severe health problems, particularly in the nervous system, making it especially dangerous for children. Because of this, lead has been phased out of many traditional uses, such as leaded gasoline, lead-based paints, and plumbing pipes.

Despite its dangers, lead remains important in specific industries. It is used in specialized alloys, radiation shielding, and some types of glass, such as lead crystal and protective glass in laboratories. Its softness and malleability allow it to be shaped easily, although it is not very strong compared to other metals.

Today, the focus is on carefully controlling and recycling lead to minimize its environmental and health impacts. Lead is a clear example of a metal with both great usefulness and serious risks, shaping human history while also presenting major challenges.

#### 10.4.7 Flerovium (Fl)

Flerovium (chemical symbol Fl) is a synthetic, superheavy element in group 14. It was first synthesized in 1998 at the Joint Institute for Nuclear Research (JINR) in Dubna, Russia, and was officially named in 2012 in honor of the Russian physicist Georgy Flyorov. Because it does not occur naturally, flerovium can only be produced in laboratories by bombarding plutonium or curium targets with high-energy ions such as calcium.

Flerovium is extremely unstable, with the most stable isotopes having half-lives of only a few seconds. This means it has no practical applications, and research is limited to studying its nuclear and chemical properties. Its rarity and instability make it one of the most difficult elements to investigate experimentally.

Theoretical predictions suggest that flerovium might behave similarly to lead, its lighter group 14 neighbor, but with unusual properties due to relativistic effects. Some models even propose that flerovium could be more volatile than expected, possibly behaving closer to a noble gas than a metal under certain conditions. However, because only a few atoms have ever been created, its exact properties remain uncertain.

Flerovium's importance lies in its contribution to the study of superheavy elements and the search for the so-called "island of stability," where longer-lived nuclei might exist. These studies help scientists better understand the limits of the periodic table and the forces that hold atomic nuclei together.

While flerovium has no industrial use, it represents the cutting edge of modern nuclear chemistry and highlights humanity's ability to push the boundaries of the known elements.

## 10.5 Metalloids:

### 10.5.1 Boron (B)

Boron is a metalloid element with atomic number 5 and is located in group 13 of the periodic table. It exhibits properties of both metals and nonmetals, making it essential in a variety of industrial and scientific applications. Pure boron is a hard, black, brittle crystalline solid at room temperature. One of its most notable characteristics is its semiconducting behavior, which makes it critical in modern electronic components. Boron has three valence electrons, allowing it to form covalent bonds with a variety of elements, especially oxygen, carbon, and nitrogen. This bonding versatility results in compounds such as borax (sodium borate), boric acid, and boron carbide, all of which have unique industrial uses.

Boron is naturally rare in its elemental form and is usually obtained from borate minerals. Its high melting point of approximately 2076 °C and low density make it useful in high-temperature applications. Boron fibers are used to reinforce materials in aerospace and defense industries due to their exceptional strength-to-weight ratio. In electronics, boron acts as a dopant in silicon-based semiconductors, enhancing electrical conductivity in precise ways, which is crucial for integrated circuits and transistors. Biologically, boron plays a minor but significant role in plant growth, particularly in cell wall formation and metabolic functions. Overall, boron's unique combination of chemical versatility, high thermal stability, and semiconducting properties make it a highly valuable metalloid in both technological and chemical industries.

### 10.5.2 Silicon (Si)

Silicon is a metalloid with atomic number 14, located in group 14 of the periodic table. It exhibits both metallic and nonmetallic properties, making it one of the most important elements in technology and industry. Pure silicon is a hard, brittle crystalline solid with a bluish-grey metallic luster. Its semiconducting properties are highly valued, as it allows precise control of electrical conductivity when doped with other elements, such as boron or phosphorus. Silicon has four valence electrons, which enables it to form strong covalent bonds, resulting in an extensive three-dimensional lattice in its crystalline form. This bonding is also responsible for the high melting point of approximately 1414 °C and its mechanical strength.

Silicon is the second most abundant element in the Earth's crust, primarily found in the form of silicate minerals such as quartz, feldspar, and mica. Industrially, it is used to manufacture glass, ceramics, concrete, and silicones. In electronics, silicon is the foundation of modern semiconductor technology, being the material for integrated circuits, transistors, diodes, and

solar cells. Its ability to form a stable oxide layer (silicon dioxide) is critical in microelectronics, acting as an insulator and protecting components. Additionally, silicon plays a minor biological role, particularly in plant structural integrity. The versatility, abundance, and unique semiconducting properties of silicon make it indispensable in both everyday materials and advanced technological applications.

### 10.5.3 Germanium (Ge)

Germanium is a metalloid element with atomic number 32, positioned in group 14 of the periodic table. It exhibits properties intermediate between metals and nonmetals, particularly notable for its semiconducting behavior. Germanium is a lustrous, hard, gray-white brittle solid that can form a diamond-like crystalline structure. Its four valence electrons allow it to form covalent bonds with other elements, giving it a tetrahedral crystal lattice similar to silicon. Germanium's semiconductor characteristics are essential in electronics, where it was historically used in early transistors and diodes before silicon largely replaced it. Despite being less abundant than silicon, germanium remains crucial for specialized electronic applications due to its higher electron mobility and effective performance at higher frequencies.

Naturally occurring germanium is usually obtained from sphalerite ores and other minerals in small quantities. It has a melting point of 938 °C and is chemically stable in air but can oxidize at high temperatures. Germanium forms compounds such as germanium dioxide ( $\text{GeO}_2$ ) and germanates, which have industrial uses in optics and electronics. In fiber optics and infrared optics, germanium's transparency to infrared radiation makes it valuable for lenses and windows. It also serves as a component in semiconductor alloys and as a dopant in silicon devices to modify electrical properties. Overall, germanium's combination of semiconducting ability, optical utility, and chemical stability secures its role in advanced technology, particularly in high-speed electronics and infrared optical systems.

### 10.5.4 Arsenic (As)

Arsenic is a metalloid with atomic number 33, located in group 15 of the periodic table. It exhibits properties of both metals and nonmetals, making it versatile in industrial and electronic applications. Pure arsenic exists in several allotropes, with gray arsenic being the most stable form at room temperature. Gray arsenic has a metallic luster, is brittle, and conducts electricity, displaying semiconducting behavior. Its five valence electrons allow it to form covalent bonds in compounds with metals and nonmetals, giving rise to a wide range of chemical species such as arsenides, arsenates, and organoarsenic compounds. The element's toxicity is well-known, and it has been historically used in pesticides, wood preservatives, and alloys, although usage is now heavily restricted due to health concerns.

In electronics, arsenic is an important dopant in semiconductors such as gallium arsenide (GaAs), which is used in high-speed and optoelectronic devices, including LEDs, laser diodes, and microwave-frequency integrated circuits. Arsenic compounds also play a role in alloying metals to improve hardness and corrosion resistance. Despite its toxicity, arsenic occurs naturally in the Earth's crust, primarily in minerals like realgar and orpiment. Industrially, arsenic is handled with extreme care due to its carcinogenic nature. Its combination of semiconducting properties, chemical versatility, and historical significance in materials science demonstrates why arsenic is an important metalloid in both technology and chemistry.

### 10.5.5 Antimony (Sb)

Antimony is a metalloid with atomic number 51, located in group 15 of the periodic table. It exhibits both metallic and nonmetallic properties, appearing as a silvery, lustrous, brittle solid at room temperature. Its five valence electrons enable it to form covalent and ionic compounds, contributing to its chemical versatility. Antimony has been known since ancient times and is used in alloys, flame retardants, and electronics. It has a relatively low melting point of 630 °C compared with other metalloids and is chemically stable in air, forming a thin oxide layer that protects it from further oxidation.

Industrial applications of antimony are diverse. When alloyed with lead, it improves hardness and mechanical strength, which is especially important in batteries, bullets, and bearings. Antimony trioxide is widely used as a flame retardant in plastics, textiles, and coatings, making it crucial in fire safety materials. In electronics, antimony acts as a dopant in semiconductors, particularly in diodes and infrared detectors. Naturally occurring antimony is found in the mineral stibnite ( $Sb_2S_3$ ), and although relatively rare, its unique combination of properties makes it valuable in metallurgical, chemical, and electronic industries. Antimony's role as a metalloid lies in its ability to bridge metallic conductivity and chemical reactivity, which makes it indispensable in certain technological applications.

### 10.5.6 Tellurium (Te)

Tellurium is a metalloid with atomic number 52, located in group 16 of the periodic table. It exhibits characteristics of both metals and nonmetals, appearing as a silvery-white, brittle crystalline solid. Tellurium has six valence electrons, allowing it to form covalent and some ionic compounds. Its semiconducting properties, combined with its chemical reactivity, make it useful in a variety of industrial and technological applications. Tellurium has a melting point of 449 °C and is chemically stable in air, forming a thin oxide layer when exposed to oxygen. It occurs naturally in the Earth's crust, primarily in combination with gold, silver, and other metals in tellurides.

Industrially, tellurium is widely used in alloys to improve hardness, machinability, and resistance to corrosion, particularly in lead and copper alloys. Its semiconducting properties are exploited in thermoelectric devices, which convert temperature differences into electrical energy, and in cadmium telluride (CdTe) solar cells, which are an important form of thin-film photovoltaic technology. Tellurium is also employed in the production of certain electronic components, including diodes and infrared detectors. Its combination of metalloid properties, electrical conductivity, and chemical versatility makes tellurium an important element in electronics, energy, and metallurgy. While relatively rare, tellurium's applications in advanced materials highlight its significance in modern industry and technology.

### 10.5.7 Polonium (Po)

Polonium is a metalloid with atomic number 84, located in group 16 of the periodic table. It is a rare and highly radioactive element, discovered by Marie and Pierre Curie in 1898. Polonium exhibits both metallic and nonmetallic properties, appearing as a silvery-gray, volatile, and brittle solid. Its radioactivity dominates its chemical behavior, as it undergoes alpha decay, producing significant heat and radiation. Polonium has six valence electrons, allowing it to form compounds with halogens, oxygen, and metals, although these compounds are less

stable due to its intense radioactivity. Its melting point is approximately 254 °C, and it is chemically reactive, particularly with oxygen and halogens, forming polonides and oxides.

Polonium occurs in trace amounts in uranium ores and is typically produced artificially by neutron irradiation of bismuth. Due to its extreme radioactivity, polonium has very limited applications, primarily in scientific research, as a heat source in space satellites, and in static eliminators used in industrial settings. Its highly toxic and radioactive nature means that it must be handled with specialized equipment and strict safety protocols. Polonium exemplifies the properties of a metalloid that straddles the boundary between metals and nonmetals, while its radioactivity introduces unique chemical and physical challenges. Despite its rarity and dangers, polonium has provided significant insight into nuclear physics and the behavior of heavy metalloids.

# The Theory of Everything

## Foundations, Applications and Corrections to General Relativity

Michael Scott Peck

Original: July 24<sup>th</sup>, 2012 | Final: May 20<sup>th</sup>, 2013 | Copyright 2013 | [Contact](#)

---

**Abstract:** Corrections to general relativity are derived from classical theory and applied to the standard model. The perspective offered is the conceptual inverse of Einstein's theory, where particles exist as localized fields. These vacuum fields undergo affine transformations that are locally invariant with respect to the space-time metric. It is demonstrated that the proper vacuum solution to the Einstein-Maxwell field equations is the limit of the single particle vacuum field solution. The existence of event horizon within Einstein's field equations is linked to the application of point-like sources in the local field theory. With vacuum field theory, it is observed that event horizon can no longer form without infinite classical energy. Gravitational waves are also discussed relative to the use of point-like sources in Einstein's field equations and similar geometric field theories. Methods for determining the space-time metric of any object on a per particle basis are provided. The continuous model of the universe is further introduced, where the solutions to several grand cosmological problems are discussed. It is demonstrated that an asymptotically flat universe will appear linear with respect to local observers. The inferred accelerated expansion is an illusion due to local geodesics deflecting towards the center of an asymptotically flat, linear universe. With recent constraints on the abundance of faint blue galaxies and observed evolution,  $\Lambda$ CDM is found to be off in galactic number densities by  $70\% \pm 15\%$  and  $104\% \pm 25\%$  at  $0.5z$  and  $1.0z$  respectively. These galaxies are also observed to be similar to local disk and irregular populations, where  $\Lambda$ CDM underestimates their size by  $200\% - 300\%$  prior to  $0.7z$ . This implies that an expanding model predicts the incorrect shape of the universe, which induces systematic lensing errors. After eliminating all viable explanations, an expanding universe is conclusively ruled out. The purposed model however agrees with all observations by applying only classical assumptions. The shape of the universe for example supports a central core, which is responsible for the cosmic background radiation. It is further argued that Einstein's field equations are incompatible with such universe due to predictions of event horizon.

---

## Preface

Due to the diversity of subjects discussed, this page is meant to provide an overview of the paper. The two theories included herein are referred to as vacuum field theory and the continuous model of the universe. These deeply interrelated theories are necessary for complete consistency between general relativity and cosmology. The cosmological aspects are further applied to rule out various theories of general relativity. The foundations of vacuum field theory arise from three postulates with respect to a unified field theory. These postulates however are only introduced for additional insight, as vacuum field theory can be derived from classical laws of physics. The first is a  $1/r$  gravitational potential for any particle, e.g. an electron or proton. The second is Einstein's equivalence principle, where a particle in one local frame will be identical to itself in any other local frame.

Rather than particles being point-like sources, it is argued that they instead exist as localized fields throughout chapters 1 and 2. The foundations of vacuum field theory can be viewed in terms of waves travelling through a relative medium. The medium is an energy density with respect to the localized field interpretation of particles. It is demonstrated that the gravitational and electric potential of a charged, non-composite particle are directly proportional to vacuum energy density. After applying the Lorentz transformation to  $1/r$  fields, deceleration of a charged particle is directly related to a change in vacuum energy density in the form of bremsstrahlung. The field dynamics within classical electrodynamics are mimicking those of vacuum field theory. It is therefore possible to formulate theories with point-like sources that agree with observations. The underlying vacuum energy density however defines a locally isotropic space-time metric for general relativity. It would therefore be incorrect to treat the space-time metric as an additional medium for the continuum limit of point-like sources to influence. By knowing how each

particle's field varies due to the background field induced by all others in consideration, the effective space-time metric can be determined for objects on a per-particle basis. With the per-particle method provided in chapter 2, it is observed that conical singularities or event horizon can no longer form without infinite energy. The application of point-like sources and coupling to space-time metric is responsible for the predictions of event horizon and gravitational waves in modern general relativity. Direct proof for the localized nature of particles will therefore arise from a null detection of gravitational waves with direct methods. Although a null result would invalidate the coupling of point-like sources to a space-time metric, the cosmological aspects already rule out any theory that allows event horizon from finite energy.

The cosmological model is the central discovery of this paper, where it is demonstrated that the universe is asymptotically flat. In other words, the inferred accelerated expansion is an illusion due to local geodesics deflecting towards the center of the universe. With redshift arising from relative motion and gravitational potential, the observed state of the universe can only be fit by accelerated expansion or an asymptotically flat shape. All observations are further in agreement with a linear, asymptotically flat universe as discussed throughout chapter 3. These include galactic number densities, angular size versus the absolute magnitude of faint blue galaxies and time-dependence. Although Hawking radiation is theorized to exist with respect to event horizon, the 3000 K temperature of the cosmic background radiation would require the core to be many orders of magnitude less massive than the Moon. Countless galaxies and clusters are however continuously flowing towards the center of the universe. The central core must therefore be more massive than any local object, i.e. the observed cosmic background radiation offers direct proof against the existence of event horizon.

## 1. Foundations

Throughout the history of modern physics, many attempts at developing a viable unified theory have been made. These attempts have diverse underlying principles, most lacking physical interpretation. Without providing the entire unified field theory, it is possible to reformulate general relativity with three fundamental postulates. These postulates are derived from classical principles, which are further discussed relative to the standard model and Einstein field equations (EFE). Methods are derived that allow the effective space-time metric of any object to be determined. These require a revision of general relativity for several reasons, which are discussed throughout the first and second chapters.

(1) EFEs are based upon a continuum limit of point-like sources, which act locally on the space-time metric. The metric is in return mimicking the localized nature of particles; i.e. similar to classical electrodynamics, space-time acts as a medium for waves. However, the actual field general relativity depicts is the underlying vacuum energy density. This includes contributions from classical and semi-classical fields, although only the electromagnetic field is thoroughly discussed herein. It is argued that the principle of locality is invalid and particles exist as localized field rather than point-like objects.

(2) Special relativity demands that a particle's field will deform from variations in relative motion. General relativity should be restricted to similar mechanisms. When multiple particles interact, the field of each is deformed due to its locally invariant nature with respect to the space-time metric.

(3) All massive particles are known to display electromagnetic behavior; however, EFEs decouple mass from the electromagnetic field. Therefore, the Schwarzschild solution cannot represent realistic objects, as even neutrons display non-zero magnetic dipole moments and electromagnetic form factors.

**Postulate I: Classical forces are mathematical constructs, approximating the time dependence of vacuum fields.** At this introductory level of vacuum field theory, the focus is directed at both classical and free field force(s). Classical force is a time-dependent variation that acts upon a point-like particle. Regardless if the discussion is general relativity or electrodynamics, force determines the time-dependence of momentum and position. Some theories also produce abstract fields that are related to position and momentum. For example, quantum mechanics provides probabilistic wave functions of an underlying semi-classical system.

In a free field theory, force refers to the action at each point in space; this is not necessarily in a classical sense. For example, say a field existed that represented a single electron. Regardless of the underlying complexity, the electron will have a classical location in space. In addition to the finite energy density at the particle's classical position, all other points in space will have finite energy density. The introduction of another electron displaced from the original would further vary the underlying field at all points in space. The force(s) between these two localized fields arise from the infinitesimal action of the effective field at each point in space. Particles are therefore localized entities displaying action at a distance.

References to quantization throughout this paper refer to reducing the localized nature of particles to point-like objects. Classical forces are then applied to determine the time dependence of position and momentum in quantum systems. This is achieved by applying the Lorentz transformation or space-time metric to a scalar field; the scalar field is related to vacuum energy density. The vacuum energy density of a single particle can further be approximated with classical theory, i.e.  $1/r$ .

**Postulate II: All particles consist of localized vacuum fields.** This is relevant to the concept of field-particle duality, where all particles display decaying fields and point-like structure. Under the most fundamental considerations, objects would cease to exist if matter was not localized. This concept can further be extended to what it means for a field to be localized. Localization requires that the underlying energy is self-reinforced, i.e. any stable vacuum field will not dissipate over time. For this to be plausible, at least two forms of field energy must exist. A complex scalar-vector field is defined below; however, this is inadequate for an interacting theory. Additional degrees of freedom are instead required for time-dependent evolution. The main objective of this paper is to bridge the gap between electrodynamics and general relativity. These additional degrees of freedom can therefore be ignored by applying classical forces.

Vacuum energy density is related to a complex scalar-vector field that is conserved throughout interactions. Due to this conservation, the envelope of fields can be approximated with linear wave-like equations. Furthermore, classical physics requires for the far-field of massive particles to decay as  $E_0/r$  in Planck units; this is later discussed in section (1.3). With the wave function  $= \emptyset + i\vec{v} \rightarrow$ , a complex Helmholtz equation (1) is needed. It is superficially similar to the Schrodinger equation of a free particle, but does not depict probability.

$$(A^2 + k^2) = 0 \quad (1)$$

Equation (1) can be divided into both real (2) and imaginary (4) parts.

$$(A^2 + k^2)\emptyset = 0 \quad (2)$$

In spherical coordinates, solutions to (2) involve spherical harmonics and Bessel functions (3).

$$\emptyset_{lm} = j_l(kr)Y_{lm}(\theta, \varphi) \quad (3)$$

The complex component of equation (1) takes a similar form (4).

$$(\vec{A}^2 + k^2)\vec{v} \rightarrow = 0 \quad (4)$$

Solutions to (4) can be determined from the scalar component via (5).

$$\vec{v} \rightarrow_{lm} = \frac{1}{k} A \emptyset_{lm} \quad (5)$$

Conservation of vacuum energy can be achieved by introducing the continuity equations (6) and (7). These relate to an underlying geometric structure after reducing the additional degrees of freedom. The equations however do not depict the correct time dependence of vacuum fields, which can instead be approximated with classical theory. All that is required after the quantization of a localized field is the Lorentz transformation and assumption of  $1/r$  gravitational potentials. The application of a single scalar-vector field is important due to the quantization process. It is therefore assumed that  $\emptyset$  remains constant at the classical position of a massive, non-composite particle. These variables are further related to a scalar invariant and motion of a point contained within <sup>3</sup>.

$$\frac{1}{c} \frac{d\emptyset}{dt} = -\vec{A} \cdot \vec{v} \rightarrow - A \emptyset \quad (6)$$

$$\frac{1}{c} \frac{d\vec{v} \rightarrow}{dt} = A \emptyset + (\vec{A} \cdot \vec{v}) \hat{\vec{v}} \quad (7)$$

The complex Hamiltonian density defines the vacuum energy density at any point in space with a quaternion norm (8).

$$\nabla(r) = \sqrt{\emptyset^*} = \sqrt{(\emptyset)^2 + (v)^2} \quad (8)$$

The linear wave solutions can be quantized with equation (9), i.e. the point of maximum field energy depicts the classical energy.

$$E = f^\infty \sqrt{\emptyset^*} \delta(r) dr \quad (9)$$

**Postulate III: The vacuum field is the result of transforming the non-linear geometric degrees of freedom.** A crucial metaphysical aspect of matter is usually overlooked in modern physics, i.e. what do particles physically consist of? Initial attempts tried to attribute a physical substance to matter, or material upon space that formed particles. However, this perspective is plagued by cyclic reasoning, i.e. if such substance existed, what would be the physical essence of it? Indeed this reasoning is no different from the modern concept of fields. For example, electromagnetic fields are mathematical constructs created in abstract to understand the universe at the quantized level. At any scale however, one fundamental property of the universe is undeniable; i.e. space itself.

In the classical perspective space is a rigid, time-independent structure that quantized mechanics is founded upon. Switching to the more abstract view of general relativity, the properties of space vary from the Euclidean model. It becomes possible to deform space, varying the location of a continuum of points in a smooth manner. However, Einstein's view of relativity is incompatible with quantum mechanics and the cosmological model discussed in chapter 3. He also applied the only physical property of the universe to a single classical force, i.e. gravity. It is demonstrated in section (1.4) that general relativity can be reinterpreted as a tool for quantization. Einstein's perspective is therefore the conceptual inverse of vacuum field theory. This opens a profound path to unification, as an underlying geometric structure can be used to depict all classical forces. Therefore, unification no longer refers to the energy scale where classical forces merge into one, but instead the manifestation of all forces from a single unified field. The essence of matter can now be attributed to something that is physically real rather than a mathematical construct.

Assuming tensors and/or a geometric foundation are capable of fulfilling the first two postulates, the necessity for additional degrees of freedom is clear. The vacuum field represents the energy of an underlying geometric structure, although the actual structure is beyond this paper's scope. Regardless, acknowledging its existence offers an intuitive explanation for the universe. It is trivial that a geometric structure should be time-dependent if it does exist. Therefore, in the most general sense matter is nothing more than fluctuations of space itself. These are much smaller than the macroscopic world, as vacuum field theory indicates structure at the Planck scale. Without assuming Planck scale fluctuations of space are responsible for fields and matter, there is literally no other way of writing a unified field theory. For example, the standard model applies several scalar-vector fields to complete symmetries and fill gaps; however, they are solely mathematical constructs.

As earlier theories developed, the original aether became a resistive medium throughout space rather than mysterious substance that formed particles. This transition was the product of the corpuscle theory of light, attributed to Newton. It was later argued against with the Michelson-Morison experiment, which tested for a variation in the speed of light relative to the local motion of Earth. This concept of anisotropy is flawed, which had been pointed out by Hendrik Lorentz<sup>[A]</sup>. As an object's momentum varies in a local frame, the field is transformed in such a way that any anisotropic affects cancel. Motion is instead relative to the vacuum field of all other particles and aether only becomes conceptually crucial for a single particle universe. In other words, the scalar vacuum energy density creates a relative medium upon space, which must further be applied to determine the effective space-time metric.

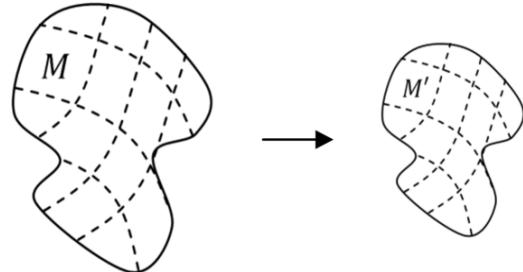
## 1.1. Space

When describing the dynamics of particles in a gravitational field, general relativity is a useful theory. However, the theory can be interpreted in two unique perspectives. The first is the mainstream view, where the four-dimensional manifold is to be taken literally; i.e. the coupled entity of space-time physically exists. Gravitational acceleration within this perspective is not induced by the curvature of space, but instead the curvature of time. The curvature of spatial components only varies the path of particles in motion or under classical force. Regardless, the physical existence of space-time is crucial for not only Einstein's interpretation, but also the validity of mainstream astronomy and cosmological models. The undermining of this perspective however originates from the predictions of singularities and gravitational waves. The use of quantized point sources in the geometric Laplace equation (EFEs) produces these artifacts. Particles instead exist as localized fields rather than point-like sources and must be treated as so.

The second view insists that the universe does not physically exist as a four-dimensional space-time manifold. Space and time should instead be treated as two independent entities; i.e. space is depicted by a classical manifold mapped to physical locations, while time is a manifestation of relativity or event comparison. Within this perspective, the vacuum field of each particle is relative to a Euclidean reference space ( $y_{\mu\nu}$ ), which can be arbitrarily chosen. Upon this reference space or frame, particles exist as localized scalar-vector fields. Affine transformations are applied to the field of each particle rather than varying the metric of space, in agreement with the principles of special relativity and electrodynamics. Any reference space is further held static so that it does not allow metric or gravitational waves.

In differential geometry, a metric is defined that maps all points ( $\xi$ ) bound to a manifold in space to a curvilinear coordinate system ( $x$ ), or vice-versa. This can be defined as an infinitesimal variation in distance between two points in space, with respect to the original configuration. A two dimensional example is depicted in figure 1.1. It is always valid to vary the points on a manifold as long as they never overlap; this is referred to as a Riemann manifold. This deformation is possible due to the infinitesimal property of nature, i.e. there exist an infinite number of infinitesimal intervals between two points in space.

In another perspective or the one previously argued for is the deformation of a ruler, which consists of a linear lattice of atoms. These atoms contain many quantized particles, although each is relatively localized at a single point in space. When vacuum energy density increases, each particle's field must remain locally invariant with respect to the space-time metric. The space-time metric is therefore encoding the deformations or affine transformations experienced by localized fields, which can further be treated as point-like objects. In a situation where the gravitational potential is increasing with respect to time, distance from the perspective of the ruler remains constant. However, an observer in the reference space will note that length contraction of the ruler has taken place. Only the lattice of atoms becomes deformed with respect to the reference state rather than space itself.



**Figure 1.1.** A smoothly connected manifold ( $M$ ) is transformed from an initial state ( $M$ ) to a final state ( $M'$ ).

## 1.2. Time

Quantization or the transformation of a localized field into a point-like object allows for the creation of systems. A system exists in a finite region of space and may contain many quantized particles, usually coupled to kinematic equations. Therefore, systems are time-dependent and obey classical energy conservation, i.e. Lagrangian formulations. Within classical mechanics, the trajectory of a particle through a potential can be determined with the Lagrangian equation (10); where  $L$  is the kinetic energy minus potential.

$$\frac{d}{dt} \left( \frac{\partial L}{\partial q_i} \right) - \frac{\partial L}{\partial q_i} = 0 \quad (10)$$

The 1-dimensional Lagrangian can be written as (11), where  $V$  is the classical potential.

$$L = \frac{1}{2} m \dot{x}^2 - V(x) \quad (11)$$

In this perspective, time is a comparative scalar related to a change in position with respect to some constant rate of observation. It allows for a coupling between quantized energy and the rate at which quantized fields move through space.

Relativity introduces additional complexity as the rate of observation can vary between different scenarios. Both special and general relativity require a Lorentz scalar, where an increase in vacuum energy density forces a field to evolve at a decreased rate. This time dependence of classical variables is seen throughout various phenomenon including decay rates, classical kinematics and intrinsic spin. For example, it is experimentally known that an unstable particle has a relatively longer half-life when moving with respect to a background field. Particle decay depends upon internal degrees of freedom, which cannot be explained by classical mechanics. The longer half-life is instead modeled by relativity, where all

variables evolve at a decreased rate. When quantum mechanics is discussed, it is demonstrated that the intrinsic spin rate of an electron also decreases as vacuum energy density increases. The Dirac equation is further related to relativistic field dynamics, discussed throughout section (1.6). It is clear that time is more fundamental than the kinematics of point particles, i.e. time is a scalar quantity that depicts change in abstract variables.

Under certain conditions, it is possible for time to be undefined. If no energy exists, then there is nothing to compare and space at all scales becomes Euclidean ( $y_{\mu\nu}$ ). The other extreme is a region of infinite vacuum energy density such as an event horizon or conical singularity. Time in this situation is undefined because the underlying field(s) cannot evolve. At this limit, quantum mechanics fails under classical considerations since the observables become static. Einstein's field equations allow such

anomalies due to the use of point-like sources in a local geometric field theory. This results in the non-linearity between vacuum energy density and space-time metric. The only way to produce infinite vacuum energy density in vacuum field theory is with infinite classical energy, which is impossible.

EEFs are more abstract with respect to time, or space-time. For example, time-dependent variations in the stress-energy tensor can produce gravitational waves. By applying quantized mass and momentum in such a local field theory, the field of point-like sources becomes geometric in nature. EEFs provide the framework to determine the time dependence of these localized fields, in which some scenarios allow quantized variables to be transformed into geometric waves. General relativity is however depicting the vacuum energy density of field(s) responsible for classical forces. Time in a more general sense is therefore a comparative scalar between the various variables of a quantized system or underlying field theory.

### 1.3. Energy

Effective energy is defined from quantized rest mass and momentum (12).

$$E = \sqrt{(pc)^2 + (m_0 c^2)^2} \quad (12)$$

Returning to the classical wave-like equation (1), effective energy can also be written as a quaternion norm. With a particle's scalar-vector field ( ), quantized energy is related to the vacuum energy density (13). From the third axiom, this density is

related to the underlying geometric degrees of freedom. The goal is to transform the non-linear geometric structure into a scalar-vector field that is linearly proportional to classical energy.

$$\nabla(r) = \sqrt{\quad^*} \quad (13)$$

In order to quantize the vacuum field, similar methods are applied with respect to the linear wave approach (9); i.e. applying a Dirac delta defines quantized energy (14).

$$E = \int_0^\infty \nabla(r) \delta(r) dr \quad (14)$$

The wave function can be written as (15), where quantized mass depicts the scalar field and momentum replaces the vector component. Scalar mass no longer exists solely at the particle's center, as it is a fundamental part of all field solutions.

$$(0) / c^2 m_0 + i \bar{c} p \rightarrow \quad (15)$$

With the quaternion norm used to define vacuum energy, it possible for scalar mass to be negative. Relative to the Dirac equation, the charge conjugate is applied to ensure only positive mass exists. Negative scalar mass however always results in positive vacuum energy density, which depicts the gravitational force. Thus  $\nabla$  will always be positive and the quantized scalar field will have the same sign as charge for non-composite massive particles.

With the second axiom and the expected inverse distance far-field, a spherically symmetric solution is possible. This is derived from the classical wave-like equation and is only meant to approximate the field's envelope. After applying spherical Bessel functions, two linear wave solutions are found for non-composite particles in Planck units (16). These solutions have an  $E_o/r$  far-field with energy density at the classical position proportional to  $E_o$ .

$$\frac{\sin(r)}{r^2} \quad \frac{\sin(r)}{r} \quad \frac{\cos(r)}{r^2} \quad (16)$$

$$\nabla(r) = E_o \sqrt{\left( \pm \frac{1}{r} \right)^2 + \left( \pm \frac{1}{r^2} \right)^2} \quad (16)$$

The approximate wavelength that corresponds to an electron or positron is therefore  $2\pi$  in Planck units. In SI units however the wavelength is  $2\pi l_p$ , where Planck length is defined as (17).

$$l_p = \sqrt{\frac{k c_o}{\Delta}} \cong 1.616199 \cdot 10^{-35} \text{ m} \quad (17)$$

As previously stated, the scalar-vector notation is not adequate for time-dependent evolution; i.e. (16) finds no real application beyond approximating the far-field vacuum energy envelope. In general, solutions for actual electrons and positrons are only approximated by equation (1) and the resulting linear solutions (16). All that is required by the purposed postulates is for vacuum energy density to be indefinitely localized in space, creating a stable  $1/r$  far-field. Planck units are used due to a relation between the space-time metric and vacuum energy density. This effectively sets  $G_o = \hbar_o = c_o = k_e = k_B = 1$ ; i.e. the far-field gravity-electric potential of an electron is defined as (18), where  $\sigma$  is the charge to mass ratio. Classical energy variations from this potential are now directly proportional to variations in vacuum energy density or  $\nabla \cong E_o/r$ .

$$U = \nabla(1 \pm \sigma) : \sigma = -15.15612(63) \frac{q_p}{m_p} \quad (18)$$

Energy (12) can be reformulated by introducing the Lorentz scalar defined by (19), which is relative to a local frame of reference or space-time metric.

$$\zeta = \sqrt{1 + (\frac{v}{c})^2} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (19)$$

The Lorentz factor scales rest energy, resulting in the effective quantized energy (20).

$$E = E_0 \gamma \quad (20)$$

The Lorentz factor (19) is simply the ratio of energy to rest energy, and has a range from 1 to  $\infty$ . For agreement between a relative and general frame, a

scalar field  $\zeta_g$  is introduced. Generalization of  $\zeta_g$  is achieved by looking for a function similar to  $\zeta$  with a range from 1 to  $\infty$ . This is accomplished by first defining the reference vacuum energy density (21), which any observer will consistently measure as constant.  $\Delta$  plays the role of  $E_o$  as defined in (20), which is a product of quantized energy being proportional to vacuum energy density.

$$\Delta = \frac{c_o^4}{G_o} \quad (21)$$

The net vacuum energy density is defined as (22). There also exists a simple relation (23) between  $\Delta$  and  $\bar{\Delta}$  similar to  $E = E_0 \zeta$ .

$$\Delta = \Delta_b + \psi \quad (22)$$

$$\bar{\Delta} = \Delta \gamma_g \quad (23)$$

The domain can also be extended from 0 to  $\infty$  when considering reference frames within a local field. However, all applications within this paper use a reference frame where the local source is removed ( $r \rightarrow \infty$ ), or as  $\nabla \rightarrow 0$ . To ensure equivalence as previously purposed, the correct equation meeting all requirements is (24).

$$\gamma_g \frac{\bar{\Delta}}{\Delta} = 1 + \frac{\nabla}{\Delta} = \frac{1}{1 - \frac{\nabla}{\Delta}} \quad (24)$$

Newtonian energy principles (25) can be derived by equating  $\zeta_g = \zeta$ . The general Lorentz scalar is defined relative to  $1_{\mu\nu}$ , while the velocity within  $\gamma$  is with respect to  $g_{\mu\nu}$ .

$$\frac{E}{E_o} = 1 + \frac{\Delta E}{E_o} = \frac{\bar{\Delta}}{\Delta} = 1 + \frac{\nabla}{\Delta} \quad (25)$$

Special relativity defines force as the change in proper momentum with respect to metric time (26). The following notation will be used for common variables;  $u = dx'/dr$  is proper velocity,  $v = dx'/dt'$  is metric velocity and  $w = dx/dt$ .

$$\mathbf{F} = \frac{dp}{dt'} = \frac{\gamma^3 m v a}{c_o^2} \quad (26)$$

The equations can be simplified by considering an object moving at escape velocity along the gradient of  $\nabla$ . If  $\nabla$  is a single static particle, the escape velocity (27) relative to the particle's field is derived via  $\zeta = \zeta_g$ . Taking the limit of (27) as  $r \rightarrow \infty$  is  $\dot{v} = 0$ , while the limit as  $M \rightarrow \infty$  results in  $\dot{v} = c_o$ . When transforming (27) to the frame relative to a distant observer or  $\Delta$ , the velocity as  $M \rightarrow \infty$  is  $\dot{w} = 0$  implying infinite vacuum energy density. This is a consequence of  $\nabla \rightarrow \infty$  rather than  $M \rightarrow \infty$ .

$$\dot{v} = c_o \sqrt{\frac{2G_o M}{r c_o^2 + G_o M} - \left( \frac{G_o M}{r c_o^2 + G_o M} \right)^2} \quad (27)$$

Acceleration is derived by differentiating (27) with respect to  $t'$  and applying the chain rule. The derivative with respect to  $r'$  can be transformed to  $r$  by applying another chain rule (28).

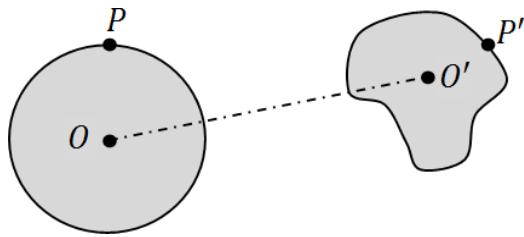
$$\frac{d}{dr'} = \frac{dr}{dr'} \frac{d}{dr} = \frac{1}{\gamma_g} \frac{d}{dr} \quad (28)$$

Acceleration relative to the space-time metric at escape velocity is therefore equal to (29).

$$a = \frac{d^2 r^7}{dt'^2} = -\frac{G_o M}{r^2 \gamma_g^4} \quad (29)$$

## 1.4. Continuum Mechanics

A localized field can be represented as a group of sections or wave fronts, where the space-time metric is applied for the quantization process. This reference to quantization is not with respect to quantum gravity, but instead the transformation of localized fields into point-like objects. From the perspective being argued within this paper, the proper view is that of localized fields deforming due to the presence of other localized fields. Classical mechanics are therefore replaced by continuum mechanics and resulting metric(s). For example, a finite manifold ( $M$ ) can be equipped with an arbitrary metric ( $g^{\mu\nu}$ ). The metric will then undergo affine transformations as depicted in figure 1.2. For simplicity, each manifold is smoothly connected and semi-rigid. Manifolds of this type are useful for describing vacuum fields, although the concept is further complicated by intrinsic spin. Relative to vacuum field theory, each particle exists as an independent manifold ( $g^{\mu\nu}$ ) that is equipped with a scalar field. The scalar field or vacuum energy density is in return relative to the preferred reference frame. Each localized vacuum field also has a reference state ( $O$ ) with respect to the space-time metric. With special relativity and classical electrodynamics, it is known that additional energy or momentum induces length contraction of the underlying field(s). This can be extended to general relativity by applying the spatial components of the contravariant metric tensor. Figure 1.3 demonstrates how general relativity allows for the quantization of a localized field.



**Figure 1.2.** A finite 2-dimensional manifold is depicted in an initial ( $O$ ) and final state ( $O'$ ). Several affine transformations are applied including translation, rotation and deformation.

Applying the contravariant and covariant metrics together results in any underlying field to be transformed back to the  $y_{\mu\nu}$  state; this is due to the equivalence principle and equation (30).

$$g^{\mu\beta}g_{\beta\nu} = \delta^\mu_\nu \quad (30)$$

Transforming each field in this manner allows for quantization, where the velocity of each point upon the field is equivalent in both direction and magnitude. The dynamics of the field can therefore be reduced to the point-like particle perspective. Without quantization, each point along a particle's manifold will travel at various velocities. This will further require continuum mechanics to determine the proper translation, rotation and deformation.

A particle's manifold will vary depending upon the background vacuum energy density that it resides in. Under realistic considerations, the particle also influences the gravitational potential generated by the background field. If the effective energy and proper velocity of each particle is known, then it is possible to determine the effective space-time metric of any object; this is discussed in chapter 2. If a particle is moving with respect to a local field, both special and general transformations must be applied. From (30), it was shown how wave fronts become equivalent to the  $y_{\mu\nu}$  frame when space-time is deformed. Any Lorentz boost (38) must be applied relative to this configuration and then mapped from  $g_{\mu\nu} \rightarrow y_{\mu\nu}$ . This ensures that all field transformations are invariant for local observers; i.e. any stationary observer will view space relative to the space-time metric.



**Figure 1.3.** Sections of two manifolds  $g^{\mu\nu}$  are depicted relative to a specific reference frame. (Left) The frame is with respect to the Euclidean frame  $\eta_{\mu\nu}$ . (Right) The frame is with respect to the deformed space-time metric  $g_{\mu\nu}$ .

Assuming a particle's classical position is with respect to the preferred reference frame ( $y_{\mu\nu}$ ),  $r_n$  denotes an independent coordinate system for each. The notation can be simplified by introducing  $\Psi_{mn}$ , which is the non-effective field from particle m to a point in the  $n^{th}$  coordinate system. The effective field from particle m at the same point relative to the  $n^{th}$  coordinate system is  $\Psi_{mn}$ . Equation (31) is based upon the equivalence principle, where each particle is relative to the effective background field or space-time metric.

$$\Psi_{mn}(r_n) \quad \left\{ \begin{array}{l} m = n : 0 \\ m \neq n : \bar{\Psi}_n(r_n) \end{array} \right. \quad (31)$$

The effective field of particle n due to all other particles is defined by equation (32),

$$\bar{\Psi}_n(r_n) = \Psi_n(r'_n) \quad (32)$$

where  $r'_n$  is determined by integration (33).

$$r'_n = r_n + \frac{1}{\Delta} \int_0^{r_n} \left( \sum_{m=1}^N \Psi_{mn}(r_m) \right) dr_m \quad (33)$$

Relative to the Euclidean line element ( $dS$ ), the transformed radial coordinate is defined as (34).

$$r'_n = \int_g(r \rightarrow n) ds = \int_0^{r_n} ds \quad (34)$$

The line element can be written in terms of the metric tensor (35).

$$(ds)^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (35)$$

If only a single non-composite particle existed,  $g = 1$  and the field would be in the original configuration with respect to the co-moving frame. Therefore, the field of a single electron or positron would exhibit the  $E_0/r$  far-field envelope in agreement with classical theory. The maximum value of this localized field is proportional to E at the classical position.

If a point or region of infinite vacuum energy density ( $\Psi$ ) exists,  $g$  also becomes infinite. The previous mapping (33) fails at boundaries of infinite vacuum energy density or points beyond them. For example, if the field of an external particle is calculated, any radial lines at or beyond the singularity will be mapped to the event horizon. The wave fronts or sections also become non-continuous, violating the assumption of smoothly connected vacuum field manifolds. It is known from classical electrodynamics, QED and QCD that particles exist as localized fields. Due to these fields following the metric of space-time, they become restricted to any boundary of infinite vacuum energy. Since these fields are in return responsible for all classical forces including gravity, a black hole with event horizon would lack an external field. This is contradictory to EFEs, where objects with event horizon display external fields.

Finite black holes are predicted to exist with respect to vacuum field theory, although their surfaces must have finite vacuum energy density. Therefore, black holes should not only demonstrate external fields, but also nearly perfect black body spectrums. With vacuum field theory, it is literally impossible to create a point of infinite vacuum energy density from finite classical energy. From the methods herein, the effective field of any object can now be determined down to the Planck scale. With the advent of QCD and resulting states of dense quark matter, it is now possible to model the finite fields of quark stars and black holes. The name given to black holes remains valid since they demonstrate near perfect black body spectrums and immense gravitational fields. Energy will escape over time due to relativistic jets and free field radiation. The non-existence of event horizon and singularities is later discussed with respect to the cosmic background radiation and observed shape of the universe.

## 1.5. Electromagnetic Fields

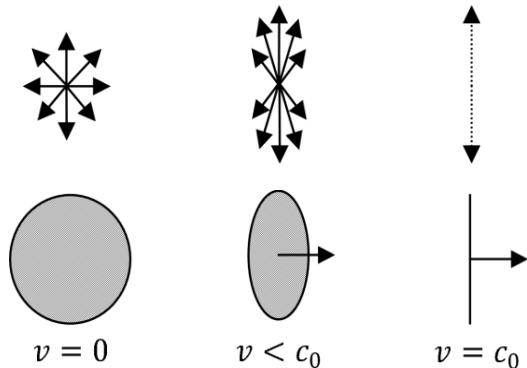
Relativistic electrodynamics provides additional insight into how the vacuum field of a particle varies due to relative velocity. As the momentum of a uniformly charged particle increases, the electric field lines and magnitude loss isotropy. The electric field in terms of the particle's classical position is given by (36)<sup>[B]</sup>, where  $\theta = 0^\circ$  is parallel with  $\vec{v} \rightarrow$ .

$$\vec{E} = k \frac{q \vec{r} \rightarrow}{r^3 (1 - (\frac{v}{c_0})^2 \sin^2(\theta))^{3/2}} \quad (36)$$

The magnitude of the electric field is therefore (37).

$$E = k \frac{q \gamma}{e r^2} [\gamma^2 \cos^2(\theta) + \sin^2(\theta)]^{-1/2} \quad (37)$$

As the particle's momentum increases, the electric field is compressed in the direction of motion. The magnitude of the field perpendicular to motion also increases, while the electric field tangent to the particle's trajectory weakens. As the limit of  $v \rightarrow c$  is approached the field becomes compressed into a cylindrical plane of infinite vacuum energy density, depicted in figure 1.4. These field dynamics with respect to the Lorentz transformation are crucial to the application of point-like sources in classical electrodynamics. This however does not indicate that the use of point-like sources in other field theories is valid.



**Figure 1.4.** (Top) The electric field of an electron is provided in several states: (Bottom) An electron's manifold undergoing length contraction in the direction of motion.

For a massive non-composite particle, vacuum energy density is not proportional to the electric field. The electric potential can instead be related to vacuum energy density as demonstrated by equation (18). Returning to the foundations of vacuum field theory, there is a distinction between quantized energy and vacuum energy density. This is because quantized energy is with respect to the amount of vacuum energy density at the classical position. It is important to notice however that the gravitational potential is both equal to vacuum energy density and the variations in classical energy due to a potential. Since quantized energy has only two components, the classical dynamics of a point-like particle can therefore be retained. The remaining field dynamics are derived by applying a Lorentz transformation to the vacuum field of a particle.

The Lorentz transformation  $\Lambda$  (38)<sup>[B]</sup> allows mathematically defined objects to be transformed in space-time. The objects in this case are individual manifolds equipped with a scalar field, which depicts the vacuum energy density of each particle.

$$\Lambda = \begin{bmatrix} \gamma & -\gamma\beta_i & -\gamma\beta_j & -\gamma\beta_k & 1 \\ -\gamma\beta_i & (1 + \beta_{ii}) & \beta_{ij} & \beta_{ik} & 0 \\ -\gamma\beta_j & \beta_{ij} & (1 + \beta_{jj}) & \beta_{jk} & 0 \\ -\gamma\beta_k & \beta_{ik} & \beta_{jk} & (1 + \beta_{ki}) & 0 \end{bmatrix} \quad (38)$$

Where  $\beta_{\mu\nu}$  is defined as (39).

$$\beta_{\mu\nu} = \frac{(\gamma - 1)\beta_\mu\beta_\nu}{\beta^2} = (\gamma - 1) \frac{\vec{v}_\mu \vec{v}_\nu}{\vec{v}^2} \quad (39)$$

In general any 4-vector can be transformed via (40), including the electromagnetic 4-potential.

$$R' = \Lambda R \quad (40)$$

Two consecutive Lorentz boost can be determined with matrix multiplication (41).

$$\Lambda(\bar{v}_1 + \bar{v}_2) = \Lambda(\bar{v}_1)\Lambda(\bar{v}_2) \quad (41)$$

With respect to the electromagnetic nature of matter, vacuum field theory would be incomplete without discussing photons. From the approximate field of a non-composite massive particle (16), vacuum energy density is determined from the classical gravitational potential and linear wave-like equation (1). This says nothing about how the electromagnetic field relates back to the unified

scalar-vector field. There is also a distinction between individually localized photons and free electromagnetic energy. Classically, an electromagnetic field can be described in terms of a superposition of waves. However, there is no guarantee that these are individual packets of vacuum energy. Unlike the electric field, superimposed vacuum fields do not display field interference; i.e. positive and negative electric field contributions will result in no electric field, while vacuum energy density is always positive. It is therefore possible to use Fourier series to create a super-position of many waves, although these will not represent actual particles. Two processes in nature provide insight into the distinction between free electromagnetic energy and localized photon (or quantized particle); these are electron-positron annihilation and bremsstrahlung.

Electron-positron annihilation demonstrates the particle nature of the electromagnetic field, where a

particle and its anti-particle produce gamma rays after colliding. The localized scalar fields of the electron and positron cancel, creating two or more massless photon. Assuming conservation of energy, the resulting photons will split quantized energy. Since the total vacuum energy of a single electron is infinite, only the conservation of classical energy can be considered. The gamma rays resulting from the annihilation process also remain localized in space indefinitely. This has implications for the CMBR, or black body spectrum observed in all directions of local space.

Each photon or localized packet of electromagnetic energy remains localized due to the reinforcement of vacuum energy density. Assuming the symmetry of a photon is cylindrical, Bessel functions of the first kind (42) are applied perpendicular to the direction of propagation.

$$J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m+n+1)} \left( \frac{x}{2} \right)^{2m+n} \quad (42)$$

If a linear wave equation is used similar to (1), the gradient of the scalar component must have the opposite sign of the vector field. This will essentially require for the central region of the field to have quantized mass. Therefore, the linear-wave approximations cannot be taken literally. An actual photon will have no scalar mass at the classical position, and will therefore be anti-symmetric with respect to the vacuum scalar field. The wave equation is instead applied to approximate the far-field envelope under consideration of vacuum energy conservation. The exact nature of the underlying vacuum field is therefore irrelevant, as only vacuum energy density is required with respect to general relativity. Applying the linear wave approximation (1), a photons field perpendicular to propagation is (43).

$$\nabla = E_o \sqrt{\left(\frac{\partial J_0}{\partial r}\right)^2 + \left(\frac{\partial J_1}{\partial \theta}\right)^2} \approx \frac{E_o}{\sqrt{r}} \quad (43)$$

In order to determine the 3-dimensional vacuum energy density of a single photon, the envelope in the z-direction is required. Without a rigorous method capable of providing the exact vacuum field of a photon, the linear wave approximation must once again be applied. The vacuum field of an individually localized photon in Planck units is therefore (44), where coordinates are with respect to the particle's classical position.

$$\nabla(r, \theta, z) = E_o e^{-|z|} \sqrt{(J_0)^2 + (J_1)^2} \approx \frac{E_o}{\sqrt{r}} e^{-|z|} \quad (44)$$

Bremsstrahlung is the second case of electromagnetic energy, where a charged particle passing close to another emits braking radiation. Unlike the previous case, electromagnetic energy is radiated as a free field over a range of frequencies. The free field should be treated as an independent manifold with respect to other localized fields and particles. In other words, the free electromagnetic field is locally invariant with respect to the space-time metric. This emission of electromagnetic energy in the framework of point-like sources is similar to the gravitational waves in general relativity. In general, theories that couple fields to point-like sources will generate waves upon the respective medium due to variations in a quantized source.

The emission of electromagnetic energy can further be related to a variation in vacuum energy density. For the non-relativistic case, the radial

Poynting flux is defined as (45)<sup>[B]</sup>.

$$S \rightarrow = \frac{q^2 a^2}{4\pi r^2 c_0^3} \sin^2(\theta) \quad (45)$$

The electromagnetic energy radiated per unit solid angle is therefore (46)<sup>[B]</sup>.

$$\frac{dW}{dtd\Omega} = \frac{q^2 a^2}{4\pi c_0^3} \sin^2(\theta) \quad (46)$$

The first order approximation of bremsstrahlung indicates that a charged particle will emit a free field electromagnetic dipole. It is important to realize that electromagnetic energy in this situation is not localized, but continuous over a range of angles and frequencies. It is distinct from the annihilation case, where two photons are emitted at unique angles in order to preserve quantized energy. With results from chapter 2, the effective vacuum far-field of a moving electron is defined by (47).

$$\nabla = \nabla(x', y', z') = \frac{E_0 \gamma}{\sqrt{(x')^2 + (y')^2 + (z\gamma)^2}} \quad (47)$$

Relative to quantized energy, the Lorentz scalar is the only free variable. Therefore  $\gamma$  will assume two states, i.e. prior to bremsstrahlung ( $(_1)$ ) and after ( $(_2)$ ). The variation of vacuum energy density between these states can be derived by subtracting  $\bar{\nabla}((_2))$  from  $\bar{\nabla}(\gamma_1)$ , i.e. equation (48).

$$\Delta \bar{\nabla}_{EM} = \frac{E_0}{r} \left( \frac{\frac{(_1)}{\sqrt{((_1 \cos \theta))^2 + (\sin \theta)^2}} - \frac{(_2)}{\sqrt{((_2 \cos \theta))^2 + (\sin \theta)^2}}}{\gamma - 1} \right) \quad (48)$$

The initial and final states from (47) are expanded via Taylor expansion at  $\gamma = 1$  resulting in (49).

$$1 + (\gamma - 1) \sin^2(\theta) - \frac{3}{2}(\gamma - 1)^2 \sin^2(\theta) \cos^2(\theta) + \dots \quad (49)$$

Ignoring higher order contributions, equation (49) reduces to a dipole field approximation (50).

$$\Delta \bar{\nabla}_{EM} \cong \frac{E_0}{r} (\Delta \gamma \sin^2(\theta)) \quad (50)$$

The Lorentz transformation derived from classical electrodynamics is therefore directly related to a fundamental scalar field. The electromagnetic field cannot be easily related to vacuum energy, since only the electric potential is proportional to  $\nabla$ . Therefore, the  $r^2$  in (45) is expected, while vacuum energy decays proportional to the inverse distance. The vacuum energy emitted in terms of non-relativistic motion becomes (51). In consideration of a unified field theory, all particles would exist upon a single field in space. However, vacuum field theory depicts each localized field as a deformable manifold relative to a reference space ( $y_{\mu\nu}$ ). In this perspective, vacuum energy density ( $\nabla$ ) radiated due to bremsstrahlung is physically detached from the electron's manifold into a free field described by Maxwell's equations.

$$\nabla \cong \frac{m_0 v a}{r} \sin^2(\theta) \quad (51)$$

## 1.6. Quantum Mechanics

Quantum mechanics was introduced by Erwin Schrodinger, who had initially attempted to create a relativistic theory. Due to the many difficulties related to the relativistic form, the time-dependent Schrodinger equation (52)<sup>[C]</sup> was instead published.

$$ik \cdot = H \cdot = (V - \frac{k^2}{2m} A^2) \quad (52)$$

In consideration of a classical potential such as the electric field produced by a proton, the probability of an electron being detected at any given position is  $|\psi|^2$ . This should not be confused with the wave-function attributed to vacuum field theory, which depicts a Hamiltonian density. The wave function of quantum mechanics can be interpreted in various ways. Vacuum field theory agrees with the path-integral approach to quantum mechanics, where each particle has a classical location in space. It attributes no physical meaning to the quantum wave other than probability. This is known as the minimalist perspective or ensemble interpretation attributed to Max Born<sup>[D]</sup>.

In order to comprehend fermion spin and mass, the field of spin 1 particles must be initially discussed. Photons are the most fundamental spin 1 particle and can be either polarized or non-polarized. For a circularly polarized photon, the spin state is either  $\pm k$ . To simplify the problem, the photon will be reduced to a plane-wave that has a helix shaped electric field. Relative to a massive particle located at a fixed point in space, the propagating EM field will appear to spin around a fixed axis. The electric field however is actually traversing space at the speed of light perpendicular to the spin plane. The field is therefore not spinning with respect to the reference frame. The quantized energy of a photon can be written with respect to the perceived angular frequency or wavelength by equation (53). Circular polarization is identical to

two super positioned plane-waves, so (53) remains true for all photons.

$$E = kw = hf = hc/\lambda \quad (53)$$

The photon's luminal field can be carried over to fermions. By applying spacetime algebra, it is observed that the field of an electron orbits the spin-plane at the speed of light. If this is true, then there must be kinematic effects due to the coupling between light-like field dynamics and space-like trajectories. The classical structure of the electron is discussed at the end of this section, for now the Dirac equation (54)<sup>[E]</sup> is examined for its connection to relativistic field dynamics.

$$ik \cdot = H \cdot = (ca \cdot \hat{p} + \beta m_e c^2) \quad (54)$$

This can be rewritten in a more intuitive way since  $m_e c^2$  is actually related to the intrinsic spin (55).

$$\frac{i \cdot \vec{\alpha}}{c \cdot \& t} = (a \cdot \hat{p} + Q \frac{m_e c}{k}) \quad (55)$$

Setting the rest energy of the electron equal to the spin angular frequency (53) results in a spin radius of (56); this is the reduced Compton wavelength. The radius ( $r_e$ ) is constant relative to the metric of space, tracking a set of points along the field.

$$r_e = k/m_e c \quad (56)$$

In consideration of the quantization process used within relativity, the objective is to demonstrate that tracking a single point upon the electron's field satisfies the equations of motion at the classical position. To simplify the motion of the field, a local orthogonal coordinate system ( $e_1, e_2, e_3$ ) is defined relative to the classical position. Historically, Schrodinger was the first to apply the Heisenberg picture in order to determine the time dependence of the position operator (57).

$$-\frac{ikx}{k} = [H, x] \quad : x = \frac{c}{k} \alpha_k \quad (57)$$

After integrating (57) twice with respect to time, the position operator becomes (58).

$$x_k(t) = x_k(0) + \frac{t c^2 p_k}{i k c} e^{i k t} + \frac{1}{2H} \left( \alpha_k(0) - \frac{p_k^2}{H} \right) (e^{-2iKt} - 1) \quad (58)$$

The last term is the complex quantum oscillation known as zitterbewegung; it is complex due the connection with spinors. Furthermore, the first two terms provide the classical trajectory of the particle, which is the average zitterbewegung path. Applying the Heisenberg picture, particle motion is combined with a non-classical rotation of the field and cannot be directly interpreted in the classical sense. From (58), the zitterbewegung angular frequency and radius are (59).

$$: w_{zwa} = \frac{2H}{k} = \frac{e}{k} \quad : r_{zwa} = \frac{c}{w_o} = \frac{k}{2mc} \quad (59)$$

Spinors in general require two rotations in order to return to the initial state. Picking half the classical radius upon the spin-plane remains a valid option. It is claimed that the field has an angular velocity equivalent to the speed of light; therefore, the angular frequency also doubles. Choosing the classical radius (56) simplifies the situation since it returns to the original state after a single rotation; i.e. the relativistic spin period is defined as (60).

$$T_c = \frac{p_h}{E_0 m_0 v} = \frac{h v}{E_0} \cong \frac{h}{m e^2} \quad (60)$$

To develop a mathematical model of the local coordinate system and spin, spacetime algebra is applied. Geometric (spacetime) algebra allows the geometric product to be defined as (61)<sup>[F]</sup>.

$$uv = u \cdot v + u A v \quad (61)$$

The orthogonal reference vectors are related to the initial set by Lorentz spinors (62)<sup>[F]</sup>, i.e.  $\text{SL}(2, \mathbb{C})$ .

$$e_\mu = R \gamma_\mu R^\sim \quad (62)$$

Spin arises as a bi-vector defined as (63)<sup>[F]</sup>, where  $\gamma_2\gamma_1$  is the spin plane.

$$S = \text{isv} = \frac{k}{2} \text{ie} e^{\frac{k}{2}} \text{R} \gamma_2 \gamma_1 \text{R} \quad (63)$$

For the free wave solution to the Dirac equation, the Dirac rotor becomes (64)<sup>[F]</sup>.

$$R = e^{\frac{Qe^2}{2}} R_0 \quad : \Omega = \frac{2mc^2}{k} e_1 e_2 \quad (64)$$

The time dependence of the local coordinates is related to the angular velocity bi-vector via (65)<sup>[F]</sup>.

$$\left( \frac{d\mathbf{e}_\mu}{dt} = \pm \boldsymbol{\Omega} \cdot \mathbf{e}_\mu \right) \quad (65)$$

On the spin plane  $\mu = 1, 2$  resulting in (66), where  $t$  is relative to metric time.

$$\begin{aligned} \frac{de_1}{dt} &= \pm \frac{2mc^2}{(k)} e_1 e_2 e_1 = \pm \frac{2mc^2}{(k)} e_2 \\ \therefore \frac{de_2}{dt} &= \pm \frac{2mc^2}{(k)} e_1 e_2 e_2 = \pm \frac{2mc^2}{(k)} e_1 \end{aligned} \quad (66)$$

These equations define the time dependence of the local coordinate system attached to the electron's classical position. Since an electron moving at the speed of light violates relativity (19), intrinsic spin and zitterbewegung must be field related.

Comparing the Heisenberg approach to the geometric algebra derivation, the only self-consistent interpretation of the Dirac equation is mechanical in nature. The Heisenberg picture demonstrates that the position operator is following a complex, light-like trajectory. Geometric algebra demonstrates that the electron has a classical velocity ( $v$ ) and an attached coordinate system at the local position. In addition, a multivector rotation is an active transformation, which acts on a field independent of the reference coordinate system. The original space-like geodesics must therefore be modified so that the field always follows time-like geodesics. This in return allows the relativistic dynamics of a localized field to be reduced to a point in space-time.

The classical structure of electrons/positrons is required to further the localized field interpretation of the Dirac equation. Quantum theory hides the localized nature of particles through Lorentz transformations. Similar to general relativity, a localized field can be reduced to a point like object. Since the Dirac equation is Lorentz invariant, knowing the trajectory of any point along the field allows all others to be determined. Therefore, it is assumed that the active transformation applied to the spin-plane carries over to all other points along the field. According to classical electrodynamics, which is implied by quantum theory via minimal coupling, a moving electric field will produce a magnetic field equal to (67).

$$\bar{B} = \frac{E_0}{c^2} (\vec{v} \times \hat{E}) \quad (67)$$

The active transformation must be light-like acting on the entire field of the electron; i.e. the magnetic field becomes (68).

$$B = \frac{k_e q_e}{cr^2} \phi \quad (68)$$

The classic definition for a magnetic dipole is (69), where  $m$  is the dipole moment.

$$\bar{B}_{dipole} = \frac{k_m m}{r^3} (2 \cos(\varphi) \hat{r} + \sin(\varphi) \varphi^\wedge) \quad (69)$$

Since it is claimed that the Dirac equation is specific to a single point on the spin-plane, equations (68, 69) are combined resulting in (70).

$$\frac{k_e q_e}{c} = \frac{k_m m}{r} \quad (70)$$

Plugging in the classical radius (56) as defined in the Dirac equation results in a spin magnetic dipole moment equal to (71).

$$\frac{m}{\mu} = \frac{m}{c} = \frac{q_e}{e} \frac{cr}{r_e} = \frac{q_e k}{m_e} \quad (71)$$

Measuring the magnetic field along the spin-plane (at  $r_e$ ) results in a magnetic dipole moment that is twice the Bohr magneton (72). If the zitterbewegung radius is used instead, the dipole moment becomes equivalent to the Bohr magneton.

$$\mu_B = \frac{q_e k}{2m_e} \quad (72)$$

It is obvious that the field generated by a spinning electron is not a true magnetic dipole. This is irrelevant until the hyperfine structure, where the nucleus interacts with the electron's far-field. The relativistic electric field also deforms with respect to the appropriate Lorentz transformations, while vacuum energy density must also be included.

Ignoring higher order effects, the approximate evolution of a spin 1/2 quantum system can be described by relativistic field dynamics. Returning to the spin bi-vector, the equation can be expanded with the geometric product (73)<sup>[G]</sup>.

$$S = isv = i(s \cdot v + s A v) \quad (73)$$

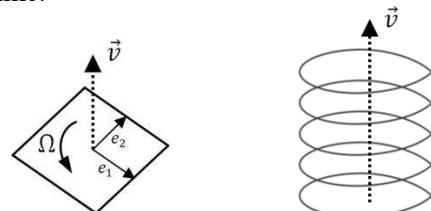
The Hodge dual (74) allows for an identity between the wedge product and cross product.

$$s A v = * (s \times v) = i(s \times v) \quad (74)$$

The spin bi-vector (75) is composed of a real scalar and imaginary vector, relative to the angle between the intrinsic spin and classical velocity.

$$S = isv[\cos(\beta) + i \sin(\beta)] \quad (75)$$

Figure 1.5 demonstrates how quantum mechanics reduces a field's relativistic spin to a single point in space-time.



**Figure 1.5.** For an electron moving in a straight line without external field,  $\beta = \pi/2$  to ensure the field's velocity remains co.

## 1.7. Applying Vacuum Fields

Vacuum field theory bridges the gap between general relativity and electrodynamics. It is argued within the previous sections that particles exist as localized field(s). This is contrary to Einstein's field equations, which depict particles as point-like sources. In perspective, EFEs take the quantized attributes of localized fields and couples them to a geometric field equation. The principle of locality is essential to Einstein's interpretation of general relativity, i.e. it allows point-like sources to have a local effect on the surrounding space-time metric. Quantum mechanics on the other hand demonstrates that local hidden variable theories are invalid via Bell's theorem. By applying vacuum field theory, the Dirac equation can be interpreted as a quantized field theory. The non-local hidden variables become visible once spacetime algebra is applied, which reveals the underlying relativistic field dynamics. Ignoring the significance of localized fields rather than point-like sources will result in theories that allow gravitational waves. For example, both EFEs and the Brans-Dicke theory allow these waves. Gravitational waves however have not been ruled out experimentally, although one has never been directly detected. Gravitational waves and the current probability of non-existence are further discussed throughout section (2.2).

The Lorentz transformation is a crucial aspect of quantum field theory or classical electrodynamics. Section (1.5) demonstrated that an accelerating charged particle emits vacuum energy density. This is related to the emission of electromagnetic energy in terms of bremsstrahlung. Therefore, application of locality in classical electrodynamics is valid because the electric potential is directly related to vacuum energy density. Application of point-like sources in quantum mechanics is also valid when considering the process of quantization. This allows

for hidden non-local variables in agreement with Bell's theorem and action at a distance. Quantum mechanics itself is based upon quantized variables such as rest mass, position and momentum. Spin only complicates the situation by offsetting the point of quantization from the classical position, i.e. the electric field will always travel at the speed of light. Satisfying the equations of motion for a point along the field will automatically solve all others.

It is difficult to define vacuum energy density in terms of the additional fields within QFT and the standard model. However, the electric potential is proportional to vacuum energy density with the far-field approximation. The mass of a single fermion is also directly proportional to an underlying scalar field. The energy density at the classical position of a particle is crucial in terms of the quantization process. It allows a localized entity to be reduced to a point-like object, where classical mechanics can be applied. Vacuum energy density can also exist without the presence of classical fields. Neutrons for example are massive compared to electrons, but demonstrate minimal electromagnetic properties. Although the charge of negative and positive quarks can cancel, the underlying vacuum energy density must be conserved. The unified field theory would therefore require classical quantized forces to be abandoned. This is not required to arrive at a theory that predicts the outcome of any experiment. For example, it is always possible to include additional factors into mathematical models for agreement with observations. However, this does not mean the resulting theory will depict what is actually taking place. It will also be difficult if not impossible to arrive at an exact formulation connecting general relativity and the standard model. With vacuum field theory, far-field approximations can instead be applied in order to arrive at a perturbative theory of everything.

The Dirac equation written in covariant form is defined as (76).

$$(\gamma^\mu \partial_\mu + i \frac{m_\varrho c}{k}) = 0 \quad (76)$$

In order to include general relativity or external fields into equation (81),  $\partial_\mu \rightarrow D_\mu$  is the covariant derivative (77) with respect to the local frame.

$$(i\gamma^a D_a - \frac{m_\varrho c}{k}) = 0 \quad (77)$$

The anholonomic Dirac matrices ( $\gamma^a$ ) are related to the Dirac matrices (78) by a vierbein field.

$$\gamma^a = e_\mu^a \gamma^\mu \quad (78)$$

The vierbein field is related to the metric tensor via equation (79).

$$g_{\mu\nu} = \frac{\partial E^a}{\partial x^\mu} \frac{\partial E^b}{\partial x^\nu} y_{ab} = e_\mu^a e_\nu^b y_{ab} \quad (79)$$

Equation (78) finds its origins from the commutator of the Dirac matrices (80).

$$[\gamma^a, \gamma^b] = 2y^{ab} \quad (80)$$

In order to relate the local tetrad frame to the metric, the general commutator is defined as (81).

$$[e^a \gamma^a, e^b \gamma^b] = 2g^{\mu\nu} \quad (81)$$

The local frame is attached to the electron's classical position and remains light-like. Therefore, the Dirac equation with respect to the preferred reference frame ( $\mu, v, \dots$ ) becomes (82).

$$(ie^\mu \gamma^\mu D_\mu - \frac{m_\varrho c}{k}) = 0 \quad (82)$$

The covariant derivative (83)<sup>[H]</sup> includes variations from the metric of space and any external electromagnetic field defined upon it.

$$D_\mu = \partial_\mu + \frac{1}{4} w_\mu^{ab} y_{ab} - \frac{ie}{kc} A_\mu \quad (83)$$

Since Planck units offer a natural scale for vacuum fields, the equations for the remainder of this section are written with  $k = c = G = 1$ . The metric relative to vacuum field theory in accordance with section (1.5) is isotropic, i.e. it must be defined by a single scalar field ( $\underline{g}$ ). Therefore, the vierbein defined in (78) is directly related to the effective vacuum energy density (84).

$$e^a = \gamma_g = 1 + \nabla \cdot \underline{g} \quad (84)$$

Neglecting any sub-structure of the nucleus or self-interactions, the static field is approximately (85).

$$\nabla_p \approx \frac{m_p}{r} \quad (85)$$

The Dirac equation with an external field is (86).

$$[i\gamma^\mu (\partial_\mu + \frac{1}{4} w_\mu^{ab} y_{ab} - ieA_\mu) - \frac{m_\varrho}{\gamma_g}] = 0 \quad (86)$$

In order to ensure that the field remains light-like, the spin connection (87)<sup>[H]</sup> must be introduced into the covariant derivative. The Christoffel symbols ( $\Gamma_{\sigma\mu}^\nu$ ) are further derived from vacuum energy density and resulting space-time metric.

$$w_\mu^{ab} = e_v^a \partial_\mu e^{vb} + e_v^a e^{\sigma b} \Gamma_{\sigma\mu}^\nu \quad (87)$$

The space-time indices can be raised or lowered by applying  $g^{\mu\nu}$  or  $g_{\mu\nu}$  respectively; i.e. (88).

$$e^{vb} = g^{\mu\nu} e_\mu^b = \frac{1}{2} [e_\mu^a \gamma^a, e_\nu^b \gamma^b] e_\nu^b \quad (88)$$

Solutions that are more exact require for the interaction between individually localized vacuum fields to be accounted for, i.e. the nucleus and electron. This usually involves solving continuous fractions by iteration discussed in section (2.6). Since the inverse distance is an approximation for the far-field, it is also necessary to apply an energy cut-off when the radius is 1 in Planck units. This ensures that vacuum energy density does not surpass the maximum value depicted by quantized or classical energy.

## 1.8. The Standard Model

The standard model is an extension of quantum field theory, which is based upon classical electrodynamics and special relativity. It includes several other fields such as the electroweak and Higgs, which model weak interactions and mass. Many of the previous principles from the Dirac equation carry over to quantum electrodynamics (QED) and the standard model. The Lagrangian density (89)<sup>[1]</sup> of QED for example consists of the Dirac equation and classical electromagnetic contributions.

$$L_{QED} = \bar{(iD - m)} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (89)$$

QED is formulated with classical fields coupled to spinning light-like manifolds. In this perspective, the mass term offers no additional insight beyond semi-classical Lagrangian dynamics. Due to this, extensions of quantum theory fail to explain the physical essence behind classical fields and mass. This however does not make the theory useless, as experiments can only measure quantized variables including rest mass, position and momentum. The electromagnetic field is also closely related to vacuum field theory and depicts time dependence of quantized charged particles.

The transition from QED to a more general theory requires the addition of neutral currents and weak interactions. The standard model unifies weak interactions and QED with electroweak theory, defined by the  $SU(2)_L \times U(1)_Y$  gauge group. Neutral vector bosons  $A_\mu$  (photon) and  $Z_\mu$  ( $Z^0$  mass eigenstates) are related to the fields responsible for weak interactions via the Weinberg angle (90)<sup>[1]</sup>. The coupling angle varies according to the energy scale or momentum transfer involved. It is also related to the running coupling constants of  $SU(2)_L$  and  $U(1)_Y$  ( $g$  and  $g'$  respectively).

$$\begin{bmatrix} A_\mu \\ Z_\mu \end{bmatrix} = \begin{bmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{bmatrix} \begin{bmatrix} B_\mu \\ W_\mu^3 \end{bmatrix} \quad (90)$$

The neutral components are included within the interaction Lagrangian density (91)<sup>[1]</sup>, along with the running coupling constants. In order to ensure that  $SU(2)_L$  invariance is not violated, a current ( $J_\mu^Y$ ) is added which preserves the symmetry.

$$L^{(W)}_{int} = -ig J^3 W^{3\mu} - i \frac{g'}{2} J^Y B^\mu \quad (91)$$

The  $SU(2)_L \times U(1)_Y$  gauge group (92)<sup>[1]</sup> shares similarities with the 2-dimensional rotors applied in space-time algebra.

$$\rightarrow_L^R = e^{i\theta_a \sigma^3} e^{i\theta_L^2} \quad (92)$$

The projections for left-handed and right-handed components are (93)<sup>[1]</sup>.

$$:_{L^2} = \frac{1}{2}(1 + \sigma^3) \quad :_{R^2} = \frac{1}{2}(1 - \sigma^3) \quad (93)$$

In canonical form, even sub-algebra solutions ( $R_{1,3}^+$ ) to the Dirac equation for  $\tilde{G} = 0$  are (94)<sup>[G]</sup>.

$$= \sqrt{\rho} e^{i\theta_a^2 R} \quad (94)$$

The connection between (92) and (94) arises because the group of 2D rotors and unitary group  $U(1)$  are locally isomorphic. However, there are two unique copies of  $U(1)$ :  $U(1)_{EM}$  and  $U(1)_Y$  with generators defined as (95)<sup>[1]</sup> respectively.

$$:Q = \frac{Y}{2} + I_3 \quad :Q = Y \quad (95)$$

The second component of (92) is locally isomorphic to  $SO(3)$ , although requires two complete rotations in order to return to the original state. The Pauli matrices define the axis of rotation, while  $\theta_a$  is the gauge parameter; i.e. the amount of rotation on each spin axis. Since the lie algebra of  $SU(2)$  and  $SO(3)$  are isomorphic, the general rotor in  $R^3$  is defined as (96). Thus  $i\sigma^a$  are the infinitesimal generators of  $SU(2)$ , similar to  $T_a$ .

$$R = \cos \frac{\theta}{2} + T_3 n \sin \frac{\theta}{2} \quad (96)$$

Returning to the conical solution of the Dirac equation, the beta factor encodes the angle between the spin-plane and velocity. Furthermore, the rotor determines the rotation of the field with respect to spin-coordinates. Combining these properties with the  $SU(2)_L \times U(1)_F$  gauge demonstrates the degrees of freedom for the underlying field. Either the spin-plane to velocity angle is varied, or an active 3-dimensional rotation is applied with respect to a rotational-axis. Similar to the Dirac equation, the process of quantization is crucial to understanding these transformations. It allows a localized field to be treated as a point-like object. Spin for example allows the quantization process to take place away from the classical position. This allows an active 3-dimensional transformation to be reduced to the motion of a single point in space-time.

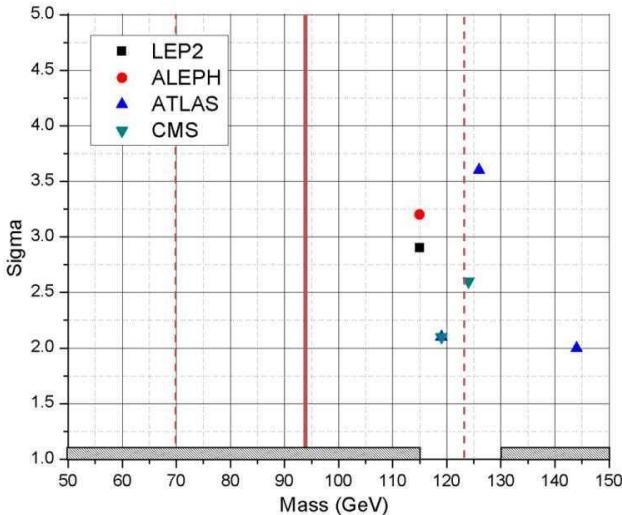
It should be no surprise that the standard model is not capable of becoming a unified field theory. Renormalization is the first of many signs that the choice of fields is poor, i.e. classical and/or multiple scalar-vector fields should not be applied. Beyond the initial choice of electromagnetism, the non-classical fields are a consequence of preserving symmetry and/or filling gaps. The standard model also requires a minimum of 19 free parameters and several running coupling constants determined experimentally<sup>[J]</sup>. Any unified field theory should have zero free variables except for the fundamental constants ( $c_0$ ,  $G_0$ ,  $\hbar$ ,  $k_e$ ), which depict underlying properties of space. For example, the electron rest mass should be determined from the only stable non-composite solution to the unified field theory. Quantized mass takes the particular value due to the non-linear nature of the field, i.e. there is only one stable value. The standard model in original form however does not attribute mass to fermions and other massive particles. An attempt to resolve this absence requires an additional field and resulting scalar particle, i.e. the Higgs field and boson respectively.

The Higgs field was introduced in 1962 by Philip Anderson to compensate for the lack of mass for gauge bosons within the standard model. The relativistic model was further developed in 1964 by independent groups who were awarded the Nobel Prize. The additional field predicted the existence of a Higgs boson, which gives mass to other particles. The mass of the Higgs boson can be theoretically determined from the mass of the top quark and W-boson. Earlier measurements of these particles predicted a Higgs boson with an expected mass of  $85^{+54}_{-34}$  GeV; however, recent world averaged values (March 2012)<sup>[K]</sup> of the top quark and W-boson vary the preferred value to  $94^{+29}_{-24}$  GeV. The standard model therefore predicts that the mass of the Higgs boson ranges from  $70$ – $123$  GeV with theoretical methods. These values are relative to the top quark having a bare mass of  $173.2 \pm 0.9$  GeV and W-boson of  $80.399 \pm 0.023$  GeV.

Direct methods of detecting the Higgs boson initially began at CERN with LEP2. Preliminary results from data collected over the year 2000 claimed that four LEP2 experiments were consistent to the 2.9 sigma level (“1.4 in 1000 chance of statistical fluctuation”) of a 115 GeV Higgs boson<sup>[L]</sup>. This was reported in November 2000, based upon an excess of events over the theoretical background rates. Results were published in December 2001 conserving the previously predicted mass, although the combined probability had decreased to  $2.4 \cdot 10^{-3}$ <sup>[M]</sup>. From the four individual experiments, ALEPH provided the most significant results. The excess of events over background was initially placed at  $3.4\sigma$ , which was further reduced to  $3.2\sigma$  in the final report. The more recent large hadron collider began operations in November 2009, which also contains multiple experiments for detecting the Higgs boson. Initial results were released in July 2011 with respect to ATLAS and CMS, showing an excess of events around 144 GeV<sup>[N]</sup>. This was compatible with a Higgs boson at the 2.9 sigma

level. An article published in nature around mid-august of the same year later revised the confidence to a sigma of 2.0<sup>[O]</sup>. Results that are more recent were published in December of 2011, where ATLAS had a signal at 126 GeV with 3.6 sigma; CMS showed an excess of events around 124 GeV with 2.6 sigma<sup>[P]</sup>. An additional weak signal was detected by both experiments around 119 GeV with a 2.1 sigma. Figure 1.6 depicts these recent results and theoretical mass as derived from the top quark and W-boson.

Between LEP2 and LHC, there seems to be a disagreement with the observed excess of events. The results of ALEPH demonstrated a sigma that is relatively close to ATLAS. If the Higgs boson exists in accordance to the standard model, both results cannot be correct. This raises concerns over the understanding of background processes and their contribution to excess events. There also lacks a single region that demonstrates a Higgs boson signal, i.e. mass has varied between individual experiments and runs. This of course can be explained by lack of data, both by the results of LEP2 and preliminary results at LHC. Relative to theoretical predictions, there exists a large margin



**Figure 1.6.** The thick red line is the theoretical Higgs mass determined from the top quark and W-boson; dashed lines are error boundaries as of March 2012 for the 68% confidence level<sup>[K]</sup>. The hatched area at the bottom depicts regions that have already been ruled out by other experiments.

between experimental results and the preferred mass range. Due to the uncertainty of the top quark's mass, recent claims of a Higgs-like particle between 4.5 to 5.0 sigma (125 GeV)<sup>[Q]</sup> are borderline acceptable up to the 68% confidence level<sup>[K]</sup>. However, only decay products of the Higgs boson are being directly detected, which coincide with the decay products of other known particles. Even at 6 sigma in agreement with the standard model, there lacks explanation for the physical essence of mass and additional deflection a particle experiences in external fields.

From vacuum field theory, the classical energy (97) of a single fermion is proportional to the point of maximum vacuum energy density. It is this central point of a quantized particle that depicts the kinematics of the entire localized field. This is in accordance with the process of quantization and affine transformations previously applied.

$$E = \sqrt{(pc)^2 + (m_0 c^2)^2} \quad (97)$$

It is also predicted that massive particles such as electrons will have symmetric scalar fields, while massless photons consist of anti-symmetric scalar fields. This symmetry allows massive particles to have a finite amount of scalar mass at the classical position. Massive particles will therefore require momentum to move through an external field due to a localized scalar field rather than a massless scalar-vector field. Mass in QED is not well defined because the Dirac equation uses it to quantize spin. This relativistic spin is in return balanced with the quantized velocity so that all points upon the spin plane move at the local speed of light. The unified field theory should instead reduce to a single scalar-vector field (98), where mass is a localized scalar field ( $\pm\emptyset$ ) depicting matter and anti-matter for each pair of fundamental particles.

$$\nabla(x) = \sqrt{\nabla^*} = \sqrt{(\emptyset)^2 + (v)^2} \quad (98)$$

## 2. Relativity and Differential Geometry

Vacuum field theory requires a single scalar field determined from quantized variables and affine transformations. From this scalar field, it is possible to define the space-time metric similar to Einstein's field equations. Identical mathematical tools are required for either field theory, i.e. differential geometry and Riemann manifolds are necessary for quantizing a field's motion. Vacuum field theory also explains the mechanism behind gravitational force. Variations in time dependence at each point in space forces a quantized field to accelerate. Similar to section (1.4), applying the space-time metric to the particle's manifold ensures the underlying field's time dependence resembles the initial  $y_{\mu\nu}$  configuration. Differential geometry and Riemann manifolds are therefore indispensable tools for general relativity, regardless of the underlying field theory.

A crucial modification to the theory of general relativity is the coupling between point-like sources and the corresponding space-time metric. If this coupling is poor, artifacts will appear under certain scenarios; i.e. gravitational waves and singularities. After discussing the correct metric from vacuum field theory, it becomes clear that EFEs are using the manifold of space-time in disguise of a localized field. The coupling between EFEs and Maxwell's equations is also poor, i.e. the contributions to the space-time metric are incorrect. The space-time metric should instead include all vacuum energy components, i.e. the electric and neutral fields. EFEs instead decouple these fields from quantized mass, depicting them as separate entities. The Schwarzschild solution in return cannot represent realistic objects since all massive particles display some electromagnetic component. Attempting to produce a proper field solution via the Einstein-Maxwell equations is also incorrect.

With respect to classical electrodynamics, a charged particle will have an effective electric field defined by (99). The velocity must be relative to the metric of space-time, or background vacuum energy density due to all other particles.

$$E^{\rightarrow} = q \frac{r^{\rightarrow}}{r^3 (\gamma^2 \cos^2(\theta) + \sin^2(\theta))^{3/2}} \quad (99)$$

The electromagnetic field does not easily transform to a particle's vacuum energy density. However, the electric potential is proportional to  $\nabla$  with respect to charged particles. Since the electric field lines are parallel with  $r^{\wedge}$ , the electric field is proportional to the radial derivative of the effective vacuum energy density (100); where  $\sigma$  is the charge to mass ratio and Planck units are applied.

$$E^{\rightarrow} = -\sigma \left( \frac{\partial}{\partial r} \nabla r^{\wedge} \right) \quad (100)$$

Assuming that the field is only compressed relative to the direction of motion, the effective vacuum field is (101).

$$\nabla = \frac{1}{r \sqrt{(\gamma \cos \theta)^2 + (\sin \theta)^2}} \quad (101)$$

The partial derivative of the transformed field (101) is therefore (102).

$$-\nabla = \frac{r^{\rightarrow}}{r^3 (\gamma^2 \cos^2(\theta) + \sin^2(\theta))^{3/2}} \quad (102)$$

Equation (100) is equivalent to the effective electric field of a moving charged particle (103).

$$-\sigma \left( \frac{\partial}{\partial r} \nabla \right) r^{\wedge} = q \frac{r^{\wedge}}{r^2 (\gamma^2 \cos^2(\theta) + \sin^2(\theta))^{3/2}} \quad (103)$$

Vacuum energy density is therefore directly related to the electric potential of a non-composite massive particle. This connection allows localized fields to be treated as point-like sources in electrodynamics.

## 2.1. General Relativity

The foundational aspects of general relativity discussed over the previous sections are sufficient for understanding the coupling between EFEs and classical electrodynamics. The Maxwell-Einstein field equations are introduced (104)<sup>[R]</sup>, where  $T^{\mu\nu}$  is the electromagnetic stress-energy tensor.

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{2G_0}{c^2\mu_0}T^{\mu\nu} \quad (104)$$

The Reissner-Nordstrom metric (105) is a static solution to the Maxwell-Einstein field equations. It defines the gravitational field around a charged, non-rotating, spherically symmetric object.

$$g_{\mu\nu} = \begin{pmatrix} \gamma'^{-2} & 0 & 0 & 0 \\ 0 & \gamma'^2 & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & r^2 \sin^2 \theta \end{pmatrix} \quad (105)$$

Where  $\gamma'$  is (106)<sup>[S]</sup> and the Schwarzschild radius is defined as  $r_c = 2E/\Delta$ .

$$\gamma'^2 = \frac{1}{1 - \frac{r_c}{r} + \left(\frac{k_e Q}{r \Delta}\right)^2} \quad (106)$$

This is the limit as  $\Delta \rightarrow \Delta$  relative to the vacuum field definition (107), after the proper electric field contribution (18, 100) is taken into consideration.

$$\gamma_g = \frac{1}{(1 - \frac{\nabla}{\Delta})^2} = \frac{1}{1 - \frac{r_c}{r} + \left(\frac{\nabla}{\Delta}\right)^2} \quad (107)$$

It is no coincidence that equations (106, 107) have similar form. Vacuum field theory unifies mass and external fields into a single entity. A simple case would be that of an electron, which is spherically symmetric with respect to any co-moving frame. According to the second axiom, the conservation of vacuum energy creates an  $E/r$  far-field via the linear wave-equation (1). Since the electromagnetic field is based upon classical dynamics, it should not

contribute to vacuum energy density with respect to a charged particle; i.e. the EM components are already included in (24). The Reissner-Nordstrom metric for a single particle therefore reduces to the derived vacuum far-field of a charged particle after proper field contributions are considered.

Returning to electromagnetic fields, quantized conservation laws can be written as (108)<sup>[B]</sup>. The electromagnetic stress-energy tensor ( $T^{\mu\nu}$ ) is once again related to classical mechanics.

$$\partial_\nu T^{\mu\nu} + y^\mu/f_\nu = 0 \quad (108)$$

The field tensor (109)<sup>[B]</sup> is derived from the electromagnetic four-potential, it is therefore related to quantized lagrangian dynamics and should not contribute to the vacuum energy density. Once again, there is a distinction between vacuum energy density and quantized energy due to classical force.

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \quad (109)$$

The four-potential (110)<sup>[B]</sup> is proportional to the scalar and vector potentials.

$$A' = \left( \frac{\phi}{c_0}, A \right) \quad (110)$$

Einstein's field equations apply quantized densities in order to determine the curvature and metric of space-time. The quantized source term on the right side of EFEs (111) allows for mass and momentum to be applied in a continuum limit. The left side couples the quantized source(s) to a local geometric field or space-time metric.

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G_0}{c_0^4}T^{\mu\nu} \quad (111)$$

With the foundations offered from vacuum field theory, it is clear that Einstein's field equations contain several flaws. Up to this point, the three reasons initially offered for a revision of general relativity have been discussed.

The first was the coupling of point-like sources to the space-time metric. The metric further mimics the underlying vacuum field from which matter originates. Similar to electrodynamics, the coupling of a localized field to a point-like source allows for waves upon the relative medium; i.e. the classical electromagnetic field or space-time metric of general relativity. The vacuum field of a single non-composite particle however is proportional to its electric potential. When the electric potential of a particle varies relative to some background field, so does the underlying vacuum energy. This was previously discussed relative to bremsstrahlung, where this variation was directly related to the variation in vacuum energy density and Lorentz transformation. Since the space-time metric is defined solely by vacuum energy density, it is incorrect to include the electric field of a single particle as a separate entity.

The second flaw is similar to the first, where particles must exist as localized fields rather than point-like objects. Similar to relativistic electrodynamics, the field of a particle becomes deformed when the background vacuum energy density varies. Therefore, the metric of space-time is of mathematical origin and plays no role in the physical structure of space. An observer will always view space relative to the metric of space-time.

However, space does not deform at the large-scale structure as purposed by Einstein. This also pertains to gravitational waves, as it is claimed by the third postulate that matter is Planck scale fluctuations of space. It is contradictory to allow space-time waves from the quantized variables of vacuum fields, when vacuum fields are a representation of Planck scale waves. Furthermore, if vacuum fields were responsible for classical force, objects with event horizon would lack external gravitational fields; this is contradictory to EFEs vacuum solutions.

The third flaw is the incorrect coupling between electrodynamics and classical gravitational field. From section (2), it was argued that the electric field is a component of the underlying vacuum field. With the conservation of vacuum energy density or classical theory, a non-composite charged particle should display an  $E_o/r$  far-field. The Maxwell-Einstein vacuum solution ( $T_{\mu\nu} = 0$ ) gives a squared Lorentz scalar of (112).

$$\gamma'^g{}^2 = \frac{1}{1 - 2 \frac{\nabla}{\Delta} + (\frac{r}{\Delta})} \quad (112)$$

For a single particle,  $T_{\mu\nu} = 0$  everywhere except at the particle's classical position. Therefore, (112) is essentially the far-field solution of a single charged particle. However, the Lorentz scalar does not obey the  $E_o/r$  law as derived under the assumption of vacuum energy conservation. According to vacuum field theory, the correct squared Lorentz scalar for a single particle reduces to (113).

$$\gamma_g{}^2 = \frac{1}{1 - 2 \frac{\nabla}{\Delta} + (\frac{\nabla}{\Delta})^2} = (1 + \frac{\nabla}{\Delta})^2 \quad (113)$$

(113) is the limit of equation (112) as  $\Delta \rightarrow \Delta$  and reduces to the Maxwell-Einstein vacuum solution after proper coupling of the electromagnetic field via equations (18, 100).

$$\gamma_g{}^2 = \frac{1}{1 - 2 \frac{\nabla}{\Delta} + (\frac{\nabla}{\Delta})^2} = \frac{1}{(1 - \frac{\nabla}{\Delta})^2} \quad (114)$$

The limit produces a singularity when  $\nabla = \Delta$ , while the proper vacuum field solution (115) is always finite. Without proper electromagnetic contributions, the singularity instead arises when  $2 \nabla = \Delta$  as demonstrated by the Schwarzschild solution.

$$\gamma_g{}^2 = \frac{1}{1 - 2 \frac{\nabla}{\Delta \gamma_g} + (\frac{\nabla}{\Delta \gamma_g})^2} = (1 + \frac{\nabla}{\Delta})^2 \quad (115)$$

## 2.2. Gravitational Waves

The existence of gravitational waves is crucial to the validity of EFEs. From a theoretical standpoint, variations in certain stress-energy moments allow quantized variables to transform into metric waves. If these waves do not exist, then the conservation of energy is clearly violated within EFEs. Indirect evidence for gravitational waves comes from the binary system PSR B1913+16. This system consists of orbiting neutron stars that emit pulsed radio signals at nearly constant periods. After measuring these pulses, their arrival was observed to oscillate over a period of about 7.75 hours. Additional observations allowed for the change in epoch of periastron to be measured, which agreed with the predictions of general relativity to within 0.2%<sup>[T]</sup>. However, it is the assumption that EFEs are correct which defines the plausible attributes of the system. It is possible for PSR B1913+16 to have parameters that vary from EFE solutions, i.e. these observations only provide indirect evidence for the existence of gravitational waves. Complications arise from uncertainties in the structure of neutron stars, their effective field, orbital parameters and classical energy flux.

Direct evidence by physically measuring the distortion due to gravitational waves appears to be the only valid option for proving their existence. Several experiments have been conducted over the previous 52 years; however, only LIGO, GEO600 and VIRGO are discussed due to precession. The probability of detecting a gravitational wave from a BH-BH event is approximated from table 2.2.

TABLE 2.2. Theoretical BH-BH detection rates

ID	Source	Published	BH-BH Detection Rates ( $\text{yr}^{-1}$ )
A	[U]	1999	0.19 <sup>(b)</sup> , 400 <sup>(a)(b)</sup>
B	[V]	2007	2, 500 <sup>(b)</sup>
C	[W]	2010	[ $2 \cdot 10^{-4}$ , 0.5] <sup>a</sup>
D	[X]	2010	[0.05, 4.9]
E	[Y]	2008	[0.01, 1.7], [0.11, 18] <sup>(a)</sup>
F	[Z]	2011	0.8
G	[AA]	2009	-
H	[AB]	2004	-
I	[AC]	2012	[0.08, 0.17] <sup>(a)</sup>

Note: (a) indicates LIGO IE and (b) are older models.

The theoretical event rates are required for determining the probability of gravitational waves existing. They have varied drastically over the previous 15 years as demonstrated by table 2.3. The running length of each experiment is also provided in table 2.4.

TABLE 2.3. Theoretical event rates

ID	Experiment	Event Rate ( $\text{yr}^{-1}$ )		
		NS-NS	NS-BH	BH-BH
A*	LIGO I	0.03	0.25	0.19
	LIGO IE	50	400	400
B*	LIGO I	100	30	500
	LIGO I	20	1	2
	LIGO I	[ $2 \cdot 10^{-4}$ , 0.2]	[ $7 \cdot 10^{-5}$ , 0.1]	[ $2 \cdot 10^{-4}$ , 0.5]
	LIGO II	[0.4, 400]	[0.2, 300]	[0.4, 1000]
	LIGO I	0.01	0.02	4.9
	LIGO II	45.1	85.8	21,425
DII	LIGO I	0.002	0.01	0.05
	LIGO II	9.5	42.8	242
	LIGO I	[0.015, 0.15]	-	[0.01, 1.7]
	LIGO IE	[0.15, 1.5]	-	[0.11, 18]
	LIGO II	[20, 200]	-	[16, 270]
	LIGO I	0.05	0.02	0.8
	LIGO II	[60, 500]	80	2,000
	Virgo I	[0.002, 0.04]	-	-
	LIGO IE	[0.02, 0.4]	-	-
	Virgo+	[0.25, 5]	-	-
	LIGO I	[0.008, 0.13]	-	-
	LIGO II	[40.2, 310.9]	-	-
	Virgo+	0.003	[0.01, 0.02]	[0.07, 0.08]
I	Virgo II	[3.0, 3.6]	[12, 19]	[35, 92]
	LIGO S5	0.004	0.02	[0.08, 0.09]
	LIGO S6	[0.008, 0.009]	[0.03, 0.04]	[0.17, 0.21]

\* indicates older theoretical models.

TABLE 2.4. Experiment runtime

Experiment	Run	Days	Run-Time	NS-NS Range
LIGO I	S1	4.8		0.08 Mpc
	S2	18.0	54.6 <sup>[AD]</sup>	0.3 Mpc
	S3	13.2		5.0 Mpc
	S4	18.6		8.6 Mpc
LIGO IE	S5	365	730 <sup>[AD][AE]</sup>	12.0 Mpc
	S6	365		~50 Mpc
Virgo I	VSR1	111	111 <sup>[AF]</sup>	12.4 Mpc
	VSR2	98		16.8 Mpc
Virgo+	VSR3	61.3	159.3 <sup>[AG]</sup>	~50 Mpc
GEO600	S4	28.8	370.8 <sup>[AH]</sup>	~LIGO I (S1)
	S5	342+		~LIGO I (S2)

Note: The more recent runs of GEO600 are included since the detection rates are similar to the earlier LIGO I runs.

The probability of gravitational waves not existing is compared to flipping a loaded coin. When a normal coin is flipped, the probability of it landing tails is 50%. This is equivalent to measuring for gravitational waves over the period required for a single event, and having a 50% chance of detecting one. For a loaded coin, the result will always be tails regardless of N. As the coin is flipped N amount of times, the probability of the coin being loaded increases if the results are always tails. Therefore, the probability that the coin is loaded is equivalent to that of gravitational waves not existing. This probability is defined as (116), where N is the expected events per total period.

$$P(\text{Null}) = 1 - 2^{-N} \quad (116)$$

Table 2.5 depicts the probability of gravitational waves not existing for each experiment and model.

**TABLE 2.5. Theoretical detection rates by experiment**

Experiment	Model	Max Events	<i>P(Null)</i>	
			Min	Max
LIGO I	<b>A*</b>	0.07	-	4.76%
	<b>B*</b>	94	-	~100.00%
	<b>B</b>	3.4	-	90.8%
	<b>C</b>	0.12	0.0049%	7.96%
	<b>D<sub>I</sub></b>	0.74	-	40.0%
	<b>D<sub>II</sub></b>	0.009	-	0.64%
	<b>F</b>	0.13	-	8.63%
LIGO IE	<b>A*</b>	1700	-	~100.00%
S5		0.114	6.96%	7.60%
S6	<b>I</b>	0.259	13.43%	16.4%
Virgo+	<b>I</b>	0.045	2.48%	3.07%
GEO600 (S4/S5)	<b>A<sup>†</sup></b>	0.48	-	28.2%
	<b>B<sup>*†</sup></b>	640	-	~100.00%
	<b>B<sup>†</sup></b>	23.4	-	~100.00%
	<b>C<sup>†</sup></b>	0.81	0.033%	43.1%
	<b>D<sub>I</sub><sup>†</sup></b>	5.01	-	96.9%
	<b>D<sub>II</sub><sup>†</sup></b>	0.063	-	4.27%
	<b>F<sup>†</sup></b>	0.88	-	45.8%

\* indicates old model

† indicates LIGO I statistics were applied to the latest runs of GEO600.

There exist several orders of magnitude between individual models. The majority of this variation is due to the merger rate and density of BH-BH events. Excluding the older models, the remaining models are grouped together in table 2.6. With respect to old models, it is clear from the number of events expected that gravitational waves could not exist ( $\sigma > 6$ ). The new models decrease expected rates by three to four orders of magnitude. However, two of the new models also indicate a  $\sigma > 6$  for the max event limit, with a third at  $\sigma = 3.7$ . The current data does not allow for a definitive answer for whether gravitational waves exist, although it does begin to raise doubts. The expected rates are also highly dependent upon the theoretical model. Assuming that these do not vary drastically in the future, the next generation of detectors should be capable of bringing all models to  $\sigma > 6$ . For example, conservative estimates of Advanced LIGO (LIGO II) project hundreds of events per year. Advanced LIGO and Advanced Virgo are expected to begin operations in 2014, so direct proof will require three to four years as of 2013. Beyond these, the Laser Interferometer Space Antenna (LISA) will provide a new range of frequencies. LISA is not expected to be in operation until after 2020; however, it will be capable of detecting massive BH-BH events if gravitational waves exist. Due to the net BH-BH mass, the waves generated would be several orders of magnitude larger than other sources.

**TABLE 2.6. Approximate combined probability**

Models	Net Max Events	Net <i>P(Null)</i>		Sigma Max
		Min	Max	
<b>B,I,I,B<sup>†</sup></b>	<b>27.2</b>	-	~100.00%	> 6
<b>C,I,I,C<sup>†</sup></b>	<b>1.35</b>	27.1%	60.8%	<b>1.7</b>
<b>D,I,I,D<sup>†</sup></b>	<b>6.17</b>	28.8%	98.6%	<b>3.7</b>
<b>E,D<sub>II</sub>,I,E<sup>†</sup></b>	<b>41.3</b>	37.0%	~100.00%	> 6
<b>F,I,I,F<sup>†</sup></b>	<b>1.43</b>	-	62.9%	<b>1.8</b>

Note: The order of listed models is LIGO I, LIGO IE, VIRGO+, GEO600; model E is also included.

### 2.3. Single Particle Metric

For a massive non-composite particle, the field should be approximately  $(E_0/r)$  in the co-moving frame. Since the metric must be isotropic, the solution in spherical coordinates is (117).

$$g_{\mu\nu} = \begin{bmatrix} \gamma^{-2} & 0 & 0 & 0 \\ F^g_0 & \gamma^2 & 0 & 0 \\ 0 & 0 & (\gamma g r)^2 & 0 \\ 1 & 0 & 0 & (\gamma g r)^2 \sin^2 \theta \end{bmatrix} \quad (117)$$

The Lorentz scalar (118) is defined as usual, where the far-field approximation is applied.

$$\gamma_g = 1 + \frac{G_0 M}{r c^2} \approx 1 + \frac{G_0 M}{r c^2} \quad (118)$$

For comparison, the Lorentz scalar relative to the Schwarzschild metric is written as (119).

$$\gamma'_g = \frac{1}{\sqrt{1 - 2 \frac{G_0 M}{r c^2}}} \quad (119)$$

The geodesic equations are applied to determine the motion of a single particle inside the previous potential. Using space-like convention  $(- + + +)$ , the space-time interval is defined as (120).

$$(ds)^2 = g_{\mu\nu} dx^\mu dx^\nu \quad (120)$$

Expanding the geodesic equations results in the proper acceleration for each component (121),  $(x^0)^2$  must be replaced with  $-(c_0 t')^2$ .

$$\begin{aligned} t'' &= -\frac{1}{2} \Gamma_{00}^{01} t'^2 - \Gamma_{10}^{01} r'^2 - \Gamma_{20}^{01} \theta'^2 - \Gamma_{30}^{01} \varphi'^2 \\ &\quad - 2 \Gamma_{12}^{01} \theta' r' - 2 \Gamma_{23}^{01} \theta' \varphi' \\ &\quad - 2 \Gamma_{13}^{01} r' \varphi' - 2 \Gamma_{23}^{01} \theta' \varphi' \end{aligned} \quad (121)$$

The geodesic equations become relative to the metric after applying the following relations (122).

$$\frac{dx'}{dr} = \gamma \frac{dx'}{dt'} \quad : \frac{d^2x'}{dr^2} = \gamma^3 \frac{d^2x'}{dt'^2} \quad (122)$$

The Christoffel symbols are determined for an isotropic, spherical vacuum field in table 2.7. To compare the radial acceleration as derived from the geodesic equations to the algebraic results (29), the motion of a particle can be restricted to the radial direction (123); i.e.  $\theta'$  and  $\varphi'$  are 0.

$$\begin{aligned} t'' &= -\frac{2 G_0 M}{r c_0^2} \frac{t'^2 r'}{g} \\ r''' &= -\frac{G_0 M}{r'^2} t'^2 + \frac{G_0 M}{r'^2} \frac{(rc_0)^2}{g} \end{aligned} \quad (123)$$

For a massive particle moving at escape velocity relative to the metric,  $r''$  is replaced by (27);  $t'$  is defined as (124), where  $k = 1$  and  $( = (g$ .

$$t' = k \gamma^2 \quad (124)$$

The acceleration at escape velocity is (125) or (29).

$$a = \frac{d^2 r'}{dt'^2} = -\frac{G_0 M}{r^2 (g^4)} \quad (125)$$

TABLE 2.7. Christoffel symbols between theories

$\Gamma_{\mu\nu}^{\sigma}$	Single Particle	Schwarzschild
$\Gamma_{01}^0, \Gamma_{10}^0$	$\frac{G_0 M}{\gamma_g (r c_0)^2} \cong \frac{1}{r} \frac{\nabla}{\Delta}$	$\gamma'^2 \frac{G_0 M}{(r c_0)^2}$
$\Gamma^{100}$	$-\frac{G_0 M}{\gamma^5 (r c_0)^2} \cong -\frac{1}{r \gamma^4} \frac{\nabla}{\Delta}$	$-\frac{G_0 M}{\gamma'^2 (r c_0)^2}$
$\Gamma_{11}^1$	$-\frac{G_0 M}{\gamma_g (r c_0)^2} \cong -\frac{1}{r} \frac{\nabla}{\Delta}$	$-\gamma'^2 \frac{G_0 M}{(r c_0)^2}$
$\Gamma_{12}^2, \Gamma_{21}^2$	$\frac{1}{\gamma_g r}$	$\frac{1}{r}$
$\Gamma_{22}^1$	$-\frac{1}{\gamma_g}$	$-\frac{1}{\gamma'^2}$
$\Gamma_{13}^3, \Gamma_{31}^3$	$1$	$1$
$\Gamma_{33}^1$	$-\frac{r}{\gamma_g} \sin^2 \theta$	$-\frac{r}{\gamma'^2} \sin^2 \theta$
$\Gamma_{23}^3, \Gamma_{32}^3$	$\cot \theta$	$\cot \theta$
$\Gamma_{33}^2$	$-\sin \theta \cos \theta$	$-\sin \theta \cos \theta$

Note: Single particle refers to massive and non-composite.

## 2.4. Gravitational Force and Potential

From sections (1.3, 2.3), it was observed that the gravitational potential is dependent upon a variation in the effective background vacuum energy density. For static fields, classical energy conservation does not depend upon the path taken between two points. The proper force (126) upon a particle is therefore similar to the relativistic Newtonian perspective. If a particle's geodesic path follows the gradient of  $\bar{\nabla}$ , the problem is always reduced to the covariant derivative of the scalar field; i.e. the vector  $\bar{g} \rightarrow_1$  will point in the direction of  $A\bar{\nabla}$ .

$$\bar{\mathbf{f}} = -A \frac{E_o}{\bar{A}} \bar{g} \rightarrow_1 U = -A \frac{E_o}{\bar{A}} \bar{A} \bar{\nabla} = -\frac{E_o}{\bar{A}} A \bar{\nabla} \quad (126)$$

The gravitational potential energy ( $U$ ) is derived by

equating  $\gamma_G = \gamma$  as applied in section (1.3); the potential is therefore defined as (127).

$$\Phi = -c \sigma^2 \frac{\bar{\nabla}}{\bar{A}} \quad (127)$$

Vacuum field theory requires that the metric of space-time is isotropic, while EFEs are anisotropic. The choice of isotropy or anisotropy is crucial for motion as observed by a distant observer. This is not locally detectable since any observer ( $g^{\mu\nu}$ ) deforms with respect to the space-time metric ( $g_{\mu\nu}$ ). Under the assumption that vacuum energy density is conserved, the vacuum far-field energy of a non-composite particle should at most remain constant or decrease when moving through an external field.

Conservation of classical energy does not always ensure conservation of vacuum energy density. The problem is complicated since actual particles will consist of an infinite amount of vacuum energy. However, it is possible to compare configurations of the vacuum field to determine if vacuum energy density is conserved. In this case, only the far-field is applied for simplicity. The conservation of vacuum energy density is discussed relative to the

transformed vacuum fields of an anisotropic and isotropic metric. Since a realistic gravitational field will complicate the field dynamics, a Cartesian coordinate system will instead be applied so that  $(r, \theta, \phi) \rightarrow (x, y, z)$ . The potential along the x-axis is anisotropic in the sense that it only increases in a single direction ( $x^\wedge$ ), similar to the Schwarzschild metric ( $\hat{t}$ ). It is difficult to consider a test particle initially placed at an infinite distance from the field. The problem is simplified by introducing an artificial potential as depicted in figure 2.1.

A massive particle placed where the local field is zero ( $y_{\mu\nu}$ ) has a far-field approximated by (128);

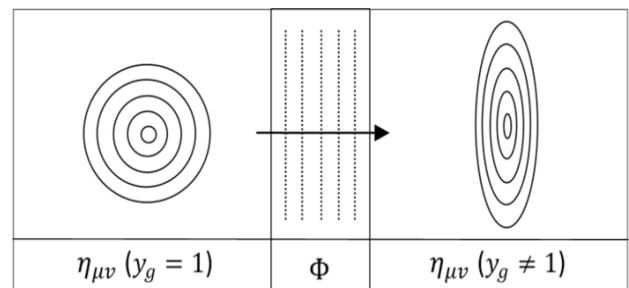
this is equivalent to  $r \rightarrow \infty$  for a spherical metric.

$$\bar{\nabla}(r) = \frac{E_o}{n} \frac{1}{r_n} \quad (128)$$

As the test particle moves into the gravitational potential, it gains quantized energy. This variation in energy is due to an influx of vacuum energy density rather than a net increase (129).

$$\bar{\nabla}(r, \theta) = \frac{E}{r_n} \frac{1}{\sqrt{(\gamma \cos \theta)^2 + (\sin \theta)^2}} \quad (129)$$

This means that the particle has gained quantized energy relative to a stationary observer; however, relative to  $y_{\mu\nu}$  vacuum energy must be conserved. Since the anisotropic transformation is similar to the special relativistic version, it can be directly substituted into (129).



**Figure 2.1.** Sections of a localized field are illustrated relative to an initial and final state. The fields relative to the preferred reference frame ( $\eta_{\mu\nu}$ ) are described by the space-time metric and Lorentz transformation.

The resulting effective field (130) is once again relative to  $y_{\mu\nu}$  or the preferred reference frame.

$$\Psi(r, \theta) = \frac{E_0}{r_n \sqrt{\frac{z^2}{(\gamma \gamma_g \cos \theta)^2 + \sin^2 \theta}}}$$
(130)

Since the test particle will be traveling at escape velocity,  $( = g$  can be applied. After reorienting the coordinate system, the Cartesian equivalent is defined as (131) with motion along the z-axis.

$$\Psi(x) = \frac{E_0}{\sqrt{\frac{x^2}{(\gamma)^2} + \frac{y^2}{(\gamma)^2} + (yz)^2}}$$
(131)

Equation (131) can further be put in spherical coordinates (132), allowing a shells net energy to be compared between configurations.

$$\bar{\Psi} = \frac{E_0}{r_n} \left[ \left( \frac{\sin(\varphi)}{\gamma} \right)^2 + (\gamma \cos(\varphi))^2 \right]^{-1/2}$$
(132)

Volumetric integration between two radii results in shell energy for the anisotropic case (133),

$$\bar{\Psi}_{\text{shell}} \Big|_A^B = \frac{\pi \gamma (A^2 - B^2)}{\alpha} [LN(\gamma^2 - \alpha) - LN(\gamma^2 + \alpha)]$$
(133)

where  $\alpha$  is defined as (134).

$$\alpha = \sqrt{\gamma^4 - 1}$$
(134)

The original field configuration (129) reduces to (135) after solving for shell energy.

$$\Psi_{\text{shell}} \Big|_A^B = 2\pi(B^2 - A^2)$$
(135)

For comparison, the variation in field energy for any shell can be defined with a factor between the two configurations. Therefore, the original field is multiplied by a function of the Lorentz scalar, defined as (136) for the anisotropic case.

$$k = \frac{1}{2\alpha} [LN(\gamma^2 + \alpha) - LN(\gamma^2 - \alpha)]$$
(136)

For an isotropic metric with identical setup, the field relative to the reference frame ( $y_{\mu\nu}$ ) is defined by (137).

$$\Psi(r, \theta) = \frac{E_0}{r_n \sqrt{\frac{(\gamma \cos \theta)^2}{(\gamma)^2} + (\sin \theta)^2}}$$
(137)

At escape velocity, the same ( $= g$  relation can be applied resulting in (138).

$$\Psi(x) = \frac{E_0}{\sqrt{(x)^2 + (y)^2 + (yz)^2}}$$
(138)

After integration, the shell energy for the isotropic metric is (139).

$$\bar{\Psi}_{\text{shell}} \Big|_A^B = \frac{\pi(A^2 - B^2)}{\sigma} [LN(\gamma + \sigma) - LN(\gamma - \sigma)]$$
(139)

Where  $\sigma$  is defined as (140).

$$\sigma = \sqrt{\gamma^2 - 1}$$
(140)

The factor between the initial configuration (135) and isotropic case is therefore (141).

$$k = \frac{1}{2\sigma} [LN(\gamma - \sigma) - LN(\gamma + \sigma)]$$
(141)

Considering vacuum energy density conservation, the isotropic and anisotropic cases are compared. As the far-field is rearranging, the peak energy of the field must also increase proportional to the classical quantized energy. There must be an influx of field energy into the central region, which is assumed smooth and finite. For field energy to be conserved, the far-field energy must be equivalent to or less than the original configuration. Plotting both (136, 141) shows that only isotropic metrics are valid, i.e. the anisotropic case initially increases far-field energy for  $( \cong [1, 2.7]$ . When  $k = 1$  (or  $( \cong 2.7015$ ), the anisotropic metric begins to lose far-field vacuum energy density. This decrease must however occur at  $( = 1$  as demonstrated by the isotropic metric. Anisotropic space-time metrics therefore cannot conserve vacuum energy density.

## 2.5. Arbitrary Space-Time Metric

In order to determine the space-time metric from quantized fields, vacuum field theory must be applied. A coordinate basis is chosen so that  $g_{\mu\nu}$  is diagonal with each term proportional to the general Lorentz scalar ( $\mathcal{G}_g$ ). From the gradient of  $\nabla(x)$ , the fundamental coordinate line is defined at each point in space with the base vector  $\hat{e}_1$ . If the gradient is zero, then the space is locally flat ( $y_{\mu\nu}$ ).

For example, if  $\nabla \cong E_o/r$  then each point would have the fundamental coordinate line parallel to the radial direction ( $\hat{r}$ ). The other base vectors would then be represented by any two orthogonal vectors perpendicular to  $\hat{e}_1$ . For other cases, the remaining orthogonal base vectors must be determined by the following methods.

Depending on choice of coordinates the field perpendicular to  $\hat{e}_1$  can vary, requiring that at least one other principle direction exists. Each aligns with the planes of maximum and minimum curvature relative to constant  $\nabla$  surfaces. The additional bases are therefore eigenvectors of the shape operator (142), which is defined by the first and second fundamental forms.

1

$$W = I^{-1}II = \frac{LG - MF}{EG - F^2} \begin{bmatrix} MG - NF & \\ ME - LF & NE - MF \end{bmatrix} \quad (142)$$

The first and second fundamental forms relative to the tangent plane of each surface  $R^2 \in R^3$  is (143).

$$:I = \begin{bmatrix} E & F \\ F & G \end{bmatrix} \quad :II = \begin{bmatrix} L & M \\ M & N \end{bmatrix} \quad (143)$$

If two unique principle directions exist beyond the initial gradient, vectors tangent to coordinate lines are ( $r_u, r_v$ ). Since the first base vector's ( $\hat{e}_1$ ) direction is determined by the normalized gradient of  $\nabla$ ,  $n$  is the normal to the surface (144)<sup>[AF]</sup>.

$$\mathbf{n} = \frac{A\nabla}{A\nabla} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\mathbf{r}_u \times \mathbf{r}_v} \quad (144)$$

From this notation, the two vectors ( $r_u, r_v$ ) are not necessarily of unit length. Vacuum field theory is however locally isotropic, i.e. the space-time metric is determined from a single scalar field or vacuum energy density. The previous vectors are therefore directly proportional to the curvilinear basis, i.e.  $r_a = (\mathcal{G}_g e_a)$ . The first fundamental form is also equivalent to the metric tensor, i.e. each component is the scalar product ( $g_{\mu\nu} = g_\mu \cdot g_\nu$ ).

The variables become  $E = r_u \cdot r_u$ ,  $F = r_u \cdot r_v$ ,  $G = r_v \cdot r_v$ . This can be extended to three dimensions, where the first fundamental form (145)<sup>[AF]</sup> is represented in quadratic form.

$$I(dx, dy, dz) = \begin{bmatrix} dx & & & \\ & dy & & \\ & & dz & \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} \quad (145)$$

The components ( $A_{\mu\nu} = r_\mu \cdot r_\nu$ ) can be defined in terms of partial derivatives of  $\nabla$  (146)<sup>[AF]</sup>; where ( $dx \rightarrow dx^1$ ), ( $dy \rightarrow dx^2$ ), ( $dz \rightarrow dx^3$ ) and partial derivatives are written as  $\nabla_i \nabla_j = i,jk$ .

$$\begin{aligned} :A_{11} &= {}_{2,31} - {}_{3,21} & :A_{22} &= {}_{3,12} - {}_{1,32} \\ && :A_{33} &= {}_{1,23} - {}_{2,13} \\ :A_{21} &= A_{12} = (1/2)({}_{3,11} - {}_{1,31} + {}_{2,32} - {}_{3,22}) & (146) \\ :A_{31} &= A_{13} = (1/2)({}_{2,33} - {}_{3,23} + {}_{1,21} - {}_{2,11}) \\ && :A_{32} &= A_{23} = (1/2)({}_{1,22} - {}_{2,12} + {}_{3,13} - {}_{1,33}) \end{aligned}$$

The second fundamental form is now introduced, which when combined with the first allows the other principle directions to be determined. From the two dimensional II (143), the variables can be defined as  $L = r_{uu} \cdot n$ ,  $M = r_{uv} \cdot n$ ,  $N = r_{vv} \cdot n$ , or as a tensor  $B_{\mu\nu} = r_{\mu\nu} \cdot n$ . Using the following method, the second fundamental form can be written in terms of partial derivatives of the vacuum energy density scalar field (147).

$$II(dx, dy, dz) = \nabla_x dx + \nabla_y dy + \nabla_z dz \quad (147)$$

For an implicit surface where  $\nabla(x, y, z) = 0$ , both fundamental forms are equal to zero. If the partial derivative of the field with respect to a given component is non-zero, that component can be solved for within  $I(dx, dy, dz)$  and substituted into  $I(dx, dy, dz)$ , arriving at the third fundamental form (III). The idea is to factor the new expression so that it is quadratic with respect to the two remaining components. If  $\nabla_z G \neq 0$ , then (148)<sup>[AF]</sup> becomes a function of  $(dx, dy)$ .

$$III(dx, dy) = U(dx)^2 + Vdxdy + W(dy)^2 = 0 \quad (148)$$

The discriminant of this equation is defined as  $\Delta = V^2 - 4UW$ , which can be used to determine the remaining principle directions (149)<sup>[AF]</sup>.

$$T = \begin{bmatrix} (-V \pm \sqrt{\Delta})\nabla_z \\ 2U\nabla_z \\ (V \pm \sqrt{\Delta})v_x - 2Uv_y \end{bmatrix} \quad (149)$$

These vectors can be solved for in other cases where  $\nabla_z = 0$  by cyclic permutation of the components<sup>[AF]</sup>. The components of  $A_{\mu\nu}$  remain constant, although  $U$ ,  $V$  and  $W$  must each be recalculated and the components of  $T$  rearranged. For example, if  $\nabla_x G \neq 0$  were true instead, the correct principle directions are defined as (150), i.e.  $(x \rightarrow y), (y \rightarrow z), (z \rightarrow x)$ .

$$T = \begin{bmatrix} (V \pm \sqrt{\Delta})\nabla_y - 2U\nabla_z \\ (-V \pm \sqrt{\Delta})\nabla_x \\ 2U\nabla_x \end{bmatrix} \quad (150)$$

The complete curvilinear basis is therefore defined as (151).

$$\hat{e}_1 = \frac{\nabla \nabla}{|\nabla \nabla|}; \quad \hat{e}_2 = -\frac{T_1}{|T_1|}; \quad \hat{e}_3 = -\frac{T_2}{|T_2|} \quad (151)$$

In general, the initial coordinate system used for the gradient and partial derivatives must be orthogonal. The resulting curvilinear coordinates will also consist of an orthogonal basis, which follows the

coordinate lines of the metric tensor ( $g_{\mu\nu}$ ). To transform from the curvilinear basis  $\hat{e}_i$  to  $g_i$ , requires a tensor so that (152) is true. Since each component is already aligned,  $F$  is diagonally symmetric and varies only in magnitude.

$$g_i = F_{ij}\hat{e}_j \quad (152)$$

Due to the parameterization of effective vacuum fields with metric distance, the transformation must be isotropic. In comparison to the relativistic case where the field is compressed in a single direction, the presence of background vacuum energy warps a particles manifold equivalently in all directions at each point. Without this feature, the perceived space-time metric cannot be attributed to a relative medium upon space, further induced by vacuum energy density or a scalar field. From the isotropic nature of general field transformations, the previous tensor  $F$  can be written as (153).

$$F = \begin{bmatrix} \gamma_g & 0 & 0 \\ 0 & \gamma_g & 0 \\ 0 & 0 & \gamma_g \end{bmatrix} = \gamma_g g_{\mu\nu} \quad (153)$$

The tensor (153) is extended to a Minkowski space via (154).

$$F = \begin{bmatrix} \gamma_g^{-1} & 0 & 0 & 0 \\ 0 & \gamma_g & 0 & 0 \\ 0 & 0 & \gamma_g & 0 \\ L & 0 & 0 & \gamma_g \end{bmatrix} \quad (154)$$

Therefore, an arbitrary space-time metric (155) can be defined relative to the curvilinear basis.

$$g_{\mu\nu} = \begin{bmatrix} \gamma_g^{-2} & 0 & 0 & 0 \\ 0 & \gamma_g^2 & 0 & 0 \\ 0 & 0 & \gamma_g^2 & 0 \\ L & 0 & 0 & \gamma_g^2 \end{bmatrix} \quad (155)$$

To determine the proper space-time metric for any object, the effective vacuum energy density must first be calculated; see section (2.6). The curvilinear basis is then determined for time dependence.

## 2.6. Numerical Methods

Since solutions are based upon each individual particles influence on local fields, realistic objects must be numerically determined. Each particle further consists of a localized field relative to a preferred reference frame ( $y_{\mu\nu}$ ). Therefore, the field of a single particle does not follow the metric induced by its own vacuum field. It instead follows the background vacuum field depicted by all other particles and free fields. This is applied to determine the effective space-time metric of any object using quantized variables. For first-order methods, the field of each particle will undergo isotropic deformations. The magnitude of these transformations will be proportional to the general Lorentz scalar at the classical position. Therefore, each particle has a unique general Lorentz scalar (156) due to the effective field of every other particle; note that ( $r_n = 0$ ) is equivalent to ( $d_n$ ).

$$(g_{,n} = 1 + \frac{1}{\Delta} \sum_{m=1}^N \bar{\Psi}_{mn} (r_m = 0)) \quad (156)$$

For two particles at rest relative to the space-time metric, the effective fields are defined as (157).

$$:\bar{\Psi}_1(r_1) = \frac{E_1}{r_1 \gamma_{g,2}} \quad : \bar{\Psi}_2(r_2) = \frac{E_2}{r_2 \gamma_{g,1}} \quad (157)$$

In Planck units, the Lorentz scalars (156) can be written as continuous fractions (158).

$$(g_{,1} = 1 + \frac{E_1}{d_1} \frac{1}{1 + \frac{E_2}{d_2} \frac{1}{\dots}}) \quad (158)$$

Setting  $a = E_1/d_1$  and  $b = E_2/d_2$  results in (159).

$$x = 1 + \frac{a}{1 + \frac{b}{x}} \quad (159)$$

The effective Lorentz scalar is therefore (160).

$$g_{,1} = \frac{1}{2} \frac{(\sqrt{a^2 + 2a - 2ab + (1+b)^2} + a - b + 1)}{(a + b + 1)} \quad (160)$$

If both particles have equivalent rest mass ( $E_o = E_n$ ), the effective field of each is (161); where  $d$  is the distance between particles.

$$\bar{\Psi}(r_n) = \frac{2E_o}{r_n} \frac{1}{(1 + \sqrt{1 + 4 \frac{E_o}{d}})} \quad (161)$$

Assuming the gradient of the local vacuum field is small, the first order approximation is useful for determining the field of large objects. However, systems with more than two particles are more difficult to deal with. The objective is to develop an iterative numerical method that is equivalent to the algebraic results.

The constant isotropic deformation is retained for dust solutions. Assuming each particle is stationary relative to the effective metric ( $g_{\mu\nu}$ ), the algorithm for two particles is the following.

1. Start with the constant, non-effective field of both particles; i.e.  $E_1/r_1$  and  $E_2/r_2$ .
2. Determine  $(g_{,1})$  and  $(g_{,2})$  from (156), relative to the non-effective fields defined in 1.
3. Iterate by updating effective fields defined as  $E_1/(r_1 (g_{,1}))$  and  $E_2/(r_2 (g_{,2}))$ .
4. Determine  $(g_{,1})$  and  $(g_{,2})$  relative to the effective fields calculated in 3.
5. Loop back to 3.

This can also be calculated by hand, where the first few iterations are (162) in arbitrary units.

$$:\bar{\Psi}_1 = \frac{E_1}{r_1} \quad (11) : \bar{\Psi}_1 = \frac{E_1}{r_1} \frac{1}{1 + \frac{E_2}{d_2} \frac{1}{\dots}} \quad (162)$$

$$I2 :\bar{\Psi}_1 = \frac{E_1}{r_1} \frac{1}{1 + \frac{E_2}{d_2} \frac{1}{1 + \frac{E_1}{d_1} \frac{1}{\dots}}} \quad (162)$$

The previous iterative method is equivalent to solving the continuous fractions of equation (156) and can be extended to N particles.

Expanding equation (156) by hand matches the iterative method for three particles, although the equations become relatively large. The effective Lorentz scalars will always be linear combinations of nested, continuous fractions. This structure is preserved by applying the iterative method outlined in figure 2.2 for any amount of particles.

```

// STEP 1 //
DOUBLE RE[N];           // Rest Energy
DOUBLE EF[N];           // Effective Energy
DOUBLE P[N][3];         // Position
DOUBLE G_C = Δ;         // Reference Energy
DOUBLE lorentz_scalar[N];

// STEP 2 //
FOR(INT A = 0, A < N, A++) {
    y_temp = 0;
    FOR(INT B = 0, B < N, B++) {
        IF(A != B) {
            dx = P[A][0] - P[B][0];
            dy = P[A][1] - P[B][1];
            dz = P[A][2] - P[B][2];
            D = sqrt(dx*dx + dy*dy + dz*dz);
            y_temp += RE[B]/D; }
    }
    lorentz_scalar[A] = 1+y_temp/G_C;
}

// STEP 3 //
EF[A] = RE[A]/lorentz_scalar[A];
}

// STEPS 4-5 //
WHILE(KEEP_ITERATING == TRUE) {
FOR(INT A = 0, A < N, A++) {
    y_temp = 0;

    FOR(INT B = 0, B < N, B++) {
        IF(A != B) {
            dx = P[A][0] - P[B][0];
            dy = P[A][1] - P[B][1];
            dz = P[A][2] - P[B][2];
            D = sqrt(dx*dx + dy*dy + dz*dz);
            y_temp += EF[B]/D; }
    }
    EF[A] = RE[A]/(1+y_temp/G_C);
}
}

```

**Figure 2.2.** C-code for calculating the effective vacuum field of composite objects with first order methods.

The exact vacuum field can be determined by applying the proper, active transformation (163) to each particle.

$$r_n^{\gamma} = r_n + \frac{1}{\Delta} \int_0^{r_n} \left( \sum_{m=1}^N \bar{\nabla}_{mn}(r_n) \right) dr_m \quad (163)$$

When the first-order approximation was defined,  $\mathbf{g}$  was held constant; i.e. the transformed radius (164) in all directions became isotropic.

$$r_n^{\gamma} = \int_0^{r_n} f_{g,n}(r_n) dr_n = r_n g_n(\mathbf{o}) \quad (164)$$

The exact vacuum field cannot be obtained without the unified field theory, which defines  $\nabla_{mn}$ . The far-field approximation is instead applied, which is in agreement with classical theory. The far-field approximation however naturally has singularities, so each particle's field must have a cut-off (165) when  $\nabla_{mn}$  surpasses the maximum quantized value or classical energy.

$$\nabla_{mn} = \begin{cases} \bar{\nabla}_{mn} & : \bar{\nabla}_{mn} \leq \bar{\nabla}_{max} \\ \bar{\nabla}_{max} & : \bar{\nabla}_{mn} > \bar{\nabla}_{max} \end{cases} \quad (165)$$

Now that the fields are finite at all points, it is possible to integrate along the effective field of each. For an exact solution, memory requirements drastically increase since the effective field of each particle must be known at each point in space. This is because each particle's field is relative to the background field of all other particles and free fields. This can be optimized by assuming many particles exist, so that the influence one has on the others is negligible. Therefore, only one effective field is defined based upon the contribution from all particles; i.e. a continuum approximation is made similar to EFEs. Identical numerical methods can be applied with respect to figure 2.2, where the initial configuration converges towards the effective field.

## 2.7. Relativistic Pressure and Bulk Flow

Temperature is a scalar quantity that depicts a system's internal kinematic energy. For a system at equilibrium, the energy distribution of individual particles is related to temperature via the Maxwell-Juttner equation (166)<sup>[AG]</sup>.

$$f(\gamma) = \gamma \sqrt{\gamma^2 - 1} \frac{e^{-\gamma/\theta}}{\Theta K_2(1/\theta)} \quad (166)$$

$K_2(z)$  is a modified Bessel function of the second kind (167) and  $\theta = k_B T / E_o$ .

$$: K_2(z) = \frac{z^2}{3} \int_1^\infty e^{-zx} (x^2 - 1)^{3/2} dx \quad (167)$$

The average Lorentz scalar for the Maxwell-Juttner equation (166) is calculated via (168).

$$\gamma_{avg} = 3\theta \frac{\int_1^\infty (e^{-\gamma/\theta} \sqrt{\gamma^2 - 1}) d\gamma}{\int_1^\infty e^{-\gamma/\theta} (\gamma^2 - 1)^{3/2} d\gamma} \quad (168)$$

For the classical limit, (168) can be approximated by Taylor expansion resulting in (169).

$$\gamma_{avg} = \frac{3}{2}\theta + 1 = \frac{3k_B T}{2E_o} + 1 \quad (169)$$

This reduces to the classical relation between average kinematic energy and temperature (170).

$$E_k \cong \frac{3}{2} k_B T \quad (170)$$

By applying the Maxwell-Juttner equation, the average kinematic energy of particles does not remain proportional to temperature. Therefore, the kinematic energy must be related to the proper force or pressure instead. When dealing with pressure at the atomic scale, a particle that collides with a perpendicular wall will experience a change in momentum via (171).

$$\Delta p_x = 2m \frac{c}{\theta} \sqrt{\gamma^2 - 1} \quad (171)$$

The number density depicts the amount of particles per unit volume. Normalizing this so there is only one particle per finite volume ( $V$ ) allows the metric to be determined. For objects at equilibrium, each particle will be confined to its own respective volume. The one-dimensional force on the plane of another particle's volume is defined by (172).

$$f_x = \frac{m_o c^2 (\gamma^2 - 1)}{V^{1/3}} = \frac{E_o}{V^{1/3}} (\gamma - \frac{1}{\gamma}) \quad (172)$$

Proper pressure assumes that the particle has an equal probability of hitting the other two walls; i.e. the proper pressure is (173).

$$P = \frac{E_o}{3V} \left( \frac{E}{E_o} - \frac{E_o}{E} \right) = \frac{m_o \gamma v^2}{3V} = \frac{E \beta^2}{3V} = \frac{\epsilon \beta^2}{3} \quad (173)$$

Temperature is related to pressure by an averaged Lorentz scalar (267), determined from the beta factor ( $\beta = v/c_o$ ) and proper energy density ( $\epsilon$ ). Pressure itself is only dependent upon average particle energy and the volume attributed to each.

With the averaged Lorentz scalar, the effective field due to pressure can be determined. A massive particle moving with respect to a field will have a vacuum field defined by (174).

$$\bar{V}(r, \theta) = \frac{E_o}{r} \frac{1}{\sqrt{(\gamma \cos \theta)^2 + \sin^2 \theta}} \quad (174)$$

Since each particle's velocity has an arbitrary direction, the field must be averaged over  $d\Omega$ . As with section (2.4), the coordinates are reoriented so that the correct integral is (175).

$$\bar{V}_{avg} = \frac{E_o}{4\pi r} \int_{d\Omega} \frac{\gamma \sin(\phi) d\theta d\phi}{\sqrt{(\gamma \cos \phi)^2 + \sin^2 \phi}} \quad (175)$$

Integrating over  $\theta$ : [0,  $2\pi$ ] and  $\phi$ : [0,  $\pi$ ] provides the averaged field for each particle (176) with respect to the statistical distribution of velocities (168).

$$\bar{V}_{avg} = \left( \frac{E_o \gamma}{2r} \right) \frac{LN(\gamma^2 + \sqrt{\gamma^4 - 1}) - LN(\gamma^2 - \sqrt{\gamma^4 - 1})}{\sqrt{\gamma^4 - 1}} \quad (176)$$

Reducing an object to its individual particles allows the effective field to be approximated with the iterative methods discussed in section (2.6). Determining the effective field of a gas under bulk flow requires a vector field ( $\bar{v} \rightarrow$ ) relative to the space-time metric. Considering a system of particles under bulk flow, the effective field of each particle must be relative to the effective field of all others. Bulk flow is therefore the transportation of kinematic energy density along the space-time metric. The averaged distribution of velocities from (166) must be properly added to the bulk flow ( $\bar{v} \rightarrow$ ). This is equivalent to applying a Lorentz boost in arbitrary directions of  $\bar{u} \rightarrow$ . Under the assumption that bulk flow moves freely in the forward direction, the problem can always be reduced to an addition of velocities. Proper velocity addition relative to the angle between a particle's velocity ( $\bar{u} \rightarrow$ ) and bulk flow ( $\bar{v} \rightarrow$ ) is (177)<sup>[B]</sup>.

$$\bar{u} \rightarrow \bar{v} \rightarrow = \frac{\frac{v}{u} - \frac{1}{u \cos(\theta)}(1 - \frac{1}{u}) + \frac{\bar{v} \rightarrow}{u}}{1 + u \cos(\theta)} \quad (177)$$

The first step is determining the final velocity relative to the metric after applying (177) for all directions of  $\bar{u} \rightarrow$ . This can be accomplished with an active rotation derived from the Rodrigues' rotation formula (178)<sup>[AH]</sup>.  $\bar{K} \rightarrow$  is the axis of rotation and  $M = K K^T$ ; the trigonometric functions are written as  $\sin(x) \rightarrow S_x$  and  $\cos(x) \rightarrow C_x$  to shorten notation.

$$R = S \begin{bmatrix} 0 & -K_3 & K \\ K & 0 & -K^2 \\ -K_3 & K_1 & 0 \end{bmatrix} + C(M + I) + M \quad (178)$$

By restricting the axis of rotation to the x-y plane

and applying a 2-dimensional rotation to  $\bar{K} \rightarrow$ , (177) can be mapped to the unit sphere (179).

$$R = \begin{bmatrix} C_\beta^2 + C_\beta S_\beta^2 & S_\beta C_\beta(C - 1) & -S_\beta S_\beta \\ S_\beta C_\beta(C - 1) & S_\beta^2 + C_\beta C_\beta^2 & -S_\beta C_\beta \\ S_\beta S_\beta & S_\beta C_\beta & C \end{bmatrix} \quad (179)$$

Since the choice of  $u_o$  is arbitrary, it will be set parallel to the bulk flow vector field ( $\bar{v} \rightarrow$ ). Therefore, the proper velocity of each particle under the ( $\alpha, \beta$ ) parameterization is (180).

$$\bar{w} \rightarrow = \frac{u + v C \left( \frac{I}{1 - \frac{v(u)}{u}} \right) - S_\beta S_\beta}{I + u v C} \left[ -S_\beta C_\beta \right] + \frac{\left( \frac{v}{\gamma(u)} \right)}{I + u v C} \left[ 0 \right] \quad (180)$$

From the composite velocities, the effective field for each is defined as (181).

$$\bar{V}(x', y', z') = E_o \left( \left( \frac{x'}{\gamma(w)} \right)^2 + \left( \frac{y'}{\gamma(w)} \right)^2 + (z')^2 \right)^{-\frac{1}{2}} \quad (181)$$

However, the coordinates within (181) are with respect to the direction  $\bar{w} \rightarrow$ . These must be mapped back to the reference frame, where the bulk flow  $\bar{v} \rightarrow$  is defined. This can be accomplished by finding the parameters of the mapping from the direction of  $\bar{w} \rightarrow$  to the z-axis via (182).

$$\begin{bmatrix} w_x \\ w_y \\ w_z \end{bmatrix} = \begin{bmatrix} C_b^2 + C_a S_b^2 & S_b C_b(C_a - 1) & -S_a S_b \\ S_b C_b(C_a - 1) & S_b^2 + C_a C_b^2 & -S_a C_b \\ S_a S_b & S_a C_b & C_a \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ Z_w Z \end{bmatrix} \quad (182)$$

After obtaining (a, b), the transformed coordinates are determined via (183).

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} C_b^2 + C_a S_b^2 & S_b C_b(C_a - 1) & -S_a S_b \\ S_b C_b(C_a - 1) & S_b^2 + C_a C_b^2 & -S_a C_b \\ S_a S_b & S_a C_b & C_a \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (183)$$

Inserting the solutions from (183) into the primed coordinates of (181) provides  $\bar{V}(r, \theta, \varphi, \alpha, \beta)$ . The last step involves integrating over the various configurations in order to arrive at the averaged vacuum field (184).

$$\bar{V}_{avg}(r, \theta, \varphi) = \frac{1}{4\pi} \int_0^{2\pi} \int_0^\pi \bar{V}(r, \theta, \varphi, \alpha, \beta) d\alpha d\beta \quad (184)$$

The problem can also be viewed as a super-position of an infinite number of configurations relative to each ( $\bar{u} \rightarrow$ ). With spherical point picking, a finite amount of these can be rotated in a 3-dimensional space and averaged.

### 3. The Universe

The big bang theory is currently the most widely accepted cosmological model, with the 2011 Nobel Prize awarded for the discovery of accelerated expansion<sup>[BM]</sup>. The model however contains several anomalies, unexplained observations and various non-classical assumptions. These aspects can be resolved by abandoning an expanding model in favor of one that is simultaneously expanding and contracting, i.e. a steady state. Current observations are already sufficient for ruling out an expanding universe. Difficulty of arriving at such conclusion arises from the recent acceptance of non-classical assumptions and lack of theoretical constraints. Dark energy for example is not predicted by the standard model and cannot be directly detected. It is widely assumed dark energy exists solely because it allows an expanding model to fit redshift versus distance modulus. The inferred expansion however is an illusion from the local deflection of geodesics, which produces a nearly spherical projection.

With insight from recent observations, aspects that conclusively rule out an expanding universe can be focused upon. Two characteristics that stand out are incorrect predictions of large-scale curvature and the perspective of time versus redshift; these are discussed throughout sections (3.3, 3.5). It is proven that the observed abundance of faint blue galaxies is due to  $\Lambda$ CDM's incorrect predictions for the curvature of the universe. Additional constraints allow all explanations for the 2 — 3x abundance of faint blue galaxies to be ruled out. These range from evolution of the local luminosity function to drastic mergers.  $\Lambda$ CDM further underestimates the size of the faint blue galaxies by 2 — 5x relative to their angular size versus luminosity. Number densities of weak MgII absorbing galaxies in section (3.6) are also in agreement with the prior conclusion. These incorrect predictions by  $\Lambda$ CDM result in systematic lensing errors as discussed in section (3.4).

Since all explanations can be ruled out relative to  $\Lambda$ CDM, the faint blue galaxy abundance is proof rather than evidence. Although proof exists against  $\Lambda$ CDM, there also exists strong evidence against an expanding universe. The proposed theory predicts for distant galaxies to be older than local ones. An expanding model predicts the opposite, which is contrary to observations. For example, galaxies are observed to cool with increasing redshift. Distant quasars contain relatively higher FeII:MgII ratios, depicting increased metallicity with redshift. There are many other firmly grounded observations not compatible with an expanding universe such as the in-fall velocity of the Bullet cluster. For clarity, the first half of chapter 3 will focus on the foundations of the continuous model. The remaining sections discuss the various proofs against  $\Lambda$ CDM, including a statistical comparison between models.

The new cosmological theory only requires the standard model and corrections to general relativity herein. From this short introduction alone, the new model is superior with respect to Occam's razor. In other words, the simplest theory that agrees with all observations is the correct theory. Similar to initial motivation behind an expanding model, the shape of the universe can be fit with a single constant. The inferred accelerated expansion is nothing more than local geodesics deflecting towards the center of an asymptotically flat, linear universe. Dark energy is therefore not required to explain redshift versus distance modulus. The trend is better fit by distant galaxies falling into an asymptotically flat universe, depicted by relativistic redshift and gravitational acceleration. The cosmic background radiation must therefore originate from the central region of an asymptotically flat universe. Classical assumptions insist that this black body radiation is emitted from a central core, which is not compatible with theories that predict event horizon such as EFEs.

### 3.1. The Big Bang Theory

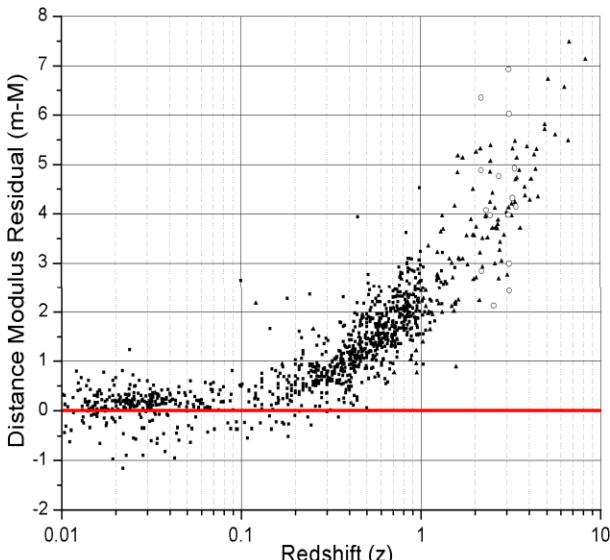
Georges Lemaître, a Belgian priest, was the first to propose the big bang theory, originally named “hypothesis of the primeval atom”<sup>[AK]</sup>. Although Lemaître was the first to discover the “Hubble constant”, it was named after Edwin Hubble. Hubble’s observations in 1929 also showed a linear relationship between the distance and redshift of local galaxies<sup>[AL]</sup>. The Hubble constant and linear trend provided the initial motivation behind an expanding model, where redshift is attributed to the recession velocity of local galaxies. According to the Hubble model, the relation between redshift and metric distance is (185).

$$d^7 = \frac{c}{H_0} \left[ \frac{(z+1)^2 - 1}{(z+1)^2 + 1} \right] = \frac{v}{H_0} \quad (185)$$

Due to the inverse square law, metric distance is related to distance modulus ( $\mu$ ) by equation (186).

$$d^7 = 10^{-\mu/51} \quad (186)$$

Combining equations (185, 186) and subtracting the result from observations, the disagreement becomes apparent. Figure 3.1 provides the error relative to the initial big bang model.



**Figure 3.1.** Data is from the NED database<sup>[AJ]</sup> and linear trend from equation (185) with  $H_0 = 73.8 \text{ km/s/Mpc}$  <sup>[AI]</sup>.

The initial big bang model is only valid under Hubble’s and Lemaître’s limited observations, i.e. for local galaxies with redshift below  $z \approx 0.15$ . It is clear that beyond this point, an expanding universe depicted by the big bang theory would need to be accelerating. Assuming the universe began as a point of infinite energy density that consequently erupted into an expanding sphere of energy, there are two plausible scenarios for recession velocity. For a homogenous universe, any initial acceleration from pressure or bulk flow should be constrained to relatively high redshift. Therefore, the first scenario requires that the mass of the universe is large enough to collapse back onto itself. The second assumes the kinematic energy imparted to matter from a big bang event is large enough to continue expansion at a constant or decreasing rate. Neither possibility fits observations, requiring the ad-hoc introduction of dark energy.

From the previous chapters, several constraints have been placed upon vacuum field theory. These include the conservation of vacuum energy density, its connection to the space-time metric and the localized nature of particles. With these additional aspects, it is clear dark energy has characteristics similar to vacuum energy density. For example, dark energy in  $\Lambda$ CDM does not force matter to become repulsive; it instead acts to expand the space-time metric. This only provides the illusion of acceleration since all observers view the universe relative to the space-time metric. With respect to only experimentally confirmed contributions to an object’s redshift, there exist two explanations for current observations. Either the universe consists almost entirely of undetectable energy and matter causing accelerated expansion, or this improperly inferred expansion is an illusion due to the universe being asymptotically flat. It is therefore important to distinguish between the angular scales and time-dependence of each model.

### 3.2. Redshift and Distance Modulus

The redshift of distant objects can be described in terms of relativistic redshift due to a variation in gravitational potential. Vacuum field theory is not required for determining redshift versus distance; however, it is necessary for large-scale curvature. The relation ( $=_g$ ) is applied to determine the average relative velocity (187) that is induced from a change in vacuum energy density between source and local observer.

$$v = \left( \frac{c_0}{\nabla} \right) \sqrt{\frac{\nabla}{\Delta} \left( \frac{\nabla}{\Delta} + 2 \right)} \quad (187)$$

The Doppler effect requires for any relative velocity to result in a redshift (188).

$$z_1 = \sqrt{\frac{1 + \beta}{1 - \beta}} - 1 \quad (188)$$

Plugging equation (187) into (188) results in the relativistic redshift (189) from a change in  $\nabla$ .

$$z_1 = \frac{1}{\Delta} \sqrt{(\sqrt{\nabla(\nabla + 2\Delta)} + \nabla + \Delta)^2} - 1 \quad (189)$$

There is also a general redshift due to variations in vacuum energy density, i.e. the second component is defined by (190).

$$z_2 = \frac{\nabla}{\Delta} \quad (190)$$

Summing both components results in the effective redshift (191) for distant galaxies and clusters.

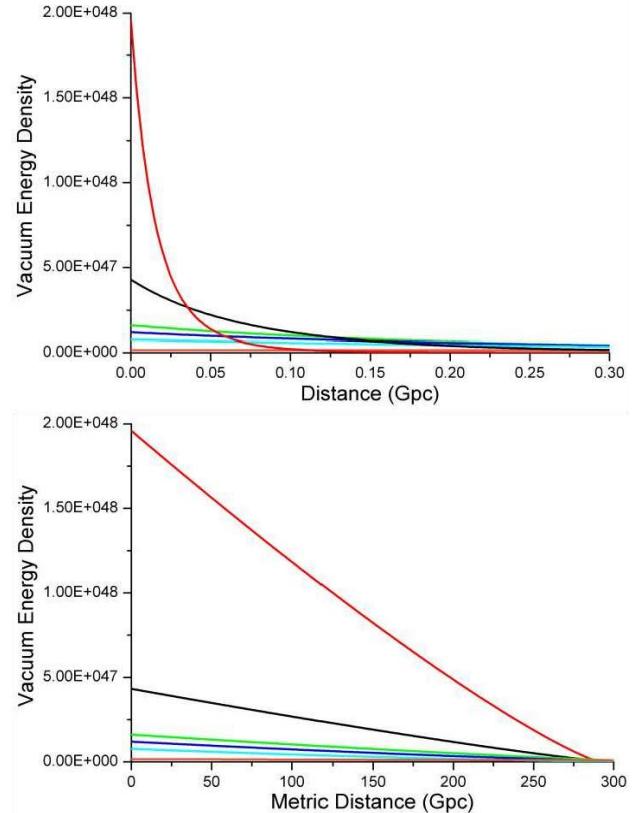
$$z_{\text{net}} = \frac{1}{\Delta} (\sqrt{\nabla(\nabla + 2\Delta)} + 2\nabla) \quad (191)$$

The above relation has a single free variable and depicts a nearly perfect linear trend for all non-local redshift. This not only provides strong evidence for an asymptotically flat universe, but also justifies the linearity of vacuum field theory.

From numerical methods, it is observed that an asymptotically flat universe will generate a field that appears nearly linear with respect to metric distance or local observers. When this is integrated to provide a plot of metric distance versus vacuum energy density, the field becomes approximately linear. Applying equation (191) to the spectral redshift of distant SNIa/GRB demonstrates a linear trend as depicted in figure 3.4. Prior to discussing actual data from SNIa and GRB observations, an ensemble of asymptotically flat universes (192) is provided in figure 3.2.

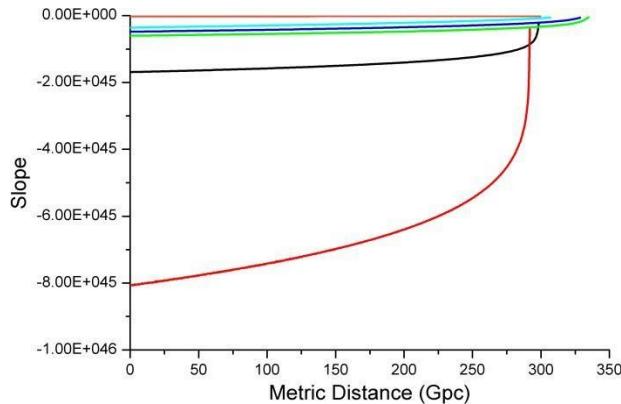
$$\nabla(r) = \frac{E}{r^n} \quad (192)$$

They demonstrate that any realistic, asymptotically flat universe will appear nearly linear relative to a local observer or space-time metric.

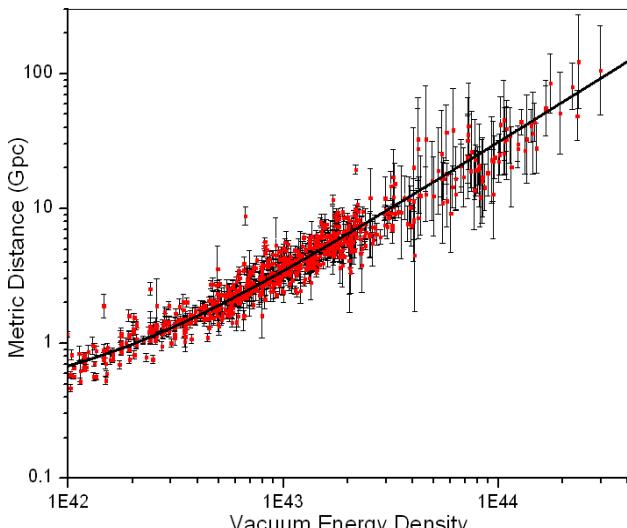


**Figure 3.2.** Each function (192) is indexed by the following colors  $n = 1, 2, 3, 4, 5, 6$ . Total metric distance is held at approximately 300 Gpc. (Top) Ensemble of asymptotically flat universes relative to the preferred reference frame and (Bottom) relative to local observers or metric distance.

The derivative of figure 3.2 with respect to metric distance provides a more accurate representation of the linear variations in figure 3.3. This linear trend is crucial for explaining dark energy or the illusion of accelerated expansion. From the NED database, metric distance to each object is determined from distance modulus. The previous net redshift relation (191) is then applied to the data, resulting in figure 3.4. The best fitting trend is with respect to data beyond 0.15z, which results in a constant slope of  $S_0 = 3.248 \pm 0.047 \cdot 10^{42} (\text{kg m s}^{-2})/\text{Gpc}$  and y-intercept  $y_0 = 0.375 \pm 0.161 \text{ Gpc}$ . Since the flow towards central core is directionally dependent, the y-intercept depicts the average distance to the start.



**figure 3.3.** The slope of each function has origin at the center of the universe. Note that the deviations are minimal when the observer is inside the localized universe, becoming less linear as the edge is approached.



**figure 3.4.** Logarithmic plot of metric distance versus vacuum energy density with respect to Earth.

From the linear slope relative to the space-time metric, an upper limit on metric or luminosity distance to the central core can be determined. The slope may slightly vary as the core is approached; however, this is not noticeable within currently observable distances ( $< 100 \text{ Gpc}$ ). From analyzing the redshift of the core's black body spectrum, the vacuum energy density at the surface is determined with respect to the local space. In other words, the observed spectrum is shifted until it matches the spectral distribution at emission. The core is found to have a redshift of  $z \approx 1089$  with a surface temperature around  $3000\text{K}^{\text{[AN]}}$ . Since this form of redshift is solely due to variations in vacuum energy density, the change in  $\nabla$  can be determined from equation (193).

$$\nabla = \Delta z \quad (193)$$

Therefore, the variation in vacuum energy density from Earth to the core's surface is approximately  $1.315 \cdot 10^{47} (\text{kg} \cdot \text{m} \cdot \text{s}^{-2})$ . With the best fitting linear slope (all data beyond  $z > 0.15$ ), the maximum metric distance to the central core is determined from equation (194).

$$S_0 d^7 = \nabla = \Delta z \quad (194)$$

Solving for  $d'$  places the core's surface at a metric distance of about 40.6 Tpc. Modern telescopes are only capable of detecting light from the most luminous objects up to 0.1 Tpc, i.e. a small portion of the universe is observable. For comparison, the far-field of a universe with perfect linear slope relative to metric distance can be written as an exponential function (195), where  $\nabla_0$  is the vacuum energy density at the observer's position. With respect to the ensemble plotted in figure 3.2, the function has a different form. However, they all produce nearly constant slopes relative to a local observer or metric distance.

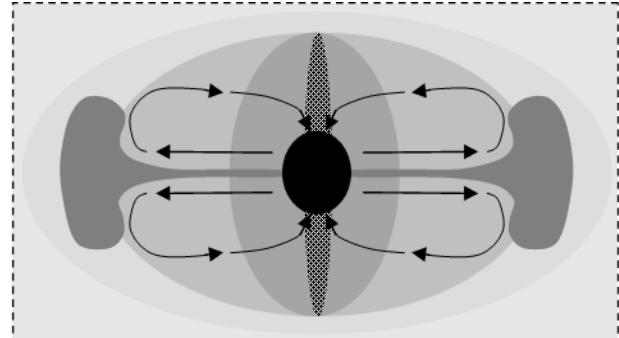
$$\nabla(r) = (S_0 d^7 + \nabla_0 + \Delta) e^{-\frac{r}{r_0}} \quad (195)$$

Relative to an asymptotically flat universe and observed cosmic background radiation (CMBR), it is clear that something large has or had existed prior to the present. If distant galaxies and clusters are falling into an asymptotically flat universe, then there must be a mechanism that transports matter from the central region outwards. A situation similar to that of an explosion is not completely out of the question. However, a mechanism already exists that can replace spherical expansion. Relativistic jets emanating from black holes are not well understood, but have been observed from numerous sources with varying intensity and duration<sup>[AM]</sup>. The first step in producing the model that agrees with all observations is assuming the local space emerged from such jet. In this perspective, distant galaxies and clusters fall back into the equatorial regions of the central core at relativistic speeds. Due to the conservation of momentum, in falling matter is ejected at the polar regions in the form of dense quark matter. This quark matter further decays into hot, x-ray emitting gas commonly seen in young clusters and galaxies. The local jet is later discussed in section (3.8) with respect to the dark flow and cleaned CMBR image.

Black holes that emit polar jets are known to exist after such events. When considering finite black holes, the best approximations available are QCD and vacuum field theory. If the CMBR is to be taken as black body radiation from a massive but cooled object, then the surface must be finite. One could argue that Hawking radiation is already theorized to be emitted from the surface of non-finite black holes. However, temperature in this perspective is inversely proportional to mass. For an Einstein black hole to emit a 3000K blackbody temperature, it would need to be several orders of magnitude less massive than the Moon. This is clearly impossible with respect to a central core, as countless black holes exist locally that are much more massive. Although the various proofs against

an expanding universe have not yet been discussed, they would clearly nullify current interpretations of the CMBR. In other words, the CMBR cannot be due to a period of recombination. It is instead classical black body radiation emitted from the central core. In an asymptotically flat universe, geodesics will begin to deflect from the local space as distance increases. After sufficient distances, the majority of local geodesics will turn towards the center of the universe. Although the projection is not perfectly spherical, it creates the illusion of accelerated expansion. The CMBR is therefore projected onto all local directions of space, as it originates from the center of the universe.

Putting all of the pieces together, a self-consistent model of the universe emerges in figure 3.5. For any steady state model to be valid, the universe must conserve energy and act as a perpetual machine. Other models similar to the cyclic big bang inherently describe the universe as such. However, they suffer from incorrect large-scale curvature similar to  $\Lambda$ CDM. This is later discussed relative to the size and number densities of distant galaxies. Galactic merger times and properties of distant clusters also insist that the universe is in a steady state. In other words, there is a constant flow of matter from the central core to the outer regions, which further flows back to the central region. This requires for distant galaxies and clusters to be older than local populations when observed from Earth.



**Figure 3.5.** A cross section of an asymptotically flat universe in a steady state. The structure takes the form of a  $Y_{20}^0$  spherical harmonic with two polar jets and annihilation boundary between hemispheres.

### **3.3. Revised Galaxy Evolution**

As inferred from section (3.2), distant galaxies are older than local populations. This is contrary to an expanding model, where objects are predicted to be younger as redshift increases. Distant galaxies and clusters with respect to Earth should therefore contain higher fractions of cold baryonic matter, increased star formation rates and high metallicity. Relative to the local space, galaxies and clusters display characteristics that are evident of an origin from hot, x-ray emitting gas. The x-ray emitting gas is the product of decaying quark matter as inferred from its connection to the dark flow. Both  $\Lambda$ CDM and the purposed model are similar in the sense that the local space emerged from dense quark matter. Relative to an expanding model, it is expected that high redshift clusters and galaxies are hotter than similar local populations; observations however depict the exact opposite. For example, local x-ray emitting clusters transform into lyman-alpha blobs beyond 2z. To temporary resolve this problem, it is usually assumed that drastic major mergers take place and heat up the intergalactic medium. However, it is illogical to have two cool clusters with high metallicity merge into a single hot cluster with low metallicity.

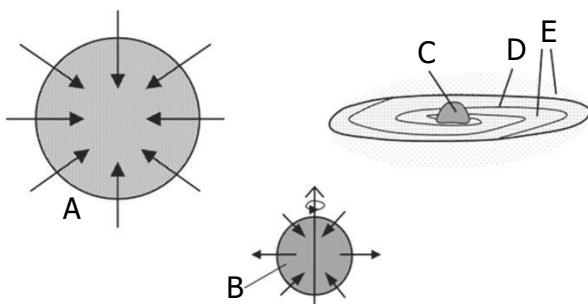
As the local jet of quark matter begins to cool and expand, it decays into a dense non-metallic gas. The oldest stars known to date are metal poor, indicating that they formed some time after this phase. All elements heavier than helium are usually produced through nuclear fusion. The following population II stars are abundant in both globular clusters and elliptical galaxies<sup>[AP]</sup>. Most elliptic galaxies contain only population II stars and large amounts of x-ray emitting gas<sup>[AT]</sup>. The source of x-rays (0.5 — 1.5 keV) is thermal bremsstrahlung due to hot ionized gas (5 to > 15 keV)<sup>[AZ]</sup>. Young stars often undergo supernova after billions of years, enriching the surrounding medium with metallic elements. Around this point, the x-ray emitting gas

originating from the core's relativistic jet begins to drastically cool. Metal-rich population I stars then form in the dense but cooler regions of galaxies and nebulas. Beyond Earth's current position in the flow, baryonic matter is observed to become increasingly colder and more metallic. Due to the abundance of preferred fuel<sup>[AQ]</sup>, galaxies between a redshift of 0.5z — 3z will demonstrate intense star formation<sup>[AR]</sup>. Late-type galaxies commonly have regions of active star formation, which is favored due to cold, dense baryonic matter<sup>[AQ][AU]</sup>. The cold interstellar gas required to ignite these galaxies was recently observed. A letter to Nature states that galaxies at redshift of 1.2z and 2.3z consist of 34% and 44% cold baryonic matter respectively, which is 3 — 10x more than local late-type galaxies<sup>[AQ]</sup>. Beyond the overly abundant blue field galaxies, red and ultra-red galaxies dominate<sup>[AV][AW][AX]</sup>. The red and ultra-red colors are an indication of abundant dust or mature star populations<sup>[AX]</sup>, both being characteristics of older galactic populations. Several of the distant red galaxies within the Hubble deep field image are also undergoing mergers, which is consistent with half of normal galaxies experiencing a major merger by 0.75z.

The previous overview of star formation does not drastically differ from current models. It is instead the evolution of galaxies that must be heavily revised. Although the purposed model and  $\Lambda$ CDM both insist the local universe originated from dense quark matter, their perspective of time versus redshift are opposite. This will vary the inferred evolution of galaxies with respect to observations. For example, predicted time-scales drastically differ between models. To explain galactic formation with respect to  $\Lambda$ CDM, large amounts of dark matter are required. These processes should instead occur over 50 — 100 Gyr, which is why excessive amounts of dark matter are required. Time-dependence is later reinforced by comparing simulated and observed merger times; discussed in section (3.5).

Relative to the purposed model or  $\Lambda$ CDM, the initial environment will consist of decaying quark matter. Galactic formation will therefore take place in a hot, non-metallic gas. With any classical gas, the system will move towards equilibrium with respect to density and pressure. As thermal pressure is overcome by gravity, regions will begin to collapse. Due to the conservation of momentum, any radial collapse will be transferred into angular momentum. The properties of galaxies also depict an evolution from early to late-type; i.e. an older galaxy will be more metallic, contain vast amounts of cold baryonic matter and be less symmetric in shape. This transition from early-type galaxies into late-type is depicted in figure 3.6.

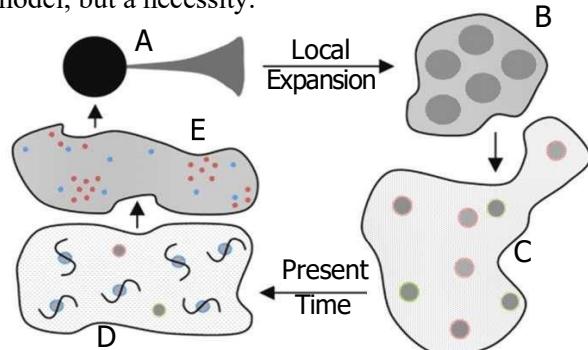
It is commonly debated whether disk galaxies merge to form ellipticals<sup>[AP]</sup>. However, mergers are insignificant with respect to galactic evolution prior to 1z. A large elliptical galaxy was also discovered to contain a rotating disk of x-ray emitting gas<sup>[AS]</sup>. In addition, elliptic galaxies at 0.5z are observed to be on average rotating faster than those in the local space<sup>[AY]</sup>. Distributions of star populations further agree with the purposed model of evolution<sup>[BA]</sup>. Due to flawed foundations however, various modern theories must be discarded. This includes drastic merger rates and dark matter, as the cooling of x-ray emitting gas and unstable rotational curves are the mechanisms behind galactic evolution.



**Figure 3.6.** (A) Galaxies initially form from hot x-ray emitting gas, which begins to collapse after sufficient cooling. (B) Early galaxies obtain a preferred axis of rotation due to local or global gravitational fields. Young metal poor stars form in the bulge due to preferred density. (C) Young metal rich stars later form in the remaining bulge (D, E) while older populations are transported outwards due to unstable rotational curves.

Approximately 60% of major galaxies are disk within the local supercluster, while the remaining are mostly elliptical. From a redshift of  $0.5z - 2.0z$  intense blue galaxies dominate, which continuously transform into blue irregulars beyond  $1z$ <sup>[AR]</sup>. It is important to notice however that galactic evolution is minimal for those involved in the faint blue galaxy problem, i.e. ones in the range of  $0.3 - 0.7z$ . This includes luminosity, color and size variations. Application of modern literature must also proceed with caution, as old calculations do not take into account the proper curvature of the universe. For example, a constant number density of elliptic galaxies out to  $1z$  with respect to  $\Lambda$ CDM<sup>[BH]</sup> would indicate a relative decrease after considering proper large-scale curvature.

Another aspect of galactic evolution arises from the statistically significant dark flow. With the bulk of galaxies and clusters falling towards the center of the universe, there must be another local flow that replenishes these populations. From WMAP, small variations in the cosmic background radiation were measured and analyzed<sup>[AZ]</sup>. These variations are due to scattering from clusters of galaxies containing large amounts of x-ray emitting gas. It is clear that hot, x-ray emitting galaxies and clusters should be younger and therefore closer to the local jet. The dark flow is therefore not only expected by the new model, but a necessity.



**Figure 3.7.** (A) Central core with jet consisting of hot, dense quark matter. (B, C) After the quark matter decays into hot x-ray emitting gas, low metallicity clusters, galaxies and stars begin to form. (C, D) Figure 3.6 overviews galactic evolution. (E) Increased merger fractions, metallicity and cold baryonic matter with respect to the local space.

The last part of this section will focus on deriving the time dependence of distant galaxies and clusters. Since all objects not in the dark flow will be falling into an asymptotically flat universe, the redshift equation can be applied to determine kinematics. The goal is to approximate the total amount of proper time a galaxy experiences while following a geodesic originating from Earth's present position, i.e. the reference frame is relative to Earth. Metric distance (196) is defined from the slope of the universe and y-intercept, although only the slope depicts the actual shape.

$$d^7 = (3.0788 \cdot 10^{-43}) \nabla + 0.3748 \quad (196)$$

Solving equation (196) in terms of vacuum energy density results in equation (197).

$$\nabla = (3.248 \cdot 10^{42}) d^7 - 1.217 \cdot 10^{42} \quad (197)$$

Plugging equation (197) into the redshift equation (191) provides the relation between redshift and metric distance. Directional dependence for local redshift however must also be considered, where the y-intercept provides the average distance to the start of flow towards central core.

To determine the time dependence of non-local galaxies, the change in redshift with respect to proper time must be determined. Each distant galaxy or cluster relative to Earth took  $r$  amount of proper time to arrive at the position where currently observed light was emitted. Therefore, the time it took for a light ray to travel from source to observer is not necessary relative to a steady state model. Directional dependence for local redshift ( $<0.2z$ ) can also be considered by varying  $y_o$  from 0 to 0.54 Gpc. For all directions relative to Earth, the average change in vacuum energy density between the start of flow ( $d'_o$ ) and finite amount of metric distance ( $D'$ ) is defined as (198).

$$\Delta \nabla = (3.248 \cdot 10^{42}) D' \quad (198)$$

From the initial position, a change in vacuum energy density can be made relative to the averaged start of flow at  $d^7 \cong 0.375$  Gpc. After normalizing units ( $\Delta = 1$ ), the y-intercept can be negated by applying  $\nabla = 0.02683 \cdot D'$ . The normalized redshift equation (199) is then applied to relate an objects spectral redshift to a variation in metric distance  $D'$  with respect to Earth.

$$z_{net} = \sqrt{\nabla(\nabla + 2)} + 2\nabla \quad (199)$$

Since galactic evolution is relative to proper time, proper velocity must be applied (200).

$$u = \frac{dx'}{dr} = v(\nabla + 1) = c_o \sqrt{\nabla(\nabla + 2)} \quad (200)$$

With both (199, 200), the proper velocity and redshift are coupled to metric distance ( $D'$ ) as the only free variable. The proper velocity versus redshift provides all information necessary in order to determine the duration of proper time a distant object has experienced. Proper velocity is relative to the amount of metric distance traveled with respect to a moving objects perspective of time. The averaged proper velocity over metric distance  $D'$  is (201). Carrying out the integral of equation (201) and applying the relation  $u_{avg} = \Delta x'/\Delta r$  provides the total proper time each galaxy has experienced without considering directional dependence.

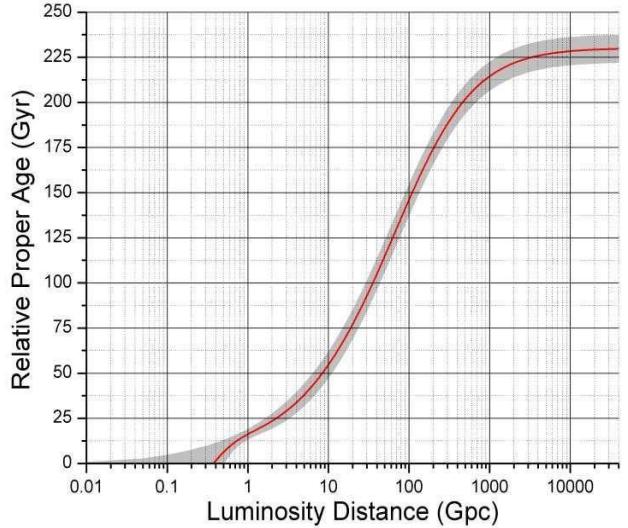
$$u_{avg} = \frac{1}{D'} \int_o^F u(D') dD' \quad (201)$$

It is also important to realize that younger galaxies and clusters exist with respect to the local jet, observed in the form of the dark flow. Neither Hubble's law nor the new redshift equation are capable of modeling this since it has a separate origin. Therefore, the error with respect to the relative age of local objects becomes large due to this uncertainty. However, the majority of objects will follow the previously derived time dependence.

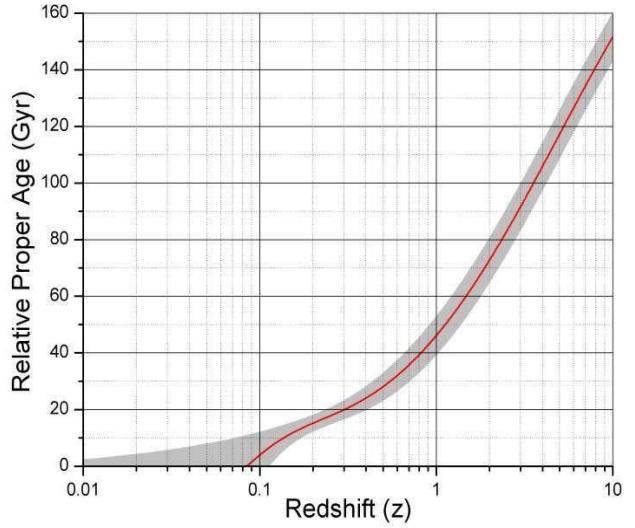
Due to matter emerging from and falling back towards the central core at relativistic speeds, there must be a turning point where relative motion is minimal. The CMBR dipole moment provides a velocity of  $627 \pm 22 \text{ km s}^{-1}$  for the local group<sup>[AZ]</sup>. This is approximately 0.209% the speed of light, which is both insignificant and expected for regions that are distant from the central core. If the metric distance from Earth's current position to the central core is about 40.6 Tpc, the most distant 98% of metric space would account for only 7.8% of all proper time experienced. The most distant objects currently observable are approximately 140 billion years older than the local group ( $z \approx 8.0z$ ). For comparison, the Sun would take about 100 billion years to consume all of its hydrogen fuel. Although this would not actually occur, it is clear that the depletion of interstellar hydrogen occurs over long time-scales. For example, UDFy-38135539 is a distant galaxy observed in the Hubble Deep Field and demonstrates strong lyman-alpha emission<sup>[BB]</sup>. The light being emitted by the object is passing through dust that has been reionized beyond what can be explained by an expanding model. The presence of neutral hydrogen gas would only be plausible if the universe is cooling as redshift increases. In addition, the galaxies redshift ( $8.55z$ ) would correspond to a proper age of about 144 Gyr with respect to the local group.

Since baryonic matter is already observed to be cooling from the local region up to  $2.3z$ , evidence supports the aging of galaxies with increasing redshift. This abundance of cold baryonic matter further induces intense star formation, which is observed for distant late-type galaxies ( $0.5z$  to  $3z$ ). These redshift correspond to proper times between 28 and 92 Gyr with respect to local populations. Therefore, the epoch of intense star formation is in agreement with the continuous model and expected conditions. Minimal evolution however occurs prior to  $0.7z$  (35.7 Gyr), which is later discussed through-

out section (3.5). With respect to proper time,  $\Lambda$ CDM predicts that the current age of the universe is  $13.75 \pm 0.13 \text{ Gyr}^{[\text{AO}]}$ . Since Earth's position is in proximity to the CMBR dipole turning point, the total proper time experienced from the core's jet back to the surface is twice that of future proper time. From figure 3.8, the amount of proper time experienced up to the surface of the central core with respect to Earth is approximately 230 Gyr. The total time experienced by an observer in the bulk of flow is therefore  $460 \pm 100 \text{ Gyr}$  relative to a complete cycle external to the core.



**Figure 3.8.** Proper age ( $r$ ) versus luminosity distance ( $D_L$ ). Error is derived from uncertainty in slope, with the y-intercept ranging from  $0 \text{ Gpc}$  to  $0.54 \text{ Gpc}$ .



**Figure 3.9.** Proper age versus effective redshift with respect to Earth; constraints are identical to figure 3.8.

### 3.4. Angular Scales and Weak Lensing

Although many characteristics can differentiate between models, the curvature of the universe is the only one that offers conclusive proof with current observations. The angular scale versus redshift for  $\Lambda$ CDM was obtained from recent constraints<sup>[BN]</sup>. The angular scale or scale-factor (kpc'') of an asymptotically flat universe is derived from the following method. The distance to an object is determined from its luminosity or metric distance. Considering that local geodesics begin to curve towards the center of the universe at relatively short distances, the projection of distant space will appear almost spherical relative to Earth. The only missing factor is the variations in metric volume induced by the vacuum field of an asymptotically flat universe. From vacuum field theory, the space-time metric is locally isotropic and defined by a scalar field. The circumference of a sphere projected from distant space (202) can therefore be scaled by  $\gamma_g$  relative to Earth's perspective.

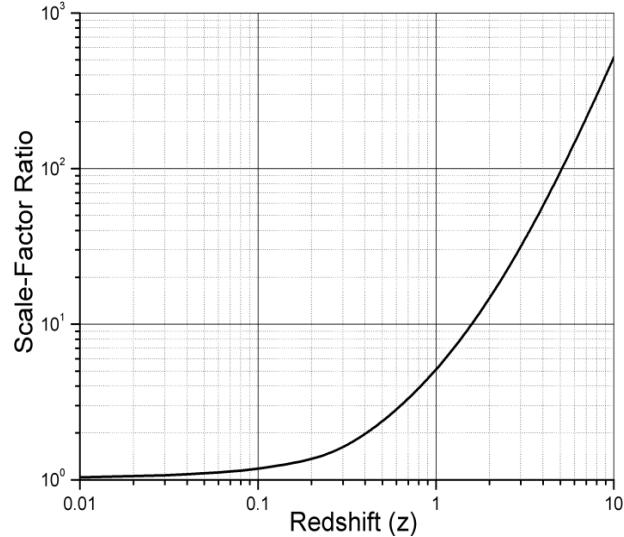
$$C = 2\pi d'\gamma_g \quad (202)$$

Figure 3.10 provides a comparison between  $\Lambda$ CDM and the continuous model.



**Figure 3.10.** The continuous model is depicted in red with  $\Lambda$ CDM in black. Due to local directional dependence, Hubble expansion with  $H_0 = 68.7$  is applied for redshift below  $0.2z$ . The continuous model however is in agreement with local redshift when applying the full range of  $y_0$ .

Plotting the scale-factor ratio between models in figure 3.11 provides several important constraints. This variation is key to ruling out an expanding universe. Although the models are opposite in several aspects, acceptance of poorly constrained hypothesis makes it difficult to rule out  $\Lambda$ CDM. For example, the amount of cold baryonic matter is observed to drastically increase from the local space to  $1.2z$  and  $2.3z$ . With only classical assumptions, galaxies should cool as they age. Although mergers can heat up ISM or ICM, such drastic increase in cold baryonic matter with redshift should be taken as strong evidence against  $\Lambda$ CDM. The processes that occur in the cooling or heating of normal galaxies however are poorly constrained. Ruling out an expanding universe therefore requires aspects that are fully constrained. This is why the angular scales are crucial, as  $\Lambda$ CDM cannot explain the  $2 - 3x$  excess of faint blue field galaxies or the disagreement between their size and luminosity. Incorrect predictions of large-scale curvature also induce systematic lensing errors, which currently provide the only direct evidence for dark matter. The continuous model on the other hand does not require the non-classical assumptions of dark matter and dark energy.



**Figure 3.11.** Scale-factor ratio between the continuous model and  $\Lambda$ CDM with respect to redshift. A  $5x$  disagreement exists by  $1z$  further increasing to  $100x$  around  $5z$ .

The disagreement between each models scale-factor allows distant gravitational lenses to appear stronger than expected from only visible matter. With respect to an expanding universe, there are three distance scales. These consist of angular diameter distance ( $D_A$ ), comoving distance ( $D_C$ ) and luminosity distance ( $D_L$ ). The angular diameter distance corresponds to the visual size of an object at a given redshift, written as (203)<sup>[BO]</sup>.

$$D_A = \frac{c_0}{H_0(1+z)} \int_0^z \frac{dz'}{\sqrt{\Omega_m(1+z)^3 + \Omega_A}} \quad (203)$$

The remaining distances are related to bolometric luminosity and flux via (204)<sup>[BO]</sup>.

$$D_L = \frac{\sqrt{F}}{4\pi F} = (1+z)D_C = (1+z)^2 D_A \quad (204)$$

The continuous model does not require comoving distances since the universe is in a steady state. With respect to distant galaxies and clusters falling into an asymptotically flat universe, the angular diameter distance will vary from that of  $\Lambda$ CDM. Distant objects in an expanding universe appear larger with respect to a flat, steady state projection. The continuous model predicts the exact opposite, where metric volume increases with distance or redshift. Distant objects will therefore appear much smaller in comparison to a flat space-time metric. Luminosity distance in this case is still equivalent to the classical flux definition; however, the angular diameter distance becomes (205).

$$D_A = D_L \left( 1 + \frac{S_0 G_0}{c_0^4} (D_L - y_0) \right) \quad (205)$$

The slope ( $S_0$ ) was discussed in section (3.2), with the  $y$ -intercept ( $y_0$ ) providing the average distance to the start of flow towards central core. Redshift becomes directionally dependent due to the local deflection of geodesics; on average, this distance is about 0.375 Gpc. The right side of equation (205) should therefore have  $D_A$  set equal to  $D_L$  prior to  $y_0$ .

With respect to dark matter, the only evidence for its existence is gravitational lensing from distant clusters. Although proof that the universe is not expanding is not discussed until section (3.5), large variations between scale-factors will clearly induce systematic lensing errors. In general, an expanding universe will overestimate distant lens efficiency with respect to visible matter. Dark matter is also self-contradictory with observations. For example, the Train Wreck cluster has lumps of dark matter that coincide with both galaxies and ICM<sup>[BP]</sup>. This is strong evidence for systematic lensing errors, as dark matter should not interact beyond gravitation. Regardless, dark matter is always located around baryonic matter. This is expected when systematic lensing errors are involved, i.e. only the visible matter existed to begin with.

Properties of the Bullet cluster with respect to  $\Lambda$ CDM are also contradictory, where it is claimed to be proof of dark matter<sup>[BQ]</sup>. At the same time, the existence of this cluster is not compatible with  $\Lambda$ CDM<sup>[BR]</sup>. Observations that are incompatible with a theory cannot be seen as proof for a specific aspect of it. The remaining reason dark matter has been inferred are the unstable rotational curves of galaxies. Assuming the virial theorem is valid in this case originates from an ad-hoc attempt at forcing observations to agree with an expanding model. This assumption is flawed as inferred from observed galactic evolution. The formation of galaxies does not begin from reionized plumes of hydrogen gas anchored to dark matter, but instead hot x-ray emitting gas. Without large amounts of dark matter, galaxies could not have formed within the time-scales predicted by an expanding model. From section 3.3, the age of the local space is well over 100 Gyr more than what an expanding model predicts. This disagreement increases with redshift, where time is viewed in the reverse of actuality. The inferred existence of dark matter therefore originates from improper foundations.

A spherically symmetric lens can be applied to compare the continuous model and  $\Lambda$ CDM. The inferred magnification of a lens is dependent upon the angular diameter distances ( $D_A$ ). As previously stated, current observations support the conclusion of systematic lensing errors, which arises from the incorrect shape of the universe. Since vacuum field theory simplifies to EFE's weak field limit, the formulation is equivalent. Therefore, the equation for a spherically symmetric lens is (206)<sup>[BS]</sup>.

$$\beta = \theta - D_{OL} \frac{4G_o M}{c^2 \theta} = \theta - k \frac{4G_o M}{c^2 \theta} \quad (206)$$

The free variables are  $D_{OL}$ ,  $D_{LS}$  and  $y$ ; these depict the distance from observer to lens, observer to source and the horizontal offset of the source respectively. The remaining variables represent the angular offset of the source ( $\beta$ ) and image ( $\theta$ ). The magnification due to a spherical lens is therefore determined by equation (207).

$$\mu = \frac{\theta d\theta}{\beta d\beta} \quad (207)$$

From the initial conditions,  $D_{OS} = D_{OL} + D_{LS}$  and  $\beta$  is determined from equation (208).

$$\beta = \arctan\left(\frac{y}{D_{OS}}\right) \quad (208)$$

Solving the lens equation (206) results in (209).

$$\theta = \frac{1}{2} \left( \beta \pm \sqrt{\beta^2 + k \frac{16G_o M}{c_o^2}} \right) \quad (209)$$

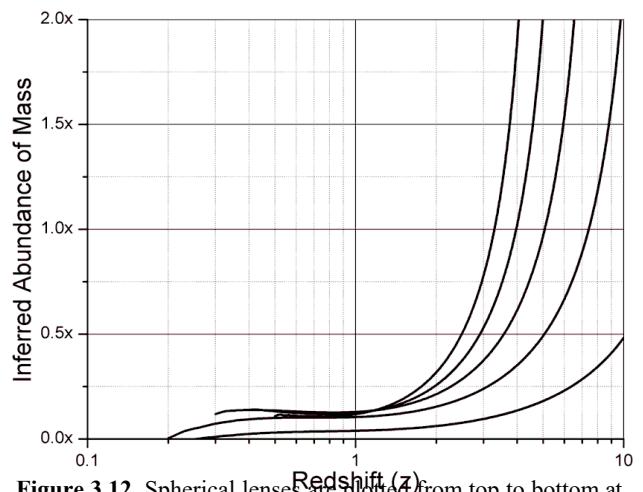
Table 3.1 provides a comparison between the Train Wreck cluster (0.201z) and Bullet cluster (0.296z).

**Table 3.1.** *k* factor between continuous model and  $\Lambda$ CDM

Redshift	Continuous Model		$\Lambda$ CDM	
	k (0.201z)	k (0.296z)	k (0.201z)	k (0.296z)
0.500z	0.7424	0.3487	0.6749	0.3537
1.000z	0.9574	0.5637	0.8673	0.4996
2.000z	1.036	0.6423	0.8976	0.5299
3.000z	1.055	0.6614	0.8506	0.4829
4.000z	1.063	0.6691	0.7856	0.4179
5.000z	1.067	0.6730	0.7140	0.3464

Although the disagreement between models is relatively small prior to 1z, the systematic error becomes apparent from moderate to high redshift. The ratio of distances provides a single variable ( $k$ ) that directly scales the mass in equation (209). This indicates that ratios between each models  $k$  is equal to the inferred abundance of mass. Figure 3.12 depicts this abundance relative to several spherical lenses. The majority of lensed sources range from 3 — 5z, where the disagreement becomes apparent. For example, the Train Wreck cluster would be inferred to have 49% more mass at 5z than actually exists with respect to  $\Lambda$ CDM. The Bullet cluster would instead be inferred to have 94% more mass at identical redshift.

As previously discussed, the initial motivation behind dark matter originates from applying the virial theorem to disk galaxies. Unstable rotational curves and cooling of x-ray emitting gas however are the mechanisms behind galactic evolution. This is reinforced by observations of entropy, angular size, number densities and stellar populations. To avoid circular reasoning, only lensing data should be compared to the visible baryonic matter within clusters. Section (3.5) further demonstrates that an expanding model or  $\Lambda$ CDM predicts incorrect large-scale curvature. This invalidates current lensing data and claimed proof of dark matter.



**Figure 3.12.** Spherical lenses are plotted from top to bottom at 0.5z, 0.4z, 0.3z, 0.2z and 0.1z respectively. The abundance is relative to the continuous model versus  $\Lambda$ CDM.

### 3.5. The Faint Blue Galaxy Problem

The abundance of faint blue galaxies (FBG) up to moderate redshift is known as a grand cosmological problem<sup>[BT]</sup>. With observed merger fractions and the angular size of these galaxies, the problem can be resolved with only classical assumptions. From the LDSS deep redshift survey, an 2x abundance exists up to  $M_B = 22.5$  with respect to no evolution<sup>[BU]</sup>. The survey provides an average redshift of  $0.32z$  at  $M_B = 21.8$ . More distant surveys indicate an 2 — 3x abundance within the limits of  $M_B = [22.5, 24]$ <sup>[BV]</sup>. The majority of recent studies also focus on the B-band, where most of the FBGs with  $M_B$  ranging from 23 to 24 exist prior to  $1.0z$ <sup>[BW]</sup>. With high-resolution imaging, distant FBGs are found to be consistent with local disk and irregular galaxies<sup>[BX]</sup>. On average, this dataset provides a similar redshift versus  $M_B$  of 21.9 at  $< 0.34z >$ . The observed OII widths of distant FBGs also indicate intense star formation across the entire disk<sup>[BX]</sup>. Many of the local FBGs ( $< 0.3z$ ) however are dwarfs. To the contrary, distant FBGs are not dwarf galaxies but instead intermediate disks and irregulars. Applying the purposed model with FBG surveys produces the various plots in figure 3.13<sup>[BU][BZ]</sup>.

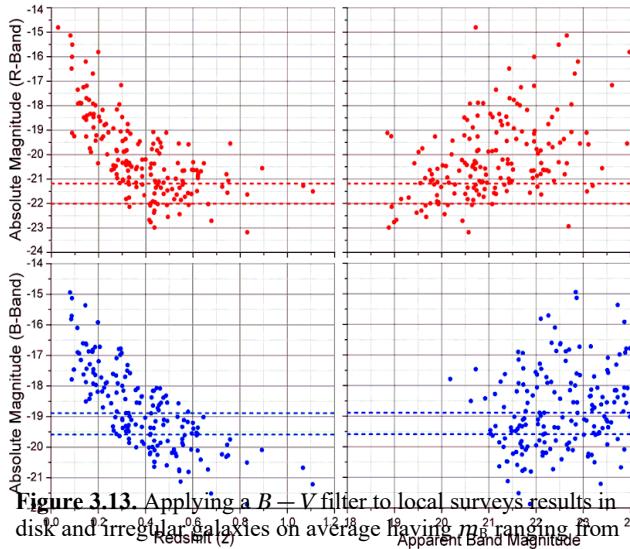
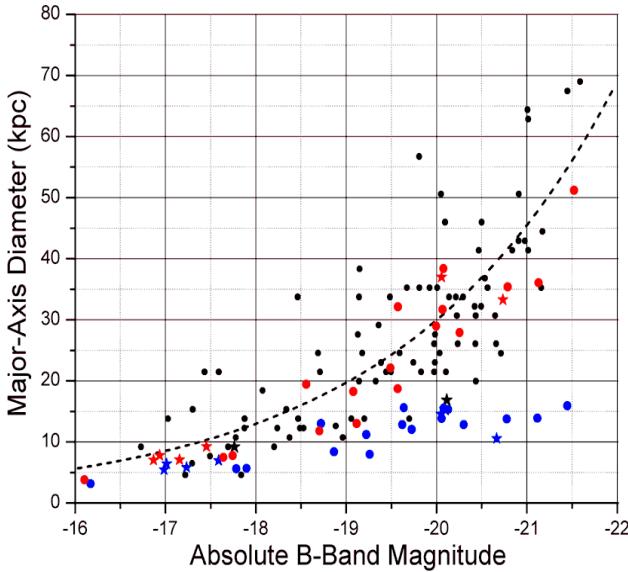


Figure 3.13. Applying a  $B + V$  filter to local surveys results in disk and irregular galaxies on average having  $m_B$  ranging from  $-21.19$  to  $-22.01$ . After combining several faint blue galaxy surveys, it is observed that many of the FBGs between  $0.3z$  and  $0.7z$  are either normal disks or irregulars.

Due to the various attempts at explaining the abundance of field galaxies, each will be discussed in detail. The local FBG abundance is completely compatible with no evolution with respect to their redshift distribution<sup>[CA]</sup>. Many have purposed that either drastic mergers or evolution of the luminosity function must take place. However, number counts in the K, R and B bands rule out evolution in the faint end of the luminosity function<sup>[CA]</sup>. Additional studies have also concluded that any evolution at the bright end of the luminosity function must be minimal below  $0.5z$ <sup>[BY]</sup>. Therefore, the observed 2x abundance placed around  $0.5z$  ( $22.5 M_B$ ) cannot be explained by evolution of the luminosity function. Furthermore, recent constraints on merger fractions limit the total amount of major mergers in these regions to  $< 30\%$  by  $0.5z$ . This would in return only reduce the 2x abundance to 1.7x. The FBG anomaly is known as a grand cosmological problem because there is no self-consistent way to explain the abundance with respect to  $\Lambda$ CDM<sup>[BT]</sup>. The local abundance of FBGs is crucial for ruling out an expanding model since the various aspects are well constrained. When the observed evolution of local blue galaxies is applied to the continuous model however, the 2x abundance is in perfect agreement with predictions.

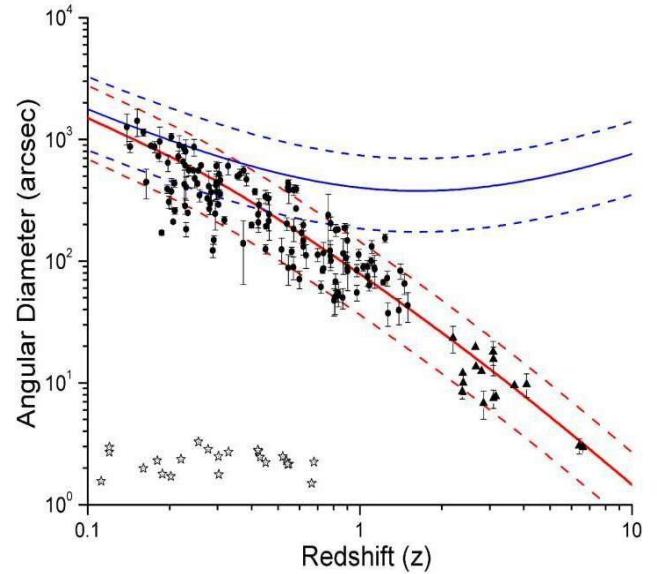
The FBG problem at moderate redshift becomes problematic as they display properties of normal sized disk and irregular galaxies. However,  $\Lambda$ CDM predicts that they are 2 — 5x smaller than similar local populations. This has lead to distant FBGs being improperly inferred to as dwarf galaxies. The luminosity function however does not drastically evolve, so this cannot be true. Taking the average of absolute magnitudes from  $0.3z$  to  $1.0z$  results in  $< m_B > = -19.51$  and  $< m_R > = -20.96$ . This further supports distant FBGs being common late-type galaxies. From high-resolution imaging of FBGs, there is no evidence for an abundance of dwarfs undergoing intense star formation<sup>[BX]</sup>. Observations clearly rule out the purposed solutions to the  $\Lambda$ CDM faint blue galaxy problem.

Several surveys allow the size of FBGs to be compared to their absolute B-band magnitudes. The available data is limited, with the majority lacking reliable redshift. Edge detection with difference of gaussians was instead applied to high resolution FBGs<sup>[BX]</sup>. This provides major-axis diameters with respect to absolute B-band magnitude as depicted in figure 3.14. Absolute magnitude is related to metric distance derived from spectroscopically confirmed redshift. As previously stated, the distant FBGs are not dwarfs. Bolometric limitations force galaxies with faint absolute magnitudes to be closer, while the remaining are more distant. This is observed in figure 3.14, where disagreement between models increases as  $m_B$  decreases. Half of the FBGs are observed to have mild to moderate star formation, usually across the entire disk. The rest range from common young to late-type disk, some of which are very blue or bulge dominated<sup>[BX]</sup>.  $\Lambda$ CDM is below even the most extreme cases from the local space. Combined with the lack of luminosity evolution, an expanding universe is ruled out due to incorrect predictions of large-scale curvature. Observations instead insist that the universe is asymptotically flat.



**Figure 3.14.** Black stars depict several local starburst galaxies, with the remaining circles being normal spiral, dwarf and irregulars. Red are the high resolution FBGs with respect to the continuous model, while blue is with respect to  $\Lambda$ CDM. The dashed trend line is relative to local galaxies with the standard logarithmic fit<sup>[CB]</sup>:  $m_B = -5.5 \cdot \log(A) - 11.9$

Further insight can be obtained from the most massive objects visible at various redshift ( $< 7z$ ). Locally these are hot x-ray emitting clusters, while the more distant populations consist of reionized hydrogen known as lyman-alpha blobs<sup>[CC][CD]</sup>. As predicted by the continuous model, distant clusters are older with respect to local clusters. The angular diameter of 148 clusters was further measured to demonstrate that the curvature of an asymptotically flat universe agrees with observations. The diameter of each is measured with respect to x-ray emissions up to  $3\sigma$  above background rates. For more distant clusters, the extent of lyman-alpha emission was instead applied. Figure 3.15 provides a plot of these clusters and expected size assuming no evolution. The majority of objects were observed by Chandra ACIS-I<sup>[CE]</sup> and the XMM cluster survey<sup>[CF]</sup>. Since the angular scale was previously verified with FBG luminosity out to  $0.7z$  in figure 3.14, there must be minimal change in cluster size up to this redshift. Mergers in figure 3.15 are depicted by large errors, where uncertainty provides the minor to major axis diameters. A major merger will peak at 30% above initial angular diameter, with an error around 25%.



**Figure 3.15.** 148 massive clusters with diameters determined by averaging the minor and major axis of each. Red is relative to the continuous model with a cluster size of  $3.25_{-1.7}^{+2.7} 5 Mpc$  and blue is with respect to  $\Lambda$ CDM. Circles depict x-ray emitting clusters, while triangles have lyman-alpha emission. Stars at the bottom represent the FBGs from figure 3.14.

The FBG problem becomes more apparent when observed merger fractions are examined. Several recent studies focus on calculating merger fractions by surveying distant galaxies. For moderate redshift ( $< 1.2z$ ), galaxies with low to intermediate mass are inferred to have merger fractions of 5 – 10%<sup>[CG]</sup>. From a survey of massive galaxies, the merger fraction ranges from 0.03 to 0.14<sup>[CH]</sup>. Another finds a morphological merger fraction less than 6% for massive disk galaxies prior to 1z<sup>[CI]</sup>. An in-depth survey focusing on both active and prior mergers allows uncertainty in merger duration to be ignored. It is observed that major mergers such as those between two medium-sized disks occur once on average by 1.4z<sup>[CJ]</sup>. Minor mergers between satellite galaxies and their host are about three times as abundant. However, these do not explain the excess of intermediate sized FBGs. In other words, the FBGs are not satellite galaxies. Distant FBGs are fully consistent with common late-type galaxies, with local populations being dwarfs similar to NGC 4214 or NGC 1310. Without drastic merger rates, there are no remaining explanations with respect to an expanding model. The purposed model however predicts for the excess to exist due to the curvature of the universe.

Merger fractions beyond 1z increase in response to the time-scales involved, i.e. these objects are at least 45 Gyr older than the local group. Red and ultra-red galaxies are found at moderate redshift. The Hubble Ultra Deep Field (HUDF) shows many of these galaxies undergoing mergers, resulting in deformed galaxies with tails or multiple cores<sup>[AW]</sup>. Products of major mergers are not consistent with local ellipticals, explaining why these red galaxies are fitted with extended star formation histories and abundant dust<sup>[CK]</sup>. From the HUDF, a peak merger fraction of 30% occurs around 2z with massive galaxies<sup>[CL]</sup>. High merger fractions of 40% to 50% are observed beyond 2.5z, where the objects are consistent with Lyman-break galaxies<sup>[CM]</sup>.

Comparing proper time between 0z and 1z, the continuous model has an additional  $38.3 \pm 6.6$  Gyr. Current studies based upon  $\Lambda$ CDM should therefore have observed merger times 4.1 — 5.8x quicker than expected from simulations. These simulations rely on fundamental physics, making it difficult to explain how the process would be occurring at five times the expected rate. Many also underestimate merger times due to the inclusion of dark matter. There is however no proof or direct evidence for dark matter, at least not in any exotic forms. Direct attempts at locally detecting dark matter have also failed<sup>[CN][CO]</sup>. Instead, dark matter is the result of systematic lensing errors and improper foundations. Without dark matter, the expected mass of galaxies and clusters will decrease. Therefore, merger times from prior simulations are likely underestimated.

From one study, the first pass on average occurs at 0.72 Gyr for Sbc galaxies. Max separation occurs on average by 1.20 Gyr, while galaxies merge at  $< 1.88$  Gyr  $>$ . After  $< 2.88$  Gyr  $>$  have passed the galaxy is considered to be a merger remnant<sup>[CP]</sup>. With the merger fractions applied on the next page, the one merger per galaxy by 1.4z fits with an average merger time of 4.0 Gyr. Both values can be compared to inferred merger times with respect to

$\Lambda$ CDM. Several estimates for  $\langle r_{oac} \rangle$  range from 0.2 Gyr to 1.0 Gyr<sup>[CQ]</sup>, which is in disagreement by 2.9x to 17.5x with simulations. A more precise ratio can be obtained from table 3.2. Taking the average of several surveys results in an average merger time of 0.65 Gyr, which results in a 4.4x to 5.4x disagreement with numerical models.

**Table 3.2. Merger times of close galaxy pairs**

Reference <sup>[CQ]</sup>	$\langle z \rangle$	$\langle r_{obs} \rangle_{S08}$	$\langle r_{obs} \rangle_{Co6}$
Patton & Atfield	0.05	0.36	0.37
Lin et al.	0.79	0.63	0.63
de Ravel et al.	0.72	1.61	1.50
Kartaltepe et al.	0.70	0.33	0.35
Bundy et al.	0.83	0.32	0.35
<b>Average</b>	<b>0.62</b>	<b>0.65</b>	<b>0.64</b>

Several sets of merger fractions are available from recent literature. Some of these could however be overestimated for several reasons. For example, mature disk or irregular galaxies can demonstrate several areas of intense star formation, which may be improperly interpreted as remnant cores from a previous merger<sup>[CR]</sup>. The application of maximum likelihood techniques also tends to overestimate merger fractions<sup>[CI]</sup>. A rough estimate is obtained by assuming one major merger per galaxy at 1.4z.

Since merger times relative to  $\Lambda$ CDM are underestimated, the averaged value from simulations is applied (2.88 Gyr). Time dependence is with respect to the purposed model as discussed in section (3.3). The fractional merger rate is defined by equation (210), depicting the fraction of galaxies completing a major merger per proper merger time.

$$R_c = \frac{f_m}{\langle r_m \rangle} \quad (210)$$

For a given redshift, the change in total galaxies is related to the fractional merger rate and number density by equation (211).

$$\frac{dN}{dr} = -R_c N \quad (211)$$

Solving equation (211) for number density results in exponential decay (212).

$$N(r) = N_0 e^{-R_c r} \quad (212)$$

Assuming the merger rate is constant relative to proper time,  $R_t$  is determined with respect to one major merger per galaxy by 1.4z; this results in  $R_t = 0.012 \pm 0.002$  Gyr<sup>-1</sup>. The 2.88 Gyr merger time translates to a constant merger fraction of 0.035 relative to the continuous model. If the 4.00 Gyr value is applied, the constant merger rate varies to 0.048. However, redshift and time are not directly proportional. For example, about 64% of proper time prior to 1.4z occurs before 0.7z. The merger fraction with respect to redshift would therefore be approximately  $< 0.035$  prior to 0.7z and greater than beyond.

With proper time as derived from the continuous model, local merger fractions are easily constrained below 0.04. For example, the fractional merger rate ( $R_t$ ) derived from a constant merger rate provides merger fractions ( $f_m$ ) ranging from 0.030 to 0.040. Distant close pair merger fractions however range from 0.066 to beyond 0.10 throughout the various surveys. Local close pair surveys provide merger fractions between 0.005 and 0.02. Therefore, the approximation of  $R_t = 0.0120$  is overestimated for redshift below 0.7z. Even after applying this value with respect to  $\Lambda$ CDM, the 2x abundance of FBGs at 0.5z only decreases by 28.6%. If the number density at 0.5z is 200 galaxies per metric volume, the final amount of galaxies would decrease to 143 by 0.0z.  $\Lambda$ CDM or an expanding model is therefore off by 43% when overestimating major mergers prior to 0.5z.

Merger fractions above 0.04 for the local space are clearly too high. The survey that is found to be the most consistent focuses on galaxy pairs with  $M < -19.8$ <sup>[CQ]</sup>; this limit coincides with common Sb/Sc disk galaxies. The projected radius for these close pairs ranges from 5 to 20 kpc (Kartaltepe et al. 2007). With respect to the average merger time obtained from simulations<sup>[CP]</sup>, the ensemble of Sbc mergers has the majority of runs starting at 11 kpc. A few runs have much greater initial radius ranging from 44 to 50 kpc. Some of the merger times in table 3.2 also apply such large distances (de Ravel et al. 2009), indicating that the ratio of expected versus observed merger time is consistent with prior parameters. Extrapolating data from Kartaltepe et al. (2007) provides table 3.3. These fractions are applied with proper time to determine the redshift dependence of major mergers.

Table 3.3. Merger fractions with respect to redshift

Redshift	$f_m$	Redshift	$f_m$
0.15z	0.012	0.75z	0.045
0.25z	0.018	0.85z	0.054
0.35z	0.025	1.00z	0.073
0.45z	0.030	1.15z	0.102
0.55z	0.034	1.30z	0.129
0.65z	0.039	-	-

Relative to the local space, merger fractions as determined from morphology require that 3% of local galaxies are merger remnants<sup>[CS]</sup>. Averaging several local pair studies places the merger fraction at 0.018, while Kartaltepe et al. (2007) provide 0.007. Relative to the purposed model, one merger on average at 1.4z should also include mergers that have already occurred prior to the local space. This requires that  $3\% \pm 1.8\%$  of the local population has already undergone mergers, which is insignificant. To include merger fractions beyond 1z, the data was extrapolated with an average between several surveys. These vary between 0.19 and 0.22 at 2.5z, with a maximum fraction for any reference being 0.30 at 2.0z. Applying these additional constraints with respect to the continuous model provides the relative amount of galaxies for a particular redshift in figure 3.16. The one major merger per galaxy occurs at a redshift of 1.2z instead of the previously referenced 1.4z. The 4.4 — 5.4x disagreement with expected merger times makes it difficult to plot the amount of galaxies with respect to  $\Lambda$ CDM. The normalized distribution of proper time between models however is nearly proportional at low z. The 3x abundance around 0.8 — 1.0z would therefore decrease by 32%, i.e.  $\Lambda$ CDM is off by  $104\% \pm 25\%$  in effective number density. The 2x abundance at 0.5z indicates an error of  $70\% \pm 15\%$ .

The ratio of scale-factors can be applied to the normalized number density in order to determine the expected abundance of FBG, depicted in figure 3.17. The 2x abundance at 0.5z is in agreement with prior constraints including lack of drastic merger fractions and minimal luminosity evolution. The abundance peaks beyond the observed 2 – 3x disagreement relative to  $M_B = [22.5, 24]$ . Apparent magnitudes of FBGs versus redshift were further predicted from the absolute B-band distribution of local blue galaxies. With the continuous model and observed mergers, 22.5  $M_B$  and 24.0  $M_B$  correspond on average to 0.54z and 1.01z respectively.

Although observations such as baryonic matter cooling with increasing redshift can be blamed on various hypothesis, there is no answer for the 2x to 3x abundance of FBGs. Combined with luminosity characteristics, they should be nearly the same size as local late-types. However, the sizes inferred from  $\Lambda$ CDM are 2 — 5x smaller than would be expected. The nearly equal disagreement between both of these aspects must be due to improper curvature of the universe. Occam's razor alone would support this conclusion; however, all viable explanations have also been ruled out.

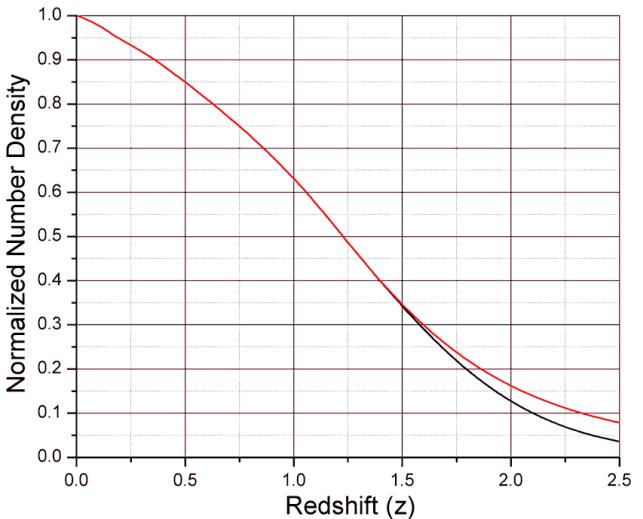


Figure 3.16. Number density of galaxies with mergers only.

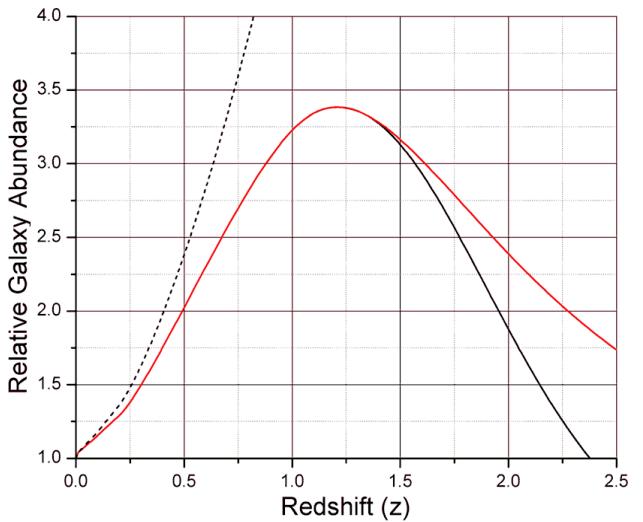


Figure 3.17. The expected abundance of galaxies with respect to no evolution  $\Lambda$ CDM is represented with a dashed line. Red corresponds to the continuous model with the merger fractions provided in table 3.3 and  $\langle r \rangle = 2.88 \text{ Gyr}$ . Black is maximum merger fractions extrapolated beyond 1.3z[CM].

### 3.6. Metallicity

Nuclear entropy is also useful for differentiating between the purposed model and  $\Lambda$ CDM. There are several properties of galaxies that can determine the redshift dependent evolution of metallicity. The focus therefore changes from that of large-scale curvature back to time dependence versus redshift.

$\Lambda$ CDM or an expanding model requires galactic age to decrease with distance, while the continuous model predicts the opposite. Time-scales involved were also found to agree with the purposed model after comparing simulations to observed merger times. FeII:MgII ratios and variations in magnesium within galaxies further provide strong evidence for the continuous model. Since the models provide opposite predictions for time dependence, evidence for the purposed model is evidence against  $\Lambda$ CDM.

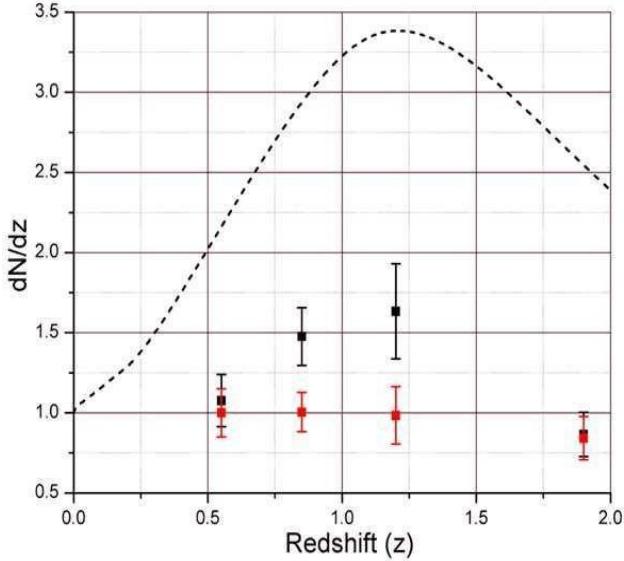
The evolution of metallicity can be inferred from galactic morphology, which trends from early-type in the local space to late-type for the more distant galaxies. Early-type galaxies consist of ellipticals and lenticulars, usually in clusters containing large amounts of hot, x-ray emitting intercluster medium (ICM). Clusters such as Abell 1367 demonstrate how galaxies contained within regions of hot ICM lack active star formation, i.e. they are considered to be passive<sup>[CT]</sup>. The majority of galaxies on the perimeter of the ICM region are either active or starburst. Similar clusters such as Abell 1656 are common in the local space, containing primarily early-type galaxies and a few late-type populations. The observed epoch of intense star formation is from 0.5z to 3z, in agreement with x-ray emitting gas inhibiting stellar formation. The ISM is also about 3.8x cooler by 1.2z relative to local galaxies. It was previously discussed how lyman-alpha blobs have dimensions similar to local x-ray emitting clusters. An expanding model however provides the wrong dimensions of these objects by several orders of magnitude. The lyman-alpha blobs are instead massive clusters that have cooled over  $\geq 50$  Gyr. All of these observations agree with the continuous model and purposed revisions to galactic evolution.

Local merger fractions are too insignificant to play a role in the thermal or nuclear entropy of galaxies prior to 1.0z. Even with a major merger, the nuclear entropy of a galaxy cannot drastically vary. Early-type galaxies for example contain an abundance of population II stars, which are metal poor. Middle-aged spiral galaxies on the other hand contain a mixture of population II and I stars. These galaxies are not metal poor, containing dust and much less x-ray emitting gas. A collision between two disk galaxies may increase the temperature of ISM; however, it will not reverse nuclear entropy or eliminate prior population I stars. There is also no clear transition from disk to elliptical galaxies in surveys. NGC 6240 is a good example of a major merger between two disks. It is similar to the major mergers occurring with red and ultra-red galaxies in the HUDF. These red populations have abundant dust, which is not similar to the x-ray emitting gas in local elliptical galaxies. They instead have very luminous cores with tails or peculiar shapes. Time-dependence of galactic entropy clearly disagrees with  $\Lambda$ CDM when considering observations.

From moderate to distant redshift (2 – 4z), the purposed model predicts 60 – 110 Gyr of evolution from the local space. Therefore, dense regions of intense star formation should begin to deplete primordial hydrogen and helium. This however does not imply that regions without star formation will become metal rich, i.e. distant galaxies or clusters are usually embedded within regions of reionized hydrogen. Galactic star formation will also depict the evolution of ISM. Type II supernova events are associated with metal poor stars, where the final Fe:Mg ratio is about 1.65x. Type Ia events are more metallic, producing Fe:Mg ratios far above type II events (393x). The size and degeneracy of stars also play roles in the type of supernova. For example, type Ia events are inferred to occur from massive stars and are much more energetic than other types. The location of SNIa are also consistent with the continuous model, which uses them to track the flow towards central core.

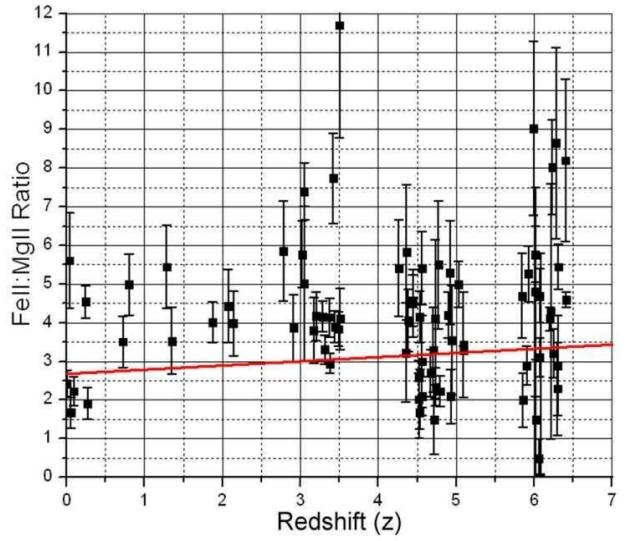
The iron and magnesium concentrations in ISM play a crucial role in determining the relative age of distant galaxies. Nuclear entropy naturally favors the production of iron over long time-scales, with magnesium slightly lower with respect to nuclear potential. Both are created from nuclear fusion and should be subjected to similar environments. For example, any mechanism besides nuclear fusion that varies FeII will proportionally vary MgII. Since rotational curves are observed to be unstable, it is unlikely that supermassive black holes at the center of galaxies will substantially vary the surrounding metallicity. Therefore, distant galaxies are expected to have increased metallicity and higher FeII:MgII ratios relative to similar local populations.

Prior to discussing FeII:MgII ratios, evidence of increasing metallicity can be inferred through other methods. The absorption of distant sources by local galaxies demonstrates an abundance of MgII. These galaxies are observed from the local space up to about  $0.9z^{[CU]}$ . Weak MgII absorbers are further observed in abundance from 0.7 to 2.2z, while vanishing beyond  $2.7z^{[CV]}$ . With observed merger fractions and continuous model, the relative amount of weak MgII absorbers is depicted in figure 3.18.



**Figure 3.18.** The dashed line represents the inferred abundance with mergers from section (3.5). Black indicates the averaged  $dN/dz$  from  $\wedge$ CDM surveys and the continuous model with evolution is in red.  $dN/dz$  in this case refers to the number density multiplied by the proper geometric cross section<sup>[CV]</sup>.

The spectrum of quasars also provides valuable information relevant to metallicity. Quasars are some of the most luminous and distant objects in the observable universe. Similar to SNIA, quasars display characteristics of highly degenerate matter. These however are related to active galactic nuclei, which harbor supermassive black holes. Relative to the continuous model, highly degenerate objects are older and therefore in the flow towards central core. Several studies were combined in figure 3.19 to determine the evolution of FeII:MgII with respect to redshift<sup>[CW][CX][CY][CZ]</sup>. The high ratios observed at 6z are indicative of galaxies that have already undergone intense star formation<sup>[CZ]</sup>. For the ISM to be enriched with an abundance of these elements, several generations of stars must have undergone supernovas. Relative to  $\wedge$ CDM, the proper age of the universe is less than 1 Gyr at 6z. Considering that the region of intense star formation is observed from  $0.5z$  to  $3z$ ,  $\wedge$ CDM does not fit observations. Other observations such as increasing cold baryonic matter with redshift, number densities, galactic evolution and the dark flow agree with galactic age increasing with redshift. An expanding universe on the other hand would violate several fundamental laws of physics including the second law of thermodynamics and nuclear entropy.



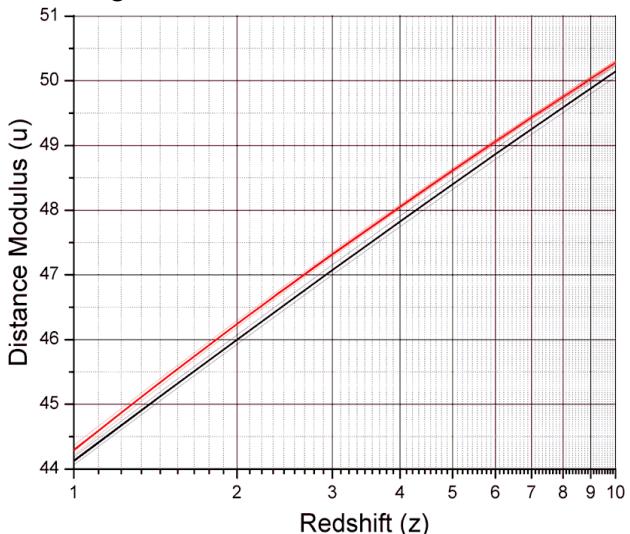
**Figure 3.19.** High FeII:MgII ratios at extreme redshift are indicative of older galaxies. A statistically significant trend is also observed with a slope of  $0.108 \pm 0.03$ .

### 3.7. Statistical Analysis

Although an expanding universe is conclusively ruled out from incorrect predictions of large-scale curvature, it is possible to compare redshift versus distance modulus between  $\Lambda$ CDM and the purposed model. Distinction between models at low redshift arises in the form of dispersion due to directional dependence. This is predicted by the continuous model from varying  $y_0$  over its complete range (0 Gpc to  $> 0.54$  Gpc). Relative to section (3.2), the slope of the universe and average y-intercept are determined from type Ia supernova (SNIa) and gamma ray burst (GRB). SNIa are superior for determining cosmological distances due to their nearly uniform properties. GRB are less reliable, but can still be used to constrain redshift versus luminosity distance. SNIa and GRBs both display characteristics of highly degenerate matter, which is an indication of relatively older galaxies. These events will therefore be statistically more abundant along the flow towards central core, where the continuous model predicts for objects to be older than the local group.

With respect to Occam's razor, the purposed model fits the shape of the universe with a single constant ( $S_0$ ).  $\Lambda$ CDM usually needs two constants in the form of dark energy and matter. Dark energy however cannot be directly detected and has no connection to the standard model. It is inferred to exist solely because it allows an expanding model to match observations. Although these non-classical modifications fit redshift versus distance modulus, they fail to agree with the observed shape of the universe. Occam's razor is therefore a necessity for arriving at the proper theory, as anyone can force a model to agree with observations by introducing purely mathematical constructs. Failure to reach parsimony and over-reliance on confirmation rather than refutation are dangerous practices for this reason. From these aspects alone, the models cannot be put on equal footing. The continuous model contains the least amount of free variables and non-classical assumptions.

$\Lambda$ CDM and the continuous model have similar redshift versus distance modulus predictions from 0.5z to about 10z. Disagreement between models in this region peaks at  $0.25\mu$  around 2.5z, making it difficult to differentiate between the two.  $\Lambda$ CDM is constrained by an interpretation of the CMBR and baryon acoustic oscillation data<sup>[BN]</sup>. These provide  $H_0 = 70.4^{+11.3}_{-11.3}$ ,  $\Omega_b = 0.0456 \pm 0.0035$ ,  $\Omega_c = 0.222 \pm 0.026$  and  $\Omega_\Lambda = 0.728^{+0.16}_{-0.16}$ . Figure 3.20 depicts an ensemble within these limits for  $\Omega_m = \Omega_b + \Omega_c$  and  $\Omega_A = 1 - \Omega_m$ . The purposed model applies the best fitting slope and y-intercept, including uncertainties previously provided in section (3.2). Since SNIa are considered standard candles, observations from 0.1z to 0.5z can easily rule out an expanding universe. The disagreement for local redshift once again arises from directional dependence due to an asymptotically flat universe. The majority of events beyond 1.0z consist of GRB, which suffer from circularity problems. In other words, most methods require that a prior cosmological model be selected in order to determine the luminosity distance to GRB sources. Attempts to avoid this problem apply various relations between GRB parameters and extrapolating SNIa data. However, some of these methods are also found to be model dependent and should not be used to independently determine cosmological distances.

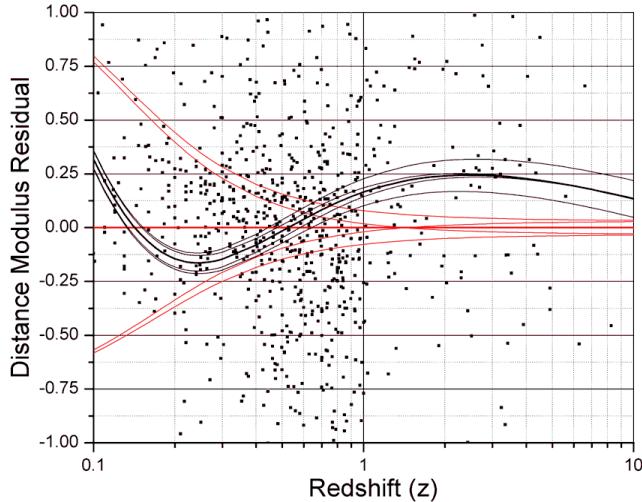


**Figure 3.20.** Red is the continuous model within limits of error, while black is  $\Lambda$ CDM within limits of error.

A more detailed comparison between models can be achieved from the distance modulus residual with respect to the continuous model. However, it is determined that the dataset obtained from NED redshift independent calculations was contaminated with respect to a few GRBs. The problem arises due to the lack of standardization for GRB sources and inclusion of fiducial data; this is apparent in figure 3.21 from 2.0z to 5.0z. Prior to 1.0z, uncertainty is too small to be fit with homogeneous expansion. For example, a portion of data is outside of either models best fit by more than  $\pm 0.50\mu$ . However, the trends for the continuous model in figure 3.21 only include uncertainty in slope and average distance to the start of flow towards central core. Considering the shortest path instead begins at 0.0 Gpc and others at distances greater than  $y_0$ , the continuous model explains this dispersion relatively well. The purposed model is also centered on the bulk of available SNIa data prior to 1z indicating a superior fit. To the contrary, homogeneous expansion does not fit observations from 0.1z to 0.5z. Extrapolating SNIa data under the assumption of  $\Lambda$ CDM will therefore produce incorrect predictions for GRBs beyond available SNIa data.

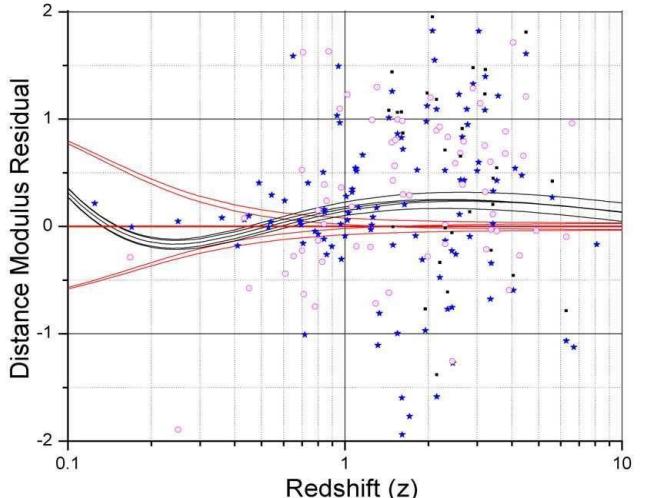
**Table 3.4.** Average SNIa/GRB error versus redshift

Redshift	Avg. Error	Redshift	Avg. Error
0.1 – 0.5z	$\pm 0.21\mu$	1.0 – 2.0z	$\pm 0.69\mu$
0.5 – 1.0z	$\pm 0.27\mu$	$\geq 2.0z$	$\pm 1.11\mu$



**Figure 3.21.** A plot of the combined SNIa/GRB dataset used to determine the slope of the universe and average y-intercept.

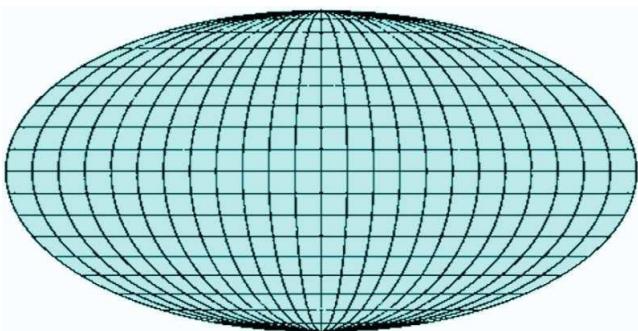
Although the dataset applied from the NED database is claimed to be redshift independent, some of the data is fiducial; i.e. metric distance is determined assuming the big bang model is correct. For example, there are several data points in figure 3.21 that are anomalously clustered around the  $\Lambda$ CDM trend from 2.0 – 5.0z. To demonstrate this, the residual is plotted in figure 3.22 with respect to several studies. As would be expected from an average error of  $\pm 1.11\mu$  beyond 2.0z, the anomaly no longer exists. After applying solely SNIa events for all available redshift, the slope ( $3.216 \cdot 10^{42}$ ) and average y-intercept (0.313 Gpc) are still within previous limits. Several contaminated data points are related to improper methods. GRB 051109A for example contains 3 of 8 values from 2009MNRAS. Removing these from the average increases distance modulus to  $46.426\mu$ , with the continuous model predicting  $46.429\mu$ . Issues with these references also include the circularity problem and use of the Amati relation. The Amati relation is found to suffer from selection effects and should not be used to probe distance<sup>[DD]</sup>. This is applied in 2010JCAP, although the data is in better agreement with SNIa. 2009MNRAS extrapolates only 55% of SNIa data with fiducial methods. This systematically offsets the 2009MNRAS GRBs with respect to the more accurate SNIa dataset between redshift of 0.015z and 1.55z.



**Figure 3.22.** Several sets of gamma ray burst are plotted from ★ 2010JCAP<sup>[DA]</sup>, ○ 2009MNRAS<sup>[DB]</sup> and ■ 2009EPJC<sup>[DC]</sup>.

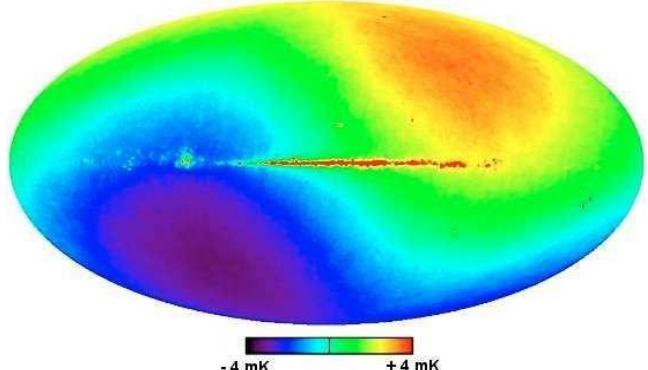
### 3.8. The Cosmic Background Radiation

The big bang theory claims that the CMBR is due to a period of recombination after the initial creation of space-time. At this point, the universe cooled until space became transparent to free photons. The big bang theory also claims that the observed blackbody radiation in all directions of space is not a free electromagnetic field, but instead localized packets of electromagnetic energy. To understand why this is not true, the source of blackbodies must be understood. All finite objects with a temperature will emit a spectrum of radiation that peaks at a given wavelength. When an object emits this blackbody spectrum, it is due to the internal kinematic energy or temperature. The free field emitted from massive objects therefore obeys a statistical distribution of internal energy, which is released at the surface boundary. Converting the observed CMBR temperature as depicted by figure 3.23 into the relative value at emission, the core's surface temperature is 3000 K. In comparison, the Sun's surface has a temperature of about 5778 K, indicating that the core likely consists of dense quark matter. Relative to an Einstein black hole with event horizon, the surface will theoretically emit black body radiation in the form of Hawking radiation. The temperature is inversely proportional to mass with a coefficient of  $6.1686 \cdot 10^{-8} M_{\odot} \cdot K$ . A 3000 K central core with respect to an Einstein black hole would have a mass of  $2.056 \cdot 10^{-11} M_{\odot}$ , compared to  $3.694 \cdot 10^{-8} M_{\odot}$  for the Moon.

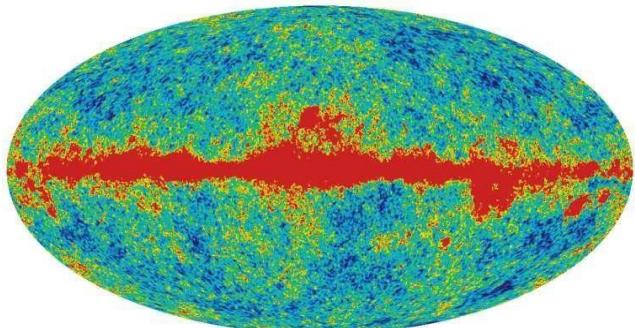


**Figure 3.23.** The locally observed cosmic background radiation with a temperature of approximately  $2.725 K$  after redshift.

The CMBR shows peculiarities such as a dipole moment, large-scale bulk flows and additional fluctuations from interaction with external matter. The CMBR temperature prior to subtracting the average value is  $2.72548 \pm 0.00057 K^{[BC]}$ . Figure 3.24 depicts the dipole moment observed after subtracting the average temperature; figure 3.25 is the CMBR with the dipole subtracted. The dipole moment is due to Earth's motion relative to the CMBR source or the core's surface. Since a spherical body will emit blackbody radiation at a nearly constant  $z$  or radius, only the Doppler Effect applies. Relative to the source of the CMBR, the solar system is moving at  $369.0 \pm 2.5 \text{ km} \cdot \text{s}^{-1}$  towards  $(l, b) = (264.26^{\circ} \pm 0.33^{\circ}, 48.22^{\circ} \pm 0.13^{\circ})^{[AZ]}$ . Taking into account the local group's motion, this velocity becomes  $627 \pm 22 \text{ km} \cdot \text{s}^{-1}$  towards the direction  $(l, b) = (276^{\circ} \pm 3^{\circ}, 30^{\circ} \pm 3^{\circ})^{[AZ]}$ .

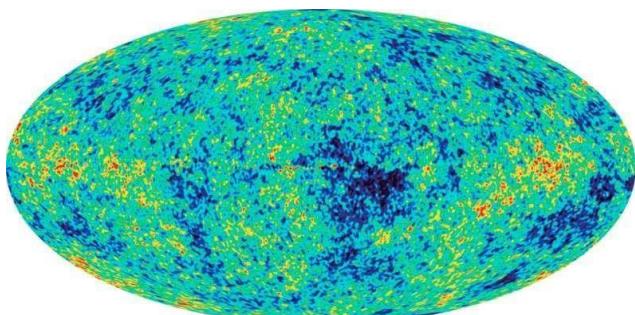


**Figure 3.24.** After subtracting the average temperature from the sample measured by COBE, the dipole dominates. **Image credited to NASA/WMAP Science Team<sup>[BD]</sup>**



**Figure 3.25.** After subtracting the dipole moment from figure 3.24, the remaining fluctuations in the CMBR occur externally from the core. Image is provided from WMAP (2003). **Image credited to NASA/WMAP Science Team<sup>[BD]</sup>**

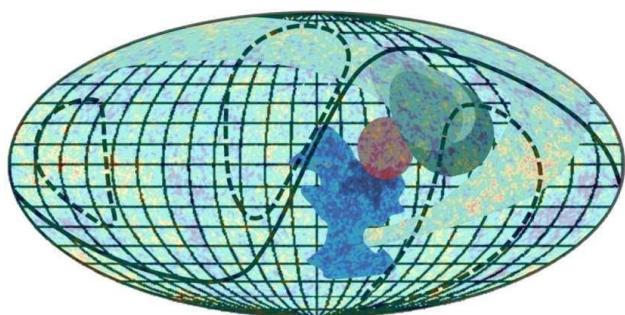
Remaining fluctuations in the CMBR are from the Milky Way and scattering of electromagnetic by matter external to the core's surface. Figure 3.26 is the CMBR after removing local foreground sources; it depicts two hot stripes and a central cold patch. The source of the CMBR itself should be at a nearly constant temperature of 3000K. Variations within the cleaned CMBR image instead occur between the foreground and background due to scattering from the x-ray emitting gas of massive clusters. The relativistic charged particles boost the black body spectrum to higher energy levels. This increases the observed black body temperature from deep blue to green and red. Analysis of x-ray emitting clusters has also shown a statistically significant bulk flow extending from the local group to  $\sim 0.77$  Gpc<sup>[AZ]</sup>. The velocity is estimated to be [600, 1000] km s<sup>-1</sup> from the thermal S-Z effect<sup>[AZ]</sup>; however, free field radiation undergoes Thomson scattering. Although many of the directions in local space lead back to the central core, there must logically be a flow of younger galaxies and clusters into Earth's present region. This is necessary to remain consistent with the foundations used in deriving non-local redshift, i.e. the universe is in a steady state. With the medium of galaxies and clusters progressing from hot x-ray emitting gas into cold metallic dust, the dark flow should consist of relatively younger clusters.



**Figure 3.26.** The cleaned CMBR is observed to contain a large hot strip originating from a central cold patch. An annihilation boundary or hot ring is surrounding the central cold patch. The base of the local jet should be visible, with the hot strip to the right being a continuation of the dark flow at extreme redshift. **Image credited to NASA/WMAP Science Team<sup>[BD]</sup>**

Conventional theory attributes the CMBR to an epoch of recombination, where photon decoupling takes place. This explanation is only valid for an expanding universe, which has a specific shape and angular scale. Section (3.5) however proved that both galaxy number densities and angular scales are incompatible with an expanding universe. This is easily observed from 0.3z to 0.7z, becoming more drastic beyond. Since an expanding model can be conclusively ruled out, current foundations for the CMBR are invalid. It was further demonstrated in sections (3.5, 3.6) that the proposed model fits the correct shape of the universe with respect to number densities, angular size and several other aspects.

The proposed or continuous model requires two polar jets originating from a central core in order to explain current observations of entropy. If the core acts as a mechanism for baryon asymmetry, an annihilation boundary should also be observed between hemispheres. The inferred local jet is depicted in figure 3.27. Anomalies may distort the cleaned image, possibly beyond use in some regions. The elliptic plane for example runs through the far right side of the local jet, which on average becomes cooler than the surrounding areas. It also correlates with zodiacal dust and several features such as quadrupole/octupole alignment and the cold strips or “fingers” in the southern hemisphere<sup>[BI]</sup>.



**Figure 3.27.** The dark flow is depicted in shades of gray, where lighter shades are more distant<sup>[BK]</sup>. Since SNIa are used to measure the flow towards central core, an observed SNIa bulk flow is in red<sup>[BL]</sup>. Hemispherical power asymmetry<sup>[BJ]</sup> could be due to the local jet, where temperature and resolution variations are more extreme closer to the central region.

### 3.9. Baryon Asymmetry

Section (3.5) demonstrated that an expanding universe is incompatible with observations. The previous sections also provide sufficient evidence for an asymptotically flat universe with central core. It is therefore important to discuss the consequences of having an always existent, steady state universe in terms of entropy and stability. The concept of an absolute beginning of reference time is flawed. Reference time in this perspective is relative to a space-time with no vacuum energy. For example, it is known that metric distance becomes infinite as an event horizon is approached with respect to the preferred reference frame. Therefore, nothing can reach the surface boundary within a finite amount of reference time. Relative to proper time, an observer falling into an event horizon would do so in a finite amount. This is an illusion since the observer will cease to evolve as the surface is approached. In comparison to a photon, the problem can still be defined in terms of metric distance. Since the photon will be travelling along a null geodesic, it must travel an infinite amount of metric distance prior to reaching the surface.

It is possible to take the limit of the inferred state of an expanding universe as  $t \rightarrow -\infty$ . With either Einstein's field equations or vacuum field theory, matter will converge at a single region in space until it is infinitesimally close to forming an event horizon. With respect to the preferred reference frame, a discontinuity forms. For any event horizon, time is undefined due to infinite vacuum energy density. From the other direction, the collapsing system cannot form into a conical singularity within a finite amount of reference time. This discontinuity between time prior to a big bang scenario and formation of event horizon indicates that an initial singularity could have never existed with respect to the reference frame. On the other hand, one could argue that the universe began at some infinite time

ago as an object infinitesimally close to a conical singularity; i.e. the interpretation of accelerated expansion is still viable. If this were true, Hubble's law would remain valid for all redshift rather than just the local. In addition, this perspective requires galaxies to become older as redshift increases, which is contrary to several recent observations. The most important of these is the increase in cold baryonic matter with redshift. Even beyond this, an expanding theory requires several non-classical assumptions such as dark matter and dark energy. It fails to explain observations of galactic evolution, number densities or the size of distant galaxies and clusters. Therefore,  $\Lambda$ CDM or an expanding model should be abandoned.

An always-existent universe has other properties based upon boundary conditions. If at some time prior to the present the universe was unstable and energy could escape, the instability would have existed at some prior point in time. With respect to a quantum system, there will be a distribution of possible events occurring over a finite period. If the system has a finite probability distribution, then it is impossible for any instability to have not occurred prior to a finite time before present. In other words, any instability must have existed for an infinite period and according to probability, any finite distribution requires that the universe is stable at  $t \rightarrow -\infty$ . This also relates to the null existence of event horizon. For example, all observations agree with a central core existing in the present universe. If black holes had event horizon as predicted by EFEs, the universe would clearly not exist in the current state. As matter approaches a central event horizon, some would be captured while the rest is ejected. Over an infinite amount of reference time prior to the present, the central event horizon would capture all matter within the universe. The CMBR temperature further provides direct proof against the existence of event horizon and Hawking radiation.

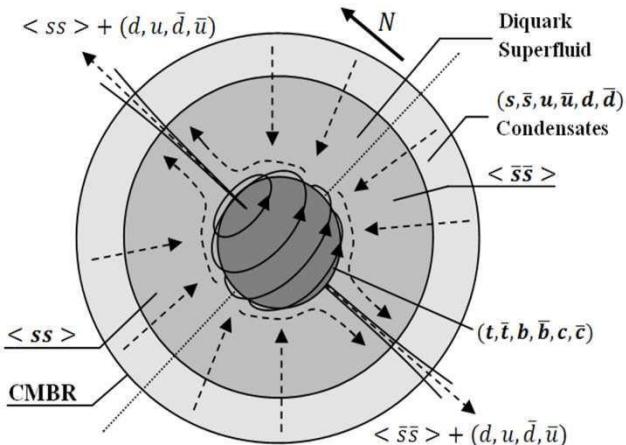
In order for an always existent universe to be in its present form, the laws of thermodynamics must be missing something. The last piece of the puzzle was previously beyond comprehension due to the belief in an expanding universe and event horizon. Since all massive objects must have finite vacuum fields, the gravitational force aids in completing the thermodynamic cycle. The entire process can be viewed as beginning from the surface of the central core in the form of a dense relativistic jet. From this point to Earth's present position in the universe, the usual thermodynamic principles apply. Dense quark matter for example will decay into x-ray emitting gas. Radiation emitted over this transition follows geodesics back to the central core, ensuring the universe is stable. However, an asymptotically flat universe by definition will have finite vacuum energy density at all points in space. The only requirement is that the universe remains localized for an infinite period, i.e. it exists in a steady state. Beyond Earth's present position, population I stars are abundant due to increased metallicity and cold baryonic matter. The bulk flow continues to move towards entropy as galaxies gain momentum falling into the center of the universe. The missing piece is where matter falls back into the central core and momentum is conserved via two polar jets. The universe therefore exists in an anisotropic state of entropy as depicted by figures 3.5 and 3.7.

The laws of thermodynamics demand that the entropy of a closed system can never be reversed without external energy. From figure 3.5, it is clear that entropy is constant for all time relative to the preferred reference frame. Therefore, the laws of thermodynamics are not violated since the universe acts as a perpetual machine. The actual mechanism that creates relativistic jets is speculated upon with QCD and modern MHD simulations. Compared to the local region of space that Earth currently resides in, the jets emanating from the central core should

be extremely large. With respect to the amount of galaxies and clusters falling into the central core, it likely contains the mass of millions or billions of galaxies and clusters. The center of the core should therefore exist in a dense, color superconducting state. The central region of the core is assumed to consist of top, bottom and charm quarks due to sheer size. With respect to modern theory, compact stars are already predicted to exist in a non-CFL color superconducting state<sup>[BF]</sup>. The layer directly adjacent to the core's central region will also likely exist in a superfluid state (CSL, Planar, A/Polar). The center could possibly rotate, further inducing a magnetic field from the London moment.

Comparing the bare mass of quarks, the core should have layers depicted by chemical potentials. For example, the up and down quarks exist within the 2 — 15 Mev range. The strange quark has a bare mass between 100 — 300 Mev, while the remaining quarks exist from 1000 Mev and beyond. Baryon asymmetry is speculated to originate from the strange quark/anti-quark layer. Considering that the core acts as a perpetual machine, the energy needed to produce the jets must be provided by the inflow alone. Due to the asymmetric shape of the universe, matter enters the equatorial regions and is funneled inwards. As density increases, the quarks at each radius become more massive until a strange diquark layer is reached. This type of condensate should be favored due to the gap between quark masses and single quark flavor. The CSL, planar or A/polar single flavor states also demonstrate superfluidity and the Meissner effect<sup>[BE]</sup>. As matter approaches the center, the magnetic field energy density begins to surpass kinematic energy density<sup>[AM]</sup>. At this point, it is expected that the inflow stops moving with the bulk of material, effectively producing a toroidal magnetic field<sup>[AM]</sup>. Due to lorentz forces, the  $\langle ss \rangle$  condensates are accelerated in one direction with  $\langle s^- s^- \rangle$  in the opposite.

Putting things into perspective, the structure of the core should be similar to figure 3.28. After applying the methods discussed in section (3.3) to derive proper time, matter approaching the surface of the core will be traveling at 99.999988% the speed of light. In consideration of momentum conservation and the internal magnetic field, the rate of inflow alone is capable of explaining the energy behind each jet. The bulk of material ejected should originate from the strange diquark layer due to its insulating property. As the quark matter begins to cool into mesons and baryons, the jet pointing in the direction of  $T$   $N$  will consist of  $K^0$ ,  $K^-$ ,  $\pi^\pm$ ,  $\pi^0$ ,  $p^+$ ,  $n^0$ ,  $A^0$ ,  $\Sigma^\pm$ ,  $\Sigma^0$ ,  $\bar{\nu}^-$  and  $\bar{\nu}^0$ ; although not exclusively. These pairings occur naturally since baryons containing a mixture of quarks and anti-quarks have not been observed experimentally. However, the anti-quarks must still be paired with something external to the surface. The resulting mesons decay into photons, electrons and neutrinos; table 3.5 provides the most common decay modes. The amount of material ejected from each shell of the core depicts the ratio of electrons and neutrinos to baryonic matter, although the composition of each layer must also be known.



**Figure 3.28.** A simplified section of a central core that would induce baryon asymmetry. The focus is placed upon the strange diquark layer, which is theorized to exist due to the gap in mass between quarks. In a steady state model, the core must have a constant inflow and outflow. This process creates an abundance of  $(p^+, e^-)$  in one hemisphere and  $(p^-, e^+)$  in the other.

The decay of unstable baryons in the northern hemisphere of the universe also results in products that are observed in abundance locally. Table 3.6 provides the common decay modes for the main constituents. This process creates an abundance of protons and electrons ( $p^+$ ,  $e^-$ ) in the northern hemisphere, with an abundance of anti-matter ( $p^-$ ,  $e^+$ ) in the other. It is concluded that Earth currently resides in the northern hemisphere beyond the point where the relativistic proton, electron and neutrino gas has cooled. For any steady state model in the current universe, a perpetual machine and explanation for the CMBR are required. With the corrections to general relativity, event horizon are no longer possible. Therefore, an asymptotically flat universe containing a central core would emit a 3000 K black body spectrum as observed. The purposed configuration further explains the hot ring surrounding the central cold spot in the cleaned CMBR image, the mechanics behind a steady state and origin of the dark flow.

**Table 3.5.** Common decay modes of mesons and fermions<sup>[BG]</sup>.

Particle(s)	Decay Mode(s)	Particle(s)	Decay Mode(s)
$n^- (\bar{u}d)$	$(\mu^- + \bar{\nu}_\mu),$ $(e^- + \bar{\nu}_e)$	$n^+ (u\bar{d})$	$(\mu^+ + \nu_\mu),$ $(e^+ + \nu_e)$
$n^0 \left( \frac{u\bar{u} - d\bar{d}}{\sqrt{2}} \right)$	$2\gamma$	$K^0 (ds)$	$A^0 + \gamma$
$K^- (\bar{u}s)$	$(\mu^- + \bar{\nu}_\mu),$ $(n^- \pm n^0),$ $(e^- + \nu_e + \nu_\mu)$	$\mu^-$	$e^- + \bar{\nu}_e + \nu_\mu$
		$e^+ + e^-$	$2\gamma$

**Table 3.6.** Common decay modes of baryons<sup>[BG]</sup>.

Baryon	Decay Mode(s)	Baryon	Decay Mode(s)
$p^+ (uud)$	Stable	$\Sigma^0 (uds)$	$A^0 + \gamma$
$n^0 (udd)$	$p^+ + e^- + \bar{\nu}_e$	$\Sigma^- (dds)$	$n^0 + n^-$
$A^0 (uds)$	$(p^+ + n^-),$ $(n^0 + n^0)$	$\Sigma^+ (uus)$	$(p^+ + n^0),$ $(n^0 + n^+)$
$\bar{\nu}^0 (uss)$	$A^0 + n^0$	$\bar{\nu}^- (dss)$	$A^0 + n^-$

## References

- [A] Edmund, Whittaker. “**A History of the Theories of Aether and Electricity from the Age of Descrates to the Close of the Nineteenth Century**”. Dublin, Ireland: Hodges, Figgis. 1910
- [B] Jackson, John. “**Classical Electrodynamics**”. New York: Wiley, 1999.
- [C] Griffiths, David. “**Introduction to Quantum Mechanics**”. Upper Saddle River: Pearson Prentice Hall, 2005.
- [D] Born, Max. “**The Statistical Interpretation of Quantum Mechanics**”. Nobel Lecture, December 11, 1954. Link (^)
- [E] Dirac, Paul. “**Theory of Electrons and Positrons**”. Nobel Lecture, December 12, 1933. Link (^)
- [F] Hestenes, David. “**The Zitterbewegung Interpretation of Quantum Mechanics**”. 1990. Link (^)
- [G] Hestenes, David. “**Real Dirac Theory**”. 1996. Link (^)
- [H] Pollock, M.D. “**The Dirac Equation in Curved Space-Time**”. March 11, 2010. Link (^)
- [I] Gaberdiel, Matthias; Ridder, Aude. “**Quantum Field Theory II**”. July 15, 2011. Link (^)
- [J] Ellis, John. “**Standard Model of Particle Physics**”. Link (^)
- [K] Grünewald, Martin. “**The LEP Electroweak Working Group**”. Link (^)
- [L] “**End of the Line for LEP**”. November 8, 2000. Link (^)
- [M] “**God particle may not exist**”. December 6, 2001. Link (^)
- [N] “**Higgs hunting at 144GeV**”. July 27, 2011. Link (^)
- [O] “**Higgs signal sinks from view**”. August 22, 2011. Link (^)
- [P] “**Detectors home in on Higgs boson**”. December 13, 2011. Link (^)
- [Q] “**Physicists find new particle, but is it the Higgs?**”. July 3, 2012. Link (^)
- [R] Mammadov, Gulmammad. “**Reissner-Nordstrom Metric**”. May 4, 2009. Link (^)
- [S] Stoica, Ovidiu-Cristinel. “**Analytic Reissner-Nordstrom Singularity**”. April 19, 2012. Link (^)
- [T] Taylor, Joseph. “**Binary Pulsars and Relativistic Gravity**”, Nobel Lecture, December 8, 1993. Link (^); Huse, Russell. “**The Discovery of the Binary Pulsar**”, Nobel Lecture, December 8, 1993. Link (^)
- [U] Brady, Patrick. “**Astrophysical Sources of Gravitational Radiation**”. 1999. Link (^)
- [V] Belczynski, Chris. “**Double Compact Object Mergers: Short-hard GRBs and Gravitational-wave Signals**”. 2007. Link (^)
- [W] Dent, Thomas. “**Searching for Inspiring and Merging Binaries in LIGO-Virgo Data**”. 2011. Link (^)
- [X] Belczynski, Chris. “**Double Black Holes: Recent Observations and Predictions**”. March 24, 2011. Link (^)
- [Y] Brown, Duncan. “**Searches for Gravitational Waves from the Inspiral of Binary Neutron Stars and Black Holes**”. April 13, 2008. Link (^).
- [Z] Gondek-Rosinska, Dorota. “**Perspective of Discovering Gravitational Waves from Astrophysical Sources**”. 2011. Link (^)
- [AA] Nikhef, Jo van den Brand. “**Gravitational Wave Detection**”. June 3, 2009. Link (^)
- [AB] Coward; Lilley; Howell; Burman; Blair. “**The ‘Probability Event Horizon’ for Neutron Star Merger Detection with Advanced LIGO**”. 2004. Link (^)
- [AC] Corvino; Ferrari; Marassi; Schneider. “**Compact Binaries Detection Rates from Gravitational Wave Interferometers: Comparison of Different Procedures**”. June 18, 2012. Link (^)
- [AD] Waldman, Samuel. “**Status of LIGO at the Start of the Fifth Science Run**”. June 2006. Link (^)
- [AE] The LIGO Scientific Collaboration and The Virgo Collaboration. “**Sensitivity Achieved by the LIGO and Virgo Gravitational Wave Detectors during LIGO’s Sixth and Virgo’s Second and Third Science Runs**”. March 15, 2012. Link (^)
- [AF] Rolland, L. “**The Status of Virgo**”. January 2009. Link (^)
- [AG] Fafone, Viviana. “**VIRGO: Where We Come From, Where We Are Going**”. June 7, 2011. Link (^)
- [AH] Strain, Ken. “**The Status of GEO600**”. July 2007. Link (^)
- [AF] Che, Wujun; Paul, Jean-Claude; Zhang, Xiaopeng. “**Lines of Curvature and Umbilical Points for Implicit Surfaces**”. April 16, 2007. Link (^)
- [AG] Kremer; Devecchi. “**Thermodynamics and Kinetic Theory of Relativistic Gases in 2-D Cosmological Models**”. February 7, 2002. Link (^)
- [AH] Belongie, Serge. “**Rodrigues’ Rotation Formula**”. Mathworld, Wolfram. Link (^)
- [AI] Reiss; Macri; Casertano; Lampeitl; Ferguson; Filippenko; Jha; Li; Chornock. “**A 3% Solution: Determining the Hubble Constant with the Hubble Space Telescope and Wide Field Camera**”. March 10, 2011. Link (^)
- [AJ] This research has made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Link (^)
- [AK] Gamow, George. “**The Creation of the Universe**”. Mineola, N.Y: Dover Publications, 2004.
- [AL] Hubble, Edwin. “**A Relation Between Distance and Radial Velocity Among Extra-Galactic Nebulae**”. January 17, 1929. Link (^)
- [AM] Mirabel, Felix; Rodriguez, Luis. “**Sources of Relativistic Jets in the Galaxy**”. February 4, 1999. Link (^)
- [AN] Riess; Strolger; Casertano; Ferguson; Mobasher; Gold; Challis; Filippenko; Jha; Li; Tonry; Foley; Kirshner; Dickinson; MacDonald; Eisenstein; Livio; Younger; Xu; Dahlen; Stern. “**New Hubble Space Telescope Discoveries of Type Ia Supernovae at  $z \geq 1$ : Narrowing Constraints on the Early Behavior of Dark Energy**”. December 20, 2006. Link (^)
- [AO] Jarosik; Bennett; Dunkley; Gold; Greason; Halpern; Hill; Hinshaw; Kogut; Komatsu; Larson; Limon; Meyer; Nolta; Odegard; Page; Smith; Spergel; Tucker; Weiland; Wollack; Wright. “**Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results**”. January 26, 2010. Link (^)
- [AP] “**Encyclopedia of astronomy and astrophysics**”. Bristol Philadelphia London New York: Institute of Physics Pub. Nature Pub. Group, 2001.
- [AQ] Tacconi; Genzel; Neri; Cox; Cooper; Shapiro; Bolatto; Bouché; Bournaud; Burkert; Combes; Comerford; Davis; Förster Schreiber; García-Burillo; Gracia-Carpio; Lutz; Naab; Omont; Shapley; Sternberg; Weiner. “**High molecular gas fractions in normal massive star-forming galaxies in the young Universe**”. December 22, 2009. Link (^)

- [AR] Driver, Simon. "Hubble Deep Fever: A faint galaxy diagnosis". February 26, 1998. Link ([^](#))
- [AS] Ohio University. "Discovery Of Giant X-Ray Disk Sheds Light On Elliptical Galaxies". December 19, 2002. Link ([^](#))
- [AT] Sarazin, Craig; "X-ray Emission from Elliptical Galaxies". December 5, 1996. Link ([^](#))
- [AU] Dekel; Birnboim; Engel; Freundlich; Goerdt; Mumcuoglu; Neistein; Pichon; Teyssier; Zinger. "Cold streams in early massive hot haloes as the main mode of galaxy formation". January 16, 2009. Link ([^](#))
- [AV] Harvard-Smithsonian Center for Astrophysics. "Strange new 'species' of ultra-red galaxy discovered". ScienceDaily, December 1, 2011. Web. 14 Jul. 2012. Link ([^](#))
- [AW] NASA. "Hubble Finds Hundreds of Young Galaxies in Early Universe". April 2006. Link ([^](#))
- [AX] Smithsonian Astrophysical Observatory. "Mysterious Red Galaxies". December 9, 2011. Link ([^](#))
- [AY] Marel, Roeland; Dokkum, Pieter. "Dynamic models of elliptical galaxies in z=0.5 clusters: I. Data-model comparison and evolution of galaxy rotation". November 17, 2006. Link ([^](#))
- [AZ] Kashlinsky; Atrio-Barandela; Kocevski; Ebeling. "A measurement of large-scale peculiar velocities of clusters of galaxies: technical details". February 4, 2009. Link ([^](#))
- [BA] Clayton, D. (1983). "Principles of Stellar Evolution and Nucleosynthesis". Chicago: University of Chicago Press.
- [BB] Lehnert; Nesvadba; Cuby; Swinbank; Morris; Clement; Evans; Bremer; Basa. "Spectroscopic Confirmation of a Galaxy at Redshift z=8.6". October 20, 2010. Link ([^](#))
- [BC] Fixsen. "The Temperature of the Cosmic Microwave Background". November 30, 2009. Link ([^](#))
- [BD] Images are credited to NASA/WMAP Science Team. "WMAP Calibration". Link ([^](#)). "Three-Year WMAP View of Early Universe". Link ([^](#)). "WMAP Resolves the Universe". Link ([^](#))
- [BE] Shovkovy, Igor. "Transport Properties of Stellar Quark Matter". September 25, 2008. Link ([^](#))
- [BF] Bowers, Jeffrey. "Color Superconducting Phases of Cold Dense Quark Matter". 1998. Link ([^](#))
- [BG] J. Beringer (Particle Data Group), J. Phys. D86, 010001 (2012). Link ([^](#))
- [BH] Totani, Tomonori; Yoshii, Yuzuru. "Does the Number Density of Elliptical Galaxies Change at  $z < 1$ ?". May 20th, 1998. Link ([^](#))
- [BI] Bennett, C. L.; Hill, R. S.; Hinshaw, G.; Larson, D.; Smith, K. M.; Dunkley, J.; Gold, B.; Halpern, M.; Jarosik, N.; Kogut, A.; Komatsu, E.; Limon, M.; Meyer, S. S.; Nolta, M. R.; Odegard, N.; Page, L.; Spergel, D. N.; Tucker, G. S.; Weiland, J. L.; Wollack, E.; Wright, E. L. "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Are There Cosmic Microwave Background Anomalies?". January 3<sup>rd</sup>, 2011. Link ([^](#))
- [BJ] Hoftuft, J.; Eriksen, H. K.; Banday, A. J.; Gorski, K. M.; Hansen, F. K.; Lilje, P. B.. "Increasing Evidence for Hemispherical Power Asymmetry in the Five-Year WMAP Data". April 20<sup>th</sup>, 2009. Link ([^](#))
- [BK] NASA/Goddard/A. Kashlinsky, et al. "Mysterious Cosmic 'Dark Flow' Tracked Deeper into Universe". Link ([^](#))
- [BL] Turnbull, Stephen J.; Hudson, Michael J.; Feldman, Hume A.; Hicken, Malcolm; Kirshner, Robert P.; Watkins, Richard. "Cosmic flows in the nearby universe from Type Ia Supernovae". November 7<sup>th</sup>, 2011. Link ([^](#))
- [BM] "The Nobel Prize in Physics 2011". Nobelprize.org. Accessed November 4<sup>th</sup>, 2012. Link ([^](#))
- [BN] Jarosik, N.; Bennett, C. L.; Dunkley, J.; Gold, B.; Greason, M. R.; Halpern, M.; Hill, R. S.; Hinshaw, G.; Kogut, A.; Komatsu, E.; Larson, D.; Limon, M.; Meyer, S. S.; Nolta, M. R.; Odegard, N.; Page, L.; Smith, K. M.; Spergel, D. N.; Tucker, G. S.; Weiland, J. L.; Wollack, E.; Wright, E. L.. "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors and Basic Results". January 26<sup>th</sup>, 2010. Link ([^](#))
- [BO] Hogg, David. "Distance Measures in Cosmology". December 16<sup>th</sup>, 2000. Link ([^](#))
- [BP] Mahdavi, Andisheh; Hoekstra, Henk; Babul, Arif; Balam, David; Capak, Peter. "A Dark Core in Abell 520". February 10<sup>th</sup>, 2007. Link ([^](#))
- [BQ] NASA's Chandra X-ray Observatory. "NASA Finds Direct Proof of Dark Matter". August 21<sup>st</sup>, 2006. Link ([^](#))
- [BR] Lee, Joonghun; Komatsu, Eiichiro. "Bullet Cluster: A Challenge to  $\Lambda$ CDM Cosmology". May 22<sup>nd</sup>, 2010. Link ([^](#))
- [BS] Narayan, Ramesh; Bartelmann, Matthias. "Lectures on Gravitational Lensing". 1995. Link ([^](#))
- [BT] He, Ping; Zhang, Yuan-Zhong. "Modelling the Number Counts of Early-Type Galaxies by Pure Luminosity Evolution". February 17<sup>th</sup>, 1998. Link ([^](#))
- [BU] Colless, Matthew; Ellis, Richard; Taylor, Keith; Hook, Richard. "The LDSS Deep Redshift Survey". November 14<sup>th</sup>, 1989. Link ([^](#))
- [BV] Driver, Simon; Couch, Warrick. "The Inferred Redshift Distribution of the Faint Blue Galaxy Excess". July 20<sup>th</sup>, 1996. Link ([^](#))
- [BW] Roche, N.; Ratnatunga, K.; Griffiths, R. E.; Im, M. "Angular Sizes of the Faint Blue Galaxies". July 15<sup>th</sup>, 1996. Link ([^](#))
- [BX] Colless, M.; Schade, D.; Broadhurst, T. J.; Ellis, R. S. "High-Resolution Imaging of Faint Blue Galaxies". September 13<sup>th</sup>, 1993. Link ([^](#))
- [BY] Broadhurst, T. J.; Ellis, R. S.; Shanks, T. "The Durham/Anglo-Australian Telescope Faint Galaxy Redshift Survey". July 14<sup>th</sup>, 1988. Link ([^](#))
- [BZ] Glazebrook, K.; Ellis, R.; Colless, M.; Broadhurst, T.; Allington-Smith, J.; Tanvir, N. "A Faint Galaxy Redshift Survey to  $B = 24$ ". March 8<sup>th</sup>, 1995. Link ([^](#))
- [CA] Loveday, Jon. "The Local Space Density of Dwarf Galaxies". May 19<sup>th</sup>, 1997. Link ([^](#))
- [CB] "Correlation Between Absolute Magnitude and Diameter". Link ([^](#)). HOLMBERG, E., Ark. Astr. 3, 387 = Uppsala astr. Obs. Medd., No. 148 (1964).
- [CC] Weijmans, Anne-Marie; Bower, Richard G.; Geach, James E.; Swinbank, A. Mark; Wilman, R. J.; Zeeuw, P. T. de; Morris, Simon L. "Dissecting the Lyman- $\alpha$  emission halo of LAB1". November 18<sup>th</sup>, 2009. Link ([^](#))
- [CD] Ouchi, Masami; Ono, Yoshiaki; Egami, Eiichi; Saito, Tomoki; Oguri, Masamune; McCarthy, Patrick J.; Farrah, Duncan; Kashikawa, Nobunari; Momcheva, Ivelina; Shimasaku, Kazuhiro; Nakanishi, Kouichiro; Furusawa, Hisanori; Akiyama, Masayuki; Dunlop, James S.; Mortier, Angela M. J.; Okamura, Sadanori; Hayashi, Masao; Cirasuolo, Michele; Dressler, Alan; Iye, Masanori; Jarvis, Matt. J.; Kodama, Tadayuki; Martin, Crystal L.; McLure, Ross J.; Ohta, Kouji; Yamada, Toru; Yoshida, Michitoshi. "Discovery of a Giant Ly- $\alpha$  Emitter Near the Reionization Epoch". February 21<sup>st</sup>, 2009. Link ([^](#))

[CE] Mehrtens, Nicola; Romer, A. Kathy; Hilton, Matt; Lloyd-Davies, E. J.; Miller, Christopher J.; Stanford, S. A.; Hosmer, Mark; Hoyle, Ben; Collins, Chris A.; Liddle, Andrew R.; Viana, Pedro T. P.; Nichol, Robert C.; Stott, John P.; Dubois, E. Naomi; Kay, Scott T.; Sahlén, Martin; Young, Owain; Short, C. J.; Christodoulou, L.; Watson, William A.; Davidson, Michael; Harrison, Craig D.; Baruah, Leon; Smith, Mathew; Burke, Claire; Mayers, Julian A.; Deadman, Paul-James; Rooney, Philip J.; Edmondson, Edward M.; West, Michael; Campbell, Heather C.; Edge, Alastair C.; Mann, Robert G.; Sabirli, Kivanc; Wake, David; Benoist, Christophe; da Costa, Luiz; Maia, Marcio A. G.; Ogando, Ricardo. "The XMM Cluster Survey: optical analysis methodology and the first data release". June 2012.

[CF] The scientific results reported in this article are based in part on observations made by the Chandra X-ray Observatory

[CG] Hopkins, Philip F.; Croton, Darren; Bundy, Kevin; Khochfar, Sadegh; Bosch, Frank van den; Somerville, Rachel S.; Wetzel, Andrew; Keres, Dusan; Hernquist, Lars; Stewart, Kyle; Younger, Joshua D.; Genel, Shy; Ma, Chung-Pei. "Mergers in  $\Lambda$ CDM: Uncertainties in Theoretical Predictions and Interpretations of the Merger Rate". June 2010. Link ([^](#))

[CH] Conselice, Christopher J.; Yang, Cui; Bluck, Asa F. L. "The Structures of Distant Galaxies - III: The Merger History of over 20,000 Massive Galaxies at  $z < 1.2$ ". December 17<sup>th</sup>, 2008. Link ([^](#))

[CI] Lopez-Sanjuan, Carlos; Balcells, Marc; Perez-Gonzalez, Pablo G.; Barro, Guillermo; Garcia-Dabo, Cesar Enrique; Gallego, Jesus; Zamorano, Jaime. "The Galaxy Major Merger Fraction to  $z \sim 1$ ". February 23<sup>rd</sup>, 2009. Link ([^](#))

[CJ] Lotz, Jennifer; Jonsson, Patrik; T.J. Cox; Croton, Darren; Primack, Joel; Somerville, Rachel; Stewart, Kyle. "Astronomers Pin Down Galaxy Collision Rate". October 27<sup>th</sup>, 2011. Link ([^](#))

[CK] Toft, S.; Dokkum, P. van; Franx, M.; Thompson, R. I.; Illingworth, G. D.; Bouwens, R. J.; Kriek, M. "Distant Red Galaxies in the Hubble Ultra Deep Field". November 16<sup>th</sup>, 2004. Link ([^](#))

[CL] Conselice, C. J.; Bluck, A.F.L.; Ravindranath, S.; Mortlock, A.; Koekemoer, A.; Buitrago, F.; Grützbauch, R.; Penny, S. "The Tumultuous Formation of the Hubble Sequence at  $z > 1$  Examined with HST/WFC3 Observations of the Hubble Ultra Deep Field". May 12<sup>th</sup>, 2011. Link ([^](#))

[CM] Conselice, Christopher J.; Bershadsky, Matthew A.; Dickinson, Mark; Papovich, Casey. "A Direct Measurement of Major Galaxy Mergers at  $z \leq 3$ ". June 5<sup>th</sup>, 2003. Link ([^](#))

[CN] XENON100 Collaboration: Aprile, E.; Arisaka, K.; Arneodo, F.; Askin, A.; Baudis, L.; Behrens, A.; Bokeloh, K.; Brown, E.; Bruch, T.; Bruno, G.; Cardoso, J. M. R.; Chen, W.T.; Choi, B.; Cline, D.; Duchovni, E.; Fattori, S.; Ferella, A. D.; Gao, F.; Giboni, K.L.; Gross, E.; Kish, A.; Lam, C. W.; Lamblin, J.; Lang, R. F.; Levy, C.; Lim, K. E.; Lin, Q.; Lindemann, S.; Lindner, M.; Lopes, J. A. M.; Lung, K.; Undagoitia, T.; Marrodan; Mei, Y.; Fernandez, A. J. Melgarejo; Ni, K.; Oberlack, U.; Orrigo, S. E. A.; Pantic, E.; Persiani, R.; Plante, G.; Ribeiro, A. C. C.; Santorelli, R.; Santos, J. M. F. dos; Sartorelli, G.; Schumann, M.; Selvi, M.; Shagin, P.; Simgen, H.; Teymourian, A.; Thers, D.; Vitells, O.; Wang, H.; Weber, M.; Weinheimer, C. "Dark Matter Results from 100 Live Days of XENON100 Data". September 7<sup>th</sup>, 2011. Link ([^](#))

[CO] XENON100 Collaboration. "Dark Matter Results from 225 Live Days of XENON100 Data". July 25<sup>th</sup>, 2012. Link ([^](#))

[CP] Lotz, Jennifer M.; Jonsson, Patrik; Cox, T.J.; Primack, Joel R. "Galaxy Merger Morphologies and Timescales from Simulations of Equal-Mass Gas-Rich Disk Mergers". May 8<sup>th</sup>, 2008. Link ([^](#))

[CQ] Lotz, Jennifer; Jonsson, Patrik; T.J. Cox; Croton, Darren; Primack, Joel; Somerville, Rachel; Stewart, Kyle. "The Major and Minor Galaxy Merger Rates at  $z < 1.5$ ". August 10<sup>th</sup>, 2011. Link ([^](#))

[CR] Hsieh, B. C.; Yee, H. K. C.; Lin, H.; Gladders, M. D.; Gilbank, D. G. "Pair Analysis of Field Galaxies from the Red-Sequence Cluster Survey". April 10<sup>th</sup>, 2008. Link ([^](#))

[CS] Xu, C. Kevin. "NIR/Optical Selected Local Mergers - Spatial Density and sSFR Enhancement". May 4<sup>th</sup>, 2012. Link ([^](#))

[CT] Edwards, Louise O.V.; Fadda, Dario. "A Multi-Wavelength Analysis of Spitzer Selected Coma Cluster Galaxies: Star Formation Rates and Masses". September 14<sup>th</sup>, 2011. Link ([^](#))

[CU] Churchill, Chris. "Mg II Absorbers: An On-line Review 'Paper'". December 1999. Link ([^](#))

[CV] Evans, Jessica L.; Churchill, Christopher W.; Murphy, Michael T. "The Redshift Distribution of Intervening Weak MgII Quasar Absorbers and a Curious Dependence on Quasar Luminosity". Draft: July 3<sup>rd</sup>, 2012. Link ([^](#))

[CW] Dietrich, M.; Hamann, F.; Appenzeller, I.; Vestergaard, M. "FeII/MgII Emission-Line Ratio in High-Redshift Quasars". June 27<sup>th</sup>, 2003. Link ([^](#))

[CX] Rosa, Gisella De; Decarli, Roberto; Walter, Fabian; Fan, Xiaohui; Jiang, Linhua; Kurk, Jaron; Pasquali, Anna; Rix, Hans-Walter. "Evidence for Non-Evolving FeII/MgII Ratios in Rapidly Accreting Z~6 QSOs". June 23<sup>rd</sup>, 2007. Link ([^](#))

[CY] Jiang, Linhua; Fan, Xiaohui; Vestergaard, Marianne; Kurk, Jaron D.; Walter, Fabian; Kelly, Brandon C.; Strauss, Michael A. "Gemini Near-Infrared Spectroscopy of Luminous z~6 Quasars: Chemical Abundances, Black Hole Masses and MgII Absorption". July 11<sup>th</sup>, 2007. Link ([^](#))

[CZ] Maiolino, R.; Juarez, Y.; Mujica, R.; Nagar, N.; Oliva, E. "Early Star Formation Traced by the Highest Redshift Quasars". September 8<sup>th</sup>, 2003. Link ([^](#))

[DA] Wei, Hao. "Observational Constraints on Cosmological Models with the Updated Long Gamma-Ray Bursts". August 16<sup>th</sup>, 2010. Link ([^](#))

[DB] Cardone, V.F.; Capozziello, S.; Dainotti, M.G. "An Updated Gamma Ray Bursts Hubble Diagram". July 27<sup>th</sup>, 2009. Link ([^](#))

[DC] Wei, Hao; Zhang, Shuang Nan. "Reconstructing the Cosmic Expansion History up to Redshift  $z = 6.29$  with the Calibrated Gamma-Ray Bursts". August 31<sup>st</sup>, 2009. Link ([^](#))

[DD] Collazzi, Andrew C.; Schaefer, Bradley E.; Goldstein, Adam; Preece, Robert D. "A Significant Problem with Using the Amati Relation for Cosmological Purposes". December 19<sup>th</sup>, 2011. Link ([^](#))

# Chemical abundances of 1111 FGK stars from the HARPS GTO planet search program

## Galactic stellar populations and planets<sup>Y,YY,YYY</sup>

V. Zh. Adibekyan<sup>1</sup>, S. G. Sousa<sup>1,2</sup>, N. C. Santos<sup>1,3</sup>, E. Delgado Mena<sup>1</sup>, J. I. González Hernández<sup>2,4</sup>, G. Israelian<sup>2,4</sup>, M. Mayor<sup>5</sup>, and G. Khachaturyan<sup>1,3</sup>

<sup>1</sup> Centro de Astrofísica da Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal  
e-mail: Vardan.Adibekyan@astro.up.pt

<sup>2</sup> Instituto de Astrofísica de Canarias, 38200 La Laguna, Tenerife, Spain

<sup>3</sup> Departamento de Física e Astronomia, Faculdade de Ciências da Universidade do Porto, Portugal

<sup>4</sup> Departamento de Astrofísica, Universidad de La Laguna, 38206 La Laguna, Tenerife, Spain

<sup>5</sup> Observatoire de Genève, Université de Genève, 51 Ch. des Maillettes, 1290 Sauverny, Switzerland

Received 12 April 2012 / Accepted 10 July 2012

### ABSTRACT

**Context.** We performed a uniform and detailed abundance analysis of 12 refractory elements (Na, Mg, Al, Si, Ca, Ti, Cr, Ni, Co, Sc, Mn, and V) for a sample of 1111 FGK dwarf stars from the HARPS GTO planet search program. Of these stars, 109 are known to harbor giant planetary companions and 26 stars are exclusively hosting Neptunians and super-Earths.

**Aims.** The two main goals of this paper are to investigate whether there are any differences between the elemental abundance trends for stars of different stellar populations and to characterize the planet host and non-host samples in terms of their [X/H]. The extensive study of this sample, focused on the abundance differences between stars with and without planets will be presented in a parallel paper.

**Methods.** The equivalent widths of spectral lines were automatically measured from HARPS spectra with the ARES code. The abundances of the chemical elements were determined using an LTE abundance analysis relative to the Sun, with the 2010 revised version of the spectral synthesis code MOOG and a grid of Kurucz ATLAS9 atmospheres. To separate the Galactic stellar populations we applied both a purely kinematical approach and a chemical method.

**Results.** We found that the chemically separated (based on the Mg, Si, and Ti abundances) thin- and thick disks are also chemically disjunct for Al, Sc, Co, and Ca. Some bifurcation might also exist for Na, V, Ni, and Mn, but there is no clear boundary of their [X/Fe] ratios. We confirm that an overabundance in giant-planet host stars is clear for all studied elements. We also confirm that stars hosting only Neptunian-like planets may be easier to detect around stars with similar metallicities than around non-planet hosts, although for some elements (particular a-elements) the lower limit of [X/H] is very abrupt.

**Key words.** stars: abundances – planetary systems – stars: fundamental parameters – Galaxy: disk – solar neighborhood – stars: kinematics and dynamics

## 1. Introduction

High-precision radial velocity measurements resulted in the detection of the first extra-solar planetary system surrounding a main-sequence star similar to our own in 1995 (Mayor & Queloz 1995). Observational progress in extra-solar planet detection and characterization is now moving rapidly on several fronts. More than 750 planetary companions have already been found orbiting late-type stars<sup>1</sup>. The total number of planet-harboring systems that are found using Doppler technique is approaching 500.

<sup>x</sup> Based on observations collected at the La Silla Paranal Observatory, ESO (Chile) with the HARPS spectrograph at the 3.6-m telescope (ESO runs ID 72.C-0488, 082.C-0212, and 085.C-0063).

<sup>xx</sup> Full Tables 4, 5, and the table with EWs of the lines are only available at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via

<http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/545/A32>

<sup>xxx</sup> Figure A.1 is available in electronic form at

<http://www.aanda.org>

<sup>1</sup> <http://exoplanet.eu/>

A strong input for this number was made by several dedicated planet-search programs that systematically monitor the sky. Among these programs, the HARPS planet search program made a special contribution. The high spectral resolution and most importantly the long-term stability of the HARPS spectrograph (Mayor et al. 2003) allowed discovering a fairly large number of new planets, including the large majority of the known planets with masses near the mass of Neptune or below (e.g. Santos et al. 2004b; Lovis et al. 2006; Mayor et al. 2009, 2011).

Shortly after the discovery of the first extra-solar planet, Gonzalez (1998), based on a small sample of eight planet-host stars (PHS), suggested that PHSs tend to be metal-rich compared with the nearby field FGK stars that are known to host no-planet. The metal-rich nature of the PHSs have been confirmed in subsequent papers (e.g. Gonzalez et al. 2001; Santos et al. 2001, 2003, 2004a, 2005; Laws et al. 2003; Fischer & Valenti 2005; Gilli et al. 2006; Udry et al. 2006; Ecuivillon et al. 2007; Sousa et al. 2008; Neves et al. 2009; Johnson et al. 2010; Kang et al. 2011). This tendency for giant planets that orbit

metal-rich stars strongly supports the core-accretion model of planet formation (e.g. Pollack et al. 1996). This implies that core accretion (Ida & Lin 2004; Mordasini et al. 2009) and not disk-instability (Boss 1997) is the main working mechanism for the formation of giant planets. Interestingly, recent studies show that Neptune and super-Earth-class planets may easier form in a low-metal-content environment (e.g. Udry et al. 2006; Sousa et al. 2008, 2011a; Ghezzi et al. 2010; Mayor et al. 2011; Buchhave et al. 2012).

Most spectroscopic studies are in general limited to small samples of a few hundred comparison stars and less than one hundred PHSs at most, and only a few studies have been based on samples as large as 1000 stars (e.g. Gazzano et al. 2010; Gazzano 2011; Petigura & Marcy 2011). In order to minimize the errors, one needs to have large and homogeneous samples with reliable measurements of their chemical features.

In this paper, we present a uniform spectroscopic analysis of 1111 FGK dwarfs observed within the context of the HARPS GTO planet search program. The paper is organized as follows: in Sect. 2, we introduce the sample used in this work. The method of the chemical abundance determination and analysis will be explained in Sect. 3. This section also includes discussion of the uncertainties and errors in our methodology as well as a comparison of our results with the literature. The calculation of the galactic space velocity data and the selection of different populations of stars, based on their kinematic and chemical properties, are presented in Sect. 4. A discussion of the [X/H] abundances of the exoplanet hosts can be found in Sect. 5. The main conclusions of the paper are finally addressed in Sect. 6. The extensive and full investigation of this sample, focused on the abundance difference between stars with and without planets will be presented in a parallel paper (Adibekyan et al. 2012).

## 2. Sample description and stellar parameters

The sample used in this work consists of 1111 FGK stars observed in the context of the HARPS GTO programs. It is a combination of three HARPS subsamples hereafter called HARPS-1 (Mayor et al. 2003), HARPS-2 (Lo Curto et al. 2010), and HARPS-4 (Santos et al. 2011). Note that the HARPS-2 planet search program is the complement of the previously started CORALIE survey (Udry et al. 2000) to fainter magnitudes and to a larger volume. The stars were selected to be suitable for radial velocity surveys. They are slowly rotating and non-evolved solar-type dwarfs with spectral type between F2 and M0 that do not show a high level of chromospheric activity either.

The individual spectra of each star were reduced using the HARPS pipeline and then combined with IRAF<sup>2</sup> after correcting for its radial velocity. The final spectra have a resolution of  $R \sim 110\,000$  and signal-to-noise ratio (S/N) ranging from  $\sim 20$  to  $\sim 2000$ , depending on the amount and quality of the original spectra. Fifty-five percent of the spectra have an S/N higher than 200, about 16% of stars have an S/N lower than 100, and less than 1% of the stars have an S/N lower than 40.

Precise stellar parameters for the entire sample were determined by Sousa et al. (2008, 2011a,b) using the same spectra as we did for this study. We refer the reader to these papers for details. The authors used a set of FeI and FeII lines whose

equivalent widths (EW) were measured using the ARES<sup>3</sup> code (automatic routine for line equivalent widths in stellar spectra – Sousa et al. 2007)<sup>4</sup>. Assuming ionization and excitation equilibrium, the parameters were derived through an iterative process until the slope of the relation between the abundances given by individual FeI lines and both the excitation potential ( $\chi_1$ ) and reduced equivalent width ( $\log EW/\lambda$ ) were zero, and until the FeI and FeII lines provided the same average abundance. The spectroscopic analysis was completed assuming local thermodynamic equilibrium (LTE) with a grid of Kurucz atmosphere models (Kurucz et al. 1993), and making use of a recent version of the MOOG<sup>5</sup> radiative transfer code (Sneden 1973). The typical precision uncertainties for the atmospheric parameters are of about 30 K for  $T_{\text{eff}}$ , 0.06 dex for  $\log g$ , 0.08 km s<sup>-1</sup> for  $\xi_t$ , and 0.03 dex for [Fe/H]. We note that there are four stars in common between HARPS-1 and HARPS-4, and 14 stars between the HARPS-2 and HARPS-4 subsamples.

The stars in the sample have derived effective temperatures from 4487 K to 7212 K, but very few stars have temperatures that are very different from those of the Sun (there are e.g. only 12 stars with  $T_{\text{eff}} > 6500$  K). The metallicities of the stars range from  $-1.39$  to  $0.55$  dex and have surface gravities from 2.68 to 4.96 dex (again there are very few “outliers”, only five stars with  $\log g < 3.8$  dex).

As already noted before, HARPS has contributed very much to the present high number of known planetary systems. Recently, Mayor et al. (2011) reported on the results of an eight-year HARPS survey with a statistical analysis of the planet and host samples. Simultaneously, they presented the list of newly discovered planets. We included these data when we updated the original GTO (Guaranteed Time Observations) catalog using data from the extra-solar planets encyclopedia<sup>6</sup>. The total number of PHSs in the current sample is now 135, of which 26 are super-Earths and Neptune-mass (the mass of the heaviest planet is less than  $30 M_{\oplus}$ ) planet hosts (hereafter NH).

## 3. Abundance analysis and uncertainties

Elemental abundances for 12 elements (Na, Mg, Al, Si, Ca, Ti, Cr, Ni, Co, Sc, Mn, and V) were determined using an LTE analysis with the Sun as reference point with the 2010 revised version of the spectral synthesis code MOOG (Sneden 1973) and a grid of Kurucz ATLAS9 plane-parallel model atmospheres (Kurucz et al. 1993). The reference abundances used in the analysis were taken from Anders & Grevesse (1989). The line list and atomic parameters of Neves et al. (2009) were used, adding the CaI line at  $\lambda 5260.39$  (excitation energy of the lower energy level  $\chi_1 = 2.52$ , and oscillator strength  $\log g f = -1.836$ ) and excluding five NiI lines ( $\lambda 4811.99, \lambda 4946.04, \lambda 4995.66, \lambda 5392.33$ , and  $\lambda 5638.75$ ), two SiI lines ( $\lambda 5517.54$  and  $\lambda 5797.87$ ), two TiII lines ( $\lambda 4657.20$  and  $\lambda 4708.67$ ) and five TiI lines ( $\lambda 4656.47, \lambda 5064.06, \lambda 5113.44, \lambda 5219.70$ , and  $\lambda 5490.16$ ). These lines were excluded because the [X/Fe] abundance ratios determined by them showed significant trends with effective temperature (see also Neves et al. 2009, for details of the lines selection). The EWs were automatically measured with the ARES code. The

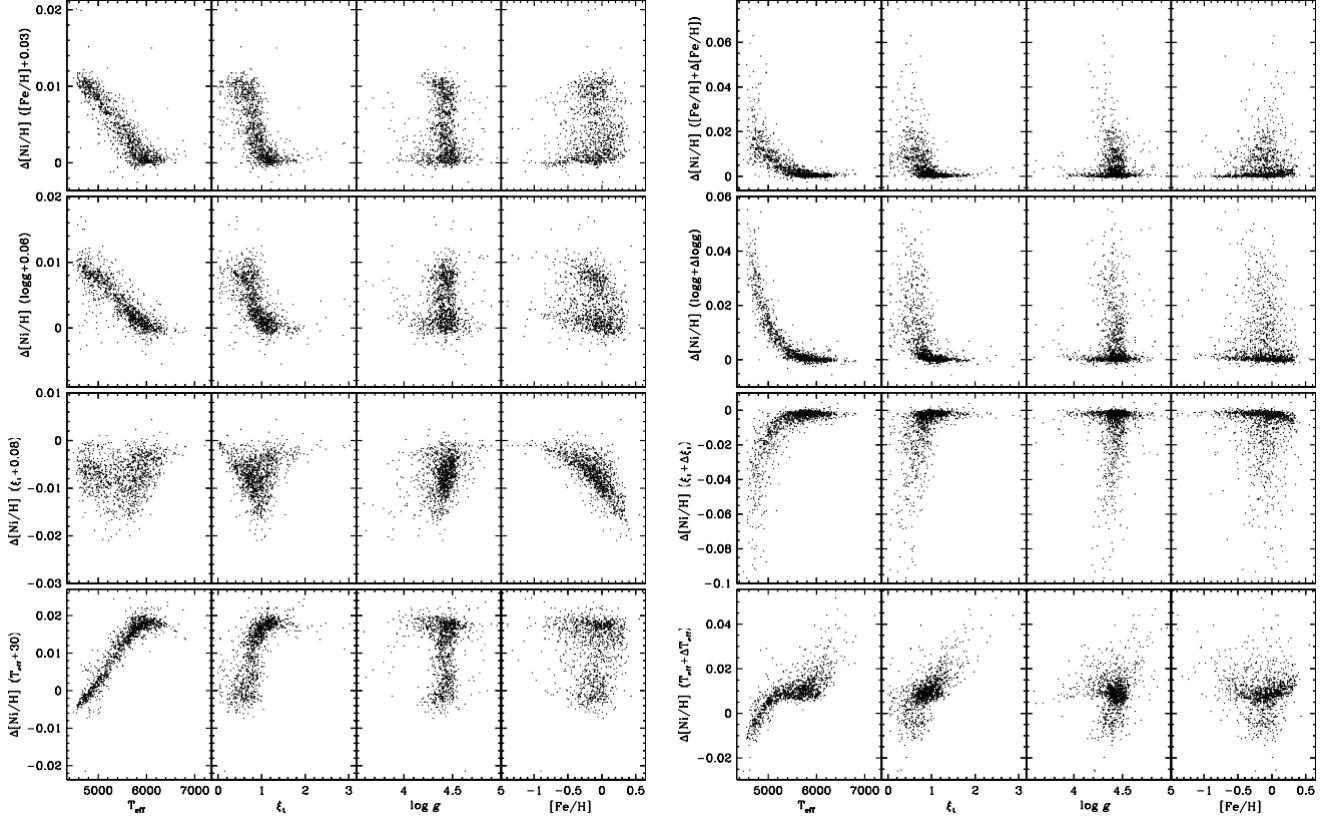
<sup>3</sup> The ARES code can be downloaded at <http://www.astro.up.pt/sousasag/ares>

<sup>4</sup> The EWs of the lines for the entire sample is available at the CDS.

<sup>5</sup> The source code of MOOG2010 can be downloaded at <http://www.as.utexas.edu/~chris/moog.html>

<sup>6</sup> <http://exoplanet.eu/>

<sup>2</sup> IRAF is distributed by National Optical Astronomy Observatories, operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation, USA.



**Fig. 1.** Ni abundance sensitivity to the stellar parameter variations as a function of model atmosphere parameters. *Left* – The variation of the atmospheric parameters are the same for all stars and are equal to the typical errors. *Right* – The variation of the atmospheric parameters are equal to their one-sigma errors taken for each star individually.

input parameters for ARES were calculated following the procedure discussed in Sousa et al. (2011b).

The final abundance for each star and element was calculated to be the average value of the abundances given by all lines detected in a given star and element. Individual lines for a given star and element with a line dispersion more than a factor of two higher than the rms were excluded. In this way we avoided the errors caused by bad pixels, cosmic rays, or other unknown effects.

### 3.1. Uncertainties

Since the abundances were determined via the measurement of EWs and using already determined stellar parameters, the errors might still come from the EW measurements, from the errors in the atomic parameters, and from the uncertainties of the atmospheric parameters that were used to make an atmosphere model. In addition to the above-mentioned errors, one should add systematic errors that can occur due to NLTE or granulation (3D) effects. To minimize the errors, it is very important to use high-quality data and as many lines as possible for each element.

It is hard to define the contribution of each error source on the abundance results separately, but we can examine the sensitivity of the abundances to the stellar parameters and test the reliability of our results by comparing the abundances with those obtained in the literature.

First, to study the sensitivity trends of the abundances to the variation of the stellar parameters in general, we performed numerical tests with variations in the model parameters by a constant value similar to their typical errors:  $\Delta T_{\text{eff}} = \pm 30$  K,

**Table 1.** Average of the stellar parameters for the subsamples with different  $T_{\text{eff}}$ .

	$T_{\text{eff}}$ (K)	$\log g$ (dex)	$[\text{Fe}/\text{H}]$ (dex)	$\xi_t$ (km s\$^{-1}\$)
low $T_{\text{eff}}$	4934	4.4	-0.17	0.59
Solar	5769	4.39	-0.12	1.01
high $T_{\text{eff}}$	6440	4.51	-0.05	1.7

$\Delta \log g = \pm 0.06$  dex,  $\Delta \xi_t = \pm 0.08$  km s\$^{-1}\$, and  $\Delta [\text{Fe}/\text{H}] = \pm 0.03$  dex. Then we calculated the abundance differences between the values obtained with and without varying the parameter. The maxima from the plus and minus cases were taken. A thorough investigation of this experiment shows that the picture is quite complicated. Changing one of the parameters will increase or decrease the abundance of a certain element depending on the stellar parameters. For example, in Fig. 1 (left panel) one can see that a variation of  $T_{\text{eff}}$  by +30 K may change the Ni abundance from about +0.02 to -0.01 dex, depending on the  $T_{\text{eff}}$  of the stars. Despite the complex picture, it can be observed that in general, the sensitivity of all element abundances to the stellar parameters also depends on the effective temperature. Following this correlation, we grouped our sample stars into three temperature groups: “low  $T_{\text{eff}}$ ” stars – stars with  $T_{\text{eff}} < 5277$  K, “solar” – stars with  $T_{\text{eff}} = T_s \pm 500$  K, and “high  $T_{\text{eff}}$ ” – stars with  $T_{\text{eff}} > 6277$  K. The average of the stellar parameters for the aforementioned groups are presented in Table 1.

The results obtained from the test for three groups of stars are displayed in Table 2. Table 2 shows that neutral species are generally more sensitive to changes in effective temperature. For

**Table 2.** Abundance sensitivities of the studied elements to changes of  $\pm 100$  K in  $T_{\text{eff}}$ ,  $\pm 0.2$  dex in  $\log g$  and [Fe/H],  $\pm 0.5$  km s $^{-1}$  in  $\xi_t$ .

	FeI	NaI	MgI	All	SiI	CaI	ScI	ScII
$\Delta T_{\text{eff}} = \pm 30$ K								
low $T_{\text{eff}}$	$\pm 0.00$	$\pm 0.02$	$\pm 0.01$	$\pm 0.02$	$\mp 0.01$	$\pm 0.03$	$\pm 0.04$	$\mp 0.00$
solar	$\pm 0.01$	$\pm 0.01$	$\pm 0.01$	$\pm 0.01$	$\pm 0.00$	$\pm 0.02$	$\pm 0.02$	$\pm 0.00$
high $T_{\text{eff}}$	$\pm 0.01$	$\pm 0.02$	$\pm 0.02$	$\pm 0.00$				
$\Delta [\text{Fe}/\text{H}] = \pm 0.03$ dex								
low $T_{\text{eff}}$	–	$\mp 0.00$	$\pm 0.00$	$\pm 0.00$	$\pm 0.01$	$\pm 0.00$	$\pm 0.00$	$\pm 0.01$
solar	–	$\pm 0.00$	$\pm 0.01$					
high $T_{\text{eff}}$	–	$\pm 0.00$						
$\Delta \log g = \pm 0.06$ dex								
low $T_{\text{eff}}$	$\mp 0.01$	$\mp 0.01$	$\mp 0.01$	$\mp 0.01$	$\pm 0.01$	$\mp 0.02$	$\mp 0.01$	$\pm 0.02$
solar	$\mp 0.01$	$\mp 0.00$	$\mp 0.01$	$\mp 0.00$	$\pm 0.00$	$\mp 0.01$	$\mp 0.00$	$\pm 0.02$
high $T_{\text{eff}}$	$\mp 0.00$	$\pm 0.02$						
$\Delta \xi_t = \pm 0.08$ km s $^{-1}$								
low $T_{\text{eff}}$	$\mp 0.01$	$\mp 0.00$	$\mp 0.00$	$\mp 0.00$	$\mp 0.00$	$\mp 0.01$	$\mp 0.01$	$\mp 0.01$
solar	$\mp 0.01$	$\mp 0.00$	$\mp 0.00$	$\mp 0.00$	$\mp 0.00$	$\mp 0.01$	$\mp 0.00$	$\mp 0.01$
high $T_{\text{eff}}$	$\mp 0.00$	$\mp 0.01$	$\mp 0.00$	$\mp 0.01$				
TiI								
$\Delta T_{\text{eff}} = \pm 30$ K								
low $T_{\text{eff}}$	$\pm 0.04$	$\mp 0.00$	$\pm 0.04$	$\pm 0.03$	$\mp 0.02$	$\pm 0.02$	$\pm 0.00$	$\pm 0.00$
solar	$\pm 0.03$	$\pm 0.00$	$\pm 0.03$	$\pm 0.02$	$\mp 0.00$	$\pm 0.02$	$\pm 0.02$	$\pm 0.01$
high $T_{\text{eff}}$	$\pm 0.02$	$\pm 0.00$	$\pm 0.02$	$\pm 0.02$	$\mp 0.00$	$\pm 0.02$	$\pm 0.02$	$\pm 0.02$
$\Delta [\text{Fe}/\text{H}] = \pm 0.03$ dex								
low $T_{\text{eff}}$	$\pm 0.00$	$\pm 0.01$	$\pm 0.00$	$\pm 0.00$	$\pm 0.01$	$\pm 0.01$	$\pm 0.01$	$\pm 0.01$
solar	$\pm 0.00$	$\pm 0.01$	$\pm 0.00$					
high $T_{\text{eff}}$	$\pm 0.00$							
$\Delta \log g = \pm 0.06$ dex								
low $T_{\text{eff}}$	$\mp 0.01$	$\pm 0.02$	$\mp 0.01$	$\mp 0.01$	$\pm 0.02$	$\mp 0.02$	$\pm 0.01$	$\pm 0.01$
solar	$\mp 0.00$	$\pm 0.02$	$\mp 0.00$	$\mp 0.00$	$\pm 0.02$	$\mp 0.01$	$\pm 0.00$	$\pm 0.00$
high $T_{\text{eff}}$	$\mp 0.00$	$\pm 0.02$	$\mp 0.00$	$\mp 0.00$	$\pm 0.02$	$\mp 0.00$	$\pm 0.00$	$\pm 0.00$
$\Delta \xi_t = \pm 0.08$ km s $^{-1}$								
low $T_{\text{eff}}$	$\mp 0.02$	$\mp 0.01$	$\mp 0.02$	$\mp 0.01$				
solar	$\mp 0.01$	$\mp 0.01$	$\mp 0.00$	$\mp 0.01$	$\mp 0.02$	$\mp 0.02$	$\mp 0.00$	$\mp 0.01$
high $T_{\text{eff}}$	$\mp 0.00$	$\mp 0.01$	$\mp 0.00$	$\mp 0.00$	$\mp 0.02$	$\mp 0.01$	$\mp 0.00$	$\mp 0.00$

gravity variations, the neutral species were hardly affected, and the variations become noticeable only for stars with low  $T_{\text{eff}}$ , but the ionized species constantly varied by the same amount independently of the effective temperature. The ions are also more sensitive to metallicity changes than the neutral elements, although the sensitivity is not as significant as that for either  $T_{\text{eff}}$  and  $\log g$ . Finally, microturbulence variations led to only very small changes in most abundances (because many species are represented only by weak lines) and only few species are an exception.

Table 2 gives an overview of the elemental abundances variation with the variation of the stellar parameters, but not the uncertainties induced by the errors in the stellar parameters for our sample. The spectroscopic stellar parameters and metallicities were derived based on the equivalent widths of the FeI and FeII weak lines by imposing excitation and ionization equilibrium assuming LTE (e.g. Sousa et al. 2011b, and references therein). The errors obtained for the stars are typically very small, especially for stars similar to the Sun. This comes directly from the method itself because a differential analysis is performed with the Sun as reference. Stars that are significantly cooler or hotter than the Sun have larger intrinsic errors. To estimate the scale errors induced by uncertainties in the model atmosphere parameters, we varied the model parameters by an amount of their one-sigma errors available for each star and then we again divided our sample stars into three temperature groups as presented above. The average errors in the  $T_{\text{eff}}$  are 70, 24, and 45 K for cool, Sun-like, and hot star groups, respectively. The average

errors in  $\log g$  are 0.15, 0.03, and 0.05, in  $\xi_t$  – 0.3, 0.04, and 0.08, and in [Fe/H] –0.04, 0.02, and 0.03 for the three groups, respectively. The right panel of Fig. 1 shows an example of the abundance variations with the variation of the stellar parameters against model parameters for Ni. From the figure it becomes clear that for stars with atmospheric parameters close to those of the Sun the uncertainties induced by the errors in the stellar parameters are very small (except for  $\log g$ ). This is because both the abundance and stellar parameters are determined using an analysis with the Sun as reference point.

We evaluated the errors in the abundances of all elements [X/H], adding quadratically the line-to-line scatter errors and errors induced by uncertainties in the model atmosphere parameters. The line-to-line scatter errors were estimated as  $\sigma/\sqrt{N}$ , where  $\sigma$  is the standard deviation of  $N$  measurements (unfortunately, for some elements we were only able to select two or three lines). The average of  $\sigma/\sqrt{N}$  and [X/H] errors for the three groups are presented in Table 3. The table shows that the  $\sigma/\sqrt{N}$  errors constitute the main part of the  $\sigma[\text{X}/\text{H}]$  total errors for the stars with  $T_{\text{eff}} = T > 500$  K. The atmospheric parameters were obtained from the FeI and FeII lines by iterating until the correlation coefficients between  $\log S(\text{FeI})$  and  $\chi_1$ , and between  $\log S(\text{FeI})$  and  $\log(W/\lambda)$  were zero, and the mean abundance given by FeI and FeII lines were the same (e.g. Santos et al. 2004a; Sousa et al. 2008). This means that the parameters are interrelated, i.e., variation of one parameter will influence others. Hence, the total error could be slightly higher due to the

**Table 3.** The average error for the element abundances [X/H], and abundance ratios [X/Fe].

Elem	$\sqrt{f}_N$	Low $T_{\text{eff}}$		solar			High $T_{\text{eff}}$	
		$\sigma[X/\text{H}]$	$\sigma[X/\text{Fe}]$	$\sqrt{f}_N$	$\sigma[X/\text{H}]$	$\sigma[X/\text{Fe}]$	$\sqrt{f}_N$	$\sigma[X/\text{H}]$
NaI	0.05	0.09	0.08	0.02	0.02	0.02	0.06	0.07
MgI	0.07	0.08	0.07	0.03	0.04	0.03	0.05	0.06
AlI	0.03	0.07	0.06	0.02	0.03	0.02	0.08	0.09
SiI	0.02	0.05	0.06	0.01	0.01	0.01	0.02	0.03
CaI	0.03	0.10	0.09	0.01	0.02	0.01	0.02	0.04
ScI	0.11	0.16	0.13	0.03	0.04	0.03	0.04	0.14
ScII	0.04	0.08	0.08	0.02	0.03	0.03	0.04	0.05
TiI	0.02	0.12	0.10	0.01	0.02	0.01	0.02	0.04
TiII	0.05	0.09	0.09	0.02	0.03	0.03	0.03	0.04
VI	0.07	0.16	0.14	0.02	0.03	0.02	0.05	0.07
CrI	0.02	0.08	0.07	0.01	0.02	0.01	0.02	0.03
CrII	0.07	0.11	0.10	0.03	0.03	0.03	0.03	0.05
MnI	0.05	0.10	0.08	0.03	0.04	0.03	0.04	0.05
CoI	0.03	0.06	0.04	0.02	0.03	0.02	0.05	0.06
NiI	0.01	0.04	0.03	0.00	0.01	0.01	0.01	0.03

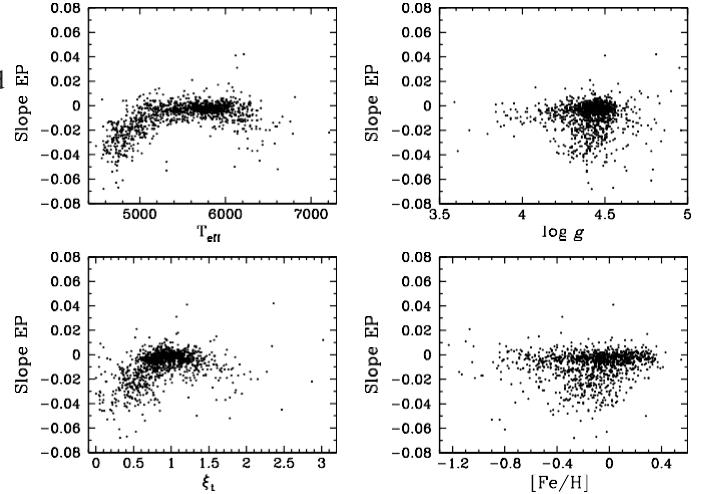
described covariance terms (e.g. Johnson et al 2002; Cayrel et al. 2004; Lai et al. 2008).

The errors in the abundance ratios, [X/Fe], were determined taking into account the differences between the sensitivities of the resulting abundance ratios to changes in the assumed atmospheric parameters and the dispersion of the abundances from individual lines of each X element. Table 2 shows that, in general, the model changes (variation of stellar parameters) induce similar effects in the abundances of different elements and Fe, so that they partially cancel out in the ratio [X/Fe]. The average error for the element abundances [X/H] and abundance ratios [X/Fe] are presented in Table 3.

### 3.2. Testing the validity of the stellar parameters

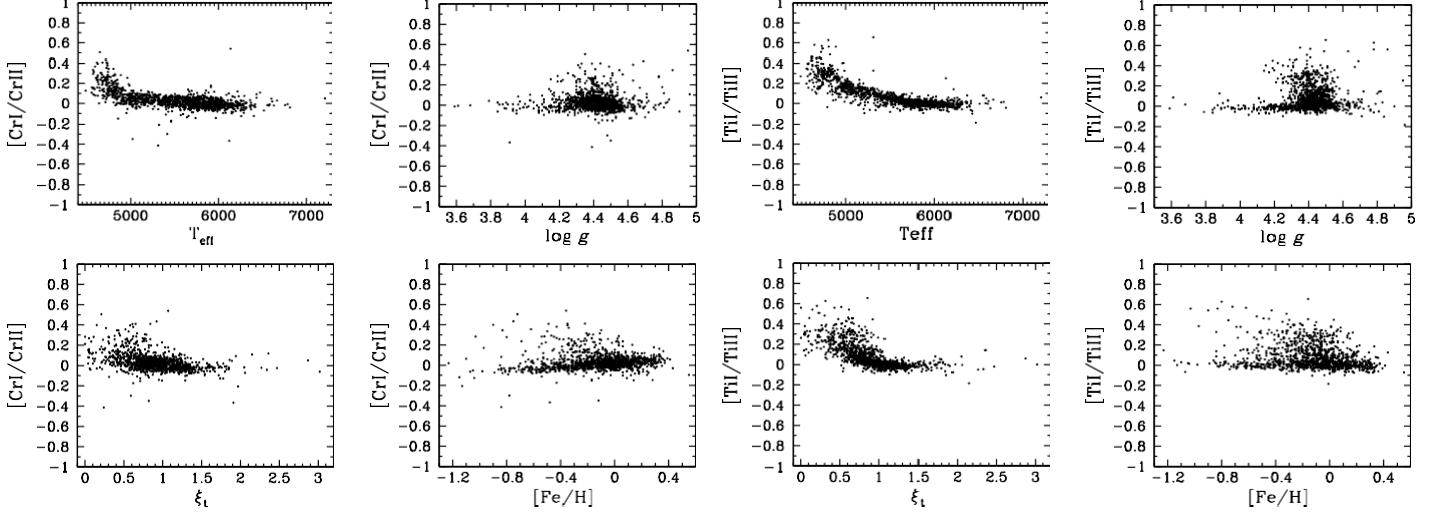
As stated before, the chemical abundances of the elements were derived by completing an LTE abundance analysis with the Sun as reference point using EW measurements. To check the validity limit of the adopted methodology in terms of stellar parameter ranges, and to test the stellar parameters themselves, we tested our results in a variety of ways. First we calculated the slopes of the derived abundances of the considered lines as a function of the excitation potential (EP) of the NiI lines. We chose nickel because its lines cover a wide range of EPs. In this way, we verified whether the excitation equilibrium enforced on the FeI lines of every star was applicable to other species. In Fig. 2, we plot the slopes of EP obtained for each star against the stellar parameters. The figure shows that there are no discernible trends of EP with  $\log g$  and [Fe/H], but there is a trend with  $T_{\text{eff}}$  and  $\xi_t$  (the  $\xi_t$  trend is just noticeable): cooler stars with  $T_{\text{eff}} \gtrsim 5000$  K, which also have low microturbulence velocities, have a systematic bias away from the expected values.

Then, in Fig. 3 we plot the [CrI/CrII] and [TiI/TiII] as a function of the stellar parameters to ensure that the ionization equilibrium enforced on the FeII lines (Sousa et al. 2008) is acceptable to other elements. The figure shows that the aforementioned ratios gradually increase with decreasing  $T_{\text{eff}}$ . Finally, plotting our abundance values of [X/Fe] as a function of the stellar parameters, we detect a significant trend for the  $T_{\text{eff}}$  plot, which is presented in Fig. 4. As seen in Fig. 4, Co and Al show a systematic trend with  $T_{\text{eff}}$  in all temperature ranges and TiI, ScI, V, CrII, and Na show a trend with  $T_{\text{eff}}$  in the low-temperature domain. The higher effective temperatures of the elements from

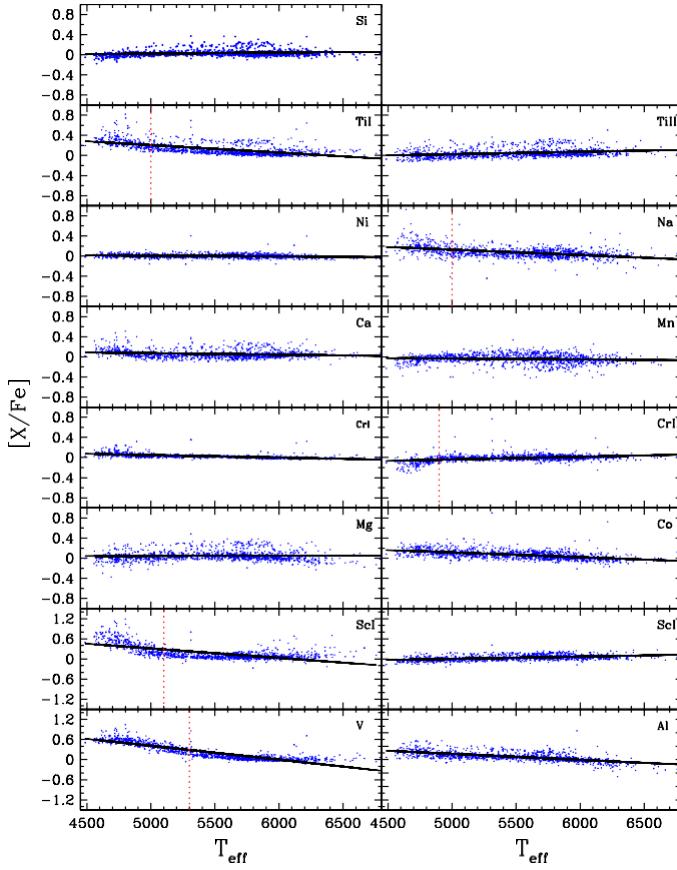


**Fig. 2.** Excitation potential slopes as a function of stellar atmospheric parameters for Ni.

which the trends appear are 4900 K for CrII, 5000 K for NaI and TiI, 5100 K for ScI and 5300 K for VI; these values are also indicated by vertical dotted lines in Fig. 4. As can be seen in Table 2, the elements and ions are very sensitive to the effective temperature, and the overestimation of the  $T_{\text{eff}}$  in the low-temperature domain might drift away from the expected abundance values. Similar trends for different elements with  $T_{\text{eff}}$  have been already noted in the literature (see e.g. Valenti & Fischer 2005; Preston et al. 2006; Gilli et al. 2006; Lai et al. 2008; Neves et al. 2009; Suda et al. 2011). As discussed in Neves et al. (2009), abundances of the cooler stars might have been overestimated due to the stronger line blending and also because the computed  $\log g$  values may be inadequate for these stars. The unexpected trends may also be connected to either deviations from excitation or ionization equilibrium, or to problems associated with the differential analysis. Finally, a possible explanation for the observed trends with  $T_{\text{eff}}$  could be an incorrect T-T relationship in the adopted model atmospheres (Lai et al. 2008). While this effect on the derived [Fe/H] abundances can be compensated for by adjusting the value of the microturbulence, this does not apply to other elements.



**Fig. 3.**  $[CrI/CrII]$  and  $[TiI/TiII]$  as a function of atmospheric parameters.



**Fig. 4.**  $[X/Fe]$  vs.  $T_{\text{eff}}$  plots. The blue dots represent the stars of the sample. The black solid lines depict the linear fits of the data. The vertical purple dotted lines indicate the “cutoff” temperature when  $[X/Fe]$  starts to show a systematic trend with  $T_{\text{eff}}$ . Each element is identified in the upper right corner of the respective plot.

This discussion indicates that, the observed trends are probably not an effect of stellar evolution, and uncertainties in atmospheric models are the dominant effect in measurements. Therefore, we chose to remove the  $T_{\text{eff}}$  trends for these elements. We fitted the data by a cubic polynomial and adding a constant

term, chosen so that the correction is zero at solar temperature. The constant term was added because simply subtracting the cubic would force the mean  $[X/Fe]$  to zero, which is an unphysical situation. A similar approach has already been applied in previous studies (see e.g. Valenti & Fischer 2005; Petigura & Marcy 2011). A sample of our results for ten stars is presented in Table 4. We present the  $[X/H]$  values before and after correction for the  $T_{\text{eff}}$  trends. The complete results are available at the CDS.

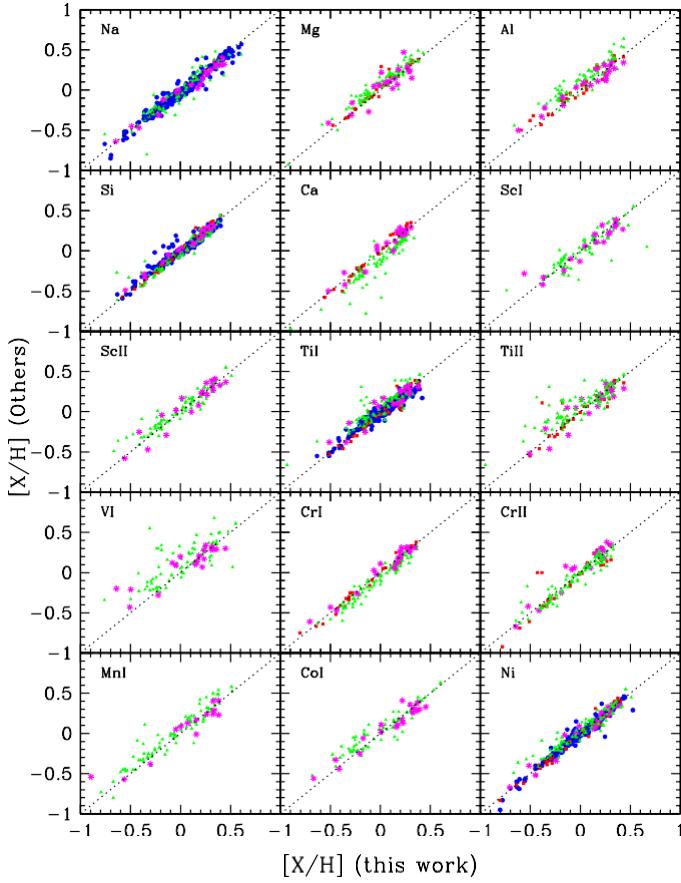
### 3.3. Comparison with previous studies

As a final check of our method and analysis, we compare our derived abundances with those obtained by Bensby et al. (2005), Valenti & Fischer (2005), Gilli et al. (2006), and Takeda (2007) for stars in common with this paper. Although we have 451 stars in common with Neves et al. (2009) and 270 with Delgado Mena et al. (2010), we do not present a comparison of the abundances, because the methods, atomic data, and the line list are almost the same. Very small differences observed for individual stars and elements during the comparison with these papers can be explained with the small differences in the line list (see the beginning of Sect. 3) and moreover for some stars we used new spectra with higher S/N compared to those used in Neves et al. (2009). We note that the comparison was performed after removing the  $T_{\text{eff}}$  trends. The results are presented in Fig. 5. As can be seen, except for the paper by Gilli et al. (2006), our results agree very well with these previous studies which lends a certain reliability to our results. Figure 5 shows that there are systematic discrepancies with Gilli et al. (2006) for most of the elements. We note that Gilli et al. (2006) also observed systematic trends with  $T_{\text{eff}}$  for some at lower effective temperatures, but they did not correct their  $[X/Fe]$  abundance ratios. Our analysis shows that the higher discrepancies show stars with  $T_{\text{eff}} < 5000$  K. Unfortunately, we do not have cool stars ( $T_{\text{eff}} < 5000$  K) in common with Bensby et al. (2005), and Takeda (2007) to test an agreement (or disagreement) at low temperatures, but we have 15 cool stars in common with Valenti & Fischer (2005), whose abundance results agree very well with those achieved in this work. We note that Valenti & Fischer (2005) also observed abundance trends with  $T_{\text{eff}}$  for some elements, and as mentioned before, they chose to remove the spurious trends. The observed discrepancies with

**Table 4.** Sample table of the derived abundances of the elements, rms, and number of measured lines for each star.

Star	...	[TiI/H]	rms	n	[TiI/H] <sub>corr*</sub>	[TiII/H]	rms	n	[VI/H]	rms	n	...
...	...	...	...	...	...	...	...	...	...	...	...	...
HD 109409	...	0.33	0.03	23	0.34	0.36	0.05	6	0.38	0.02	7	...
HD 109423	...	0.05	0.06	23	-0.06	-0.07	0.07	6	0.24	0.18	8	...
HD 109684	...	-0.27	0.05	24	-0.26	-0.24	0.03	6	-0.33	0.03	8	...
HD 109723	...	-0.01	0.06	25	-0.02	-0.10	0.03	6	0.00	0.06	8	...
HD 109988	...	0.23	0.05	23	0.14	0.15	0.07	6	0.53	0.20	8	...
HD 110291	...	0.03	0.05	23	-0.01	-0.03	0.05	6	0.08	0.07	8	...
HD 110557	...	0.04	0.04	23	-0.03	-0.05	0.04	6	0.17	0.15	8	...
HD 110619	...	-0.30	0.03	23	-0.31	-0.35	0.02	6	-0.35	0.01	8	...
HD 110668	...	0.16	0.05	23	0.16	0.22	0.01	5	0.22	0.04	8	...
HD 111031	...	0.30	0.03	24	0.30	0.29	0.03	6	0.35	0.02	7	...
...	...	...	...	...	...	...	...	...	...	...	...	...

**Notes.** (\*) The [X/H] abundances after correction for the  $T_{\text{eff}}$  trends. The full table is available at the CDS.



**Fig. 5.** Comparison of our abundance to those derived in other studies: Bensby et al. (2005) (red squares), Valenti & Fischer (2005) (blue dots), Gilli et al. (2006) (green triangles), and Takeda (2007) (magenta asterisks). The element label is located at the upper left corner of each plot.

Gilli et al. (2006) and at the same time perfect agreement with other studies confirm that applying corrections to remove the observed trends with  $T_{\text{eff}}$  is a correct approach.

#### 4. Kinematics, chemistry, and stellar populations

The Milky Way (MW) has a composite structure with several stellar subsystems. The main three stellar populations of the MW in the solar neighborhood are the thin disk, thick disk, and the halo, although most of the stars belong to the thin disk.

These populations have different kinematic and chemical properties. Generally, the thick disk is composed of relatively old (e.g. Bensby et al. 2005; Adibekyan et al. 2011), metal-poor and  $\alpha$ -enhanced (Fuhrmann 1998; Prochaska et al. 2000; Feltzing et al. 2003; Mishenina et al. 2004; Reddy et al. 2006; Haywood 2008b; Lee et al. 2011) stars that move in Galactic orbits with a large-scale height and long-scale length (Robin et al. 1996; Buser et al. 2001; Juric' et al. 2008). However, recent analyses of the geometric decompositions of the Galactic disk based on the elemental-abundance selection of the sample stars yielded strikingly different results (see e.g., Bovy et al. 2012a,b; Liu & van de Ven 2012). The latter authors found a chemo-orbital evidence that the thicker component of the MW disk is not distinct from the thin component (the MW has no thick disk – Bovy et al. 2012a), which can be explained by smooth internal evolution through radial migration (Liu & van de Ven 2012). The exceptions are the old metal-poor stars with different orbital properties that could be part of a distinct thick-disk component formed through an external mechanism (Liu & van de Ven 2012).

The first and important step in developing an understanding of the differences between the thin (thinner) and the thick (thicker) disks is to find an accurate and reliable method of assigning a star to a certain population. There is no obvious pre-determined way to identify purely thick or thin disk stars in the solar neighborhood. The main essential ways of distinguishing local thick and thin disk stars are a purely kinematical approach (e.g. Bensby et al. 2003, hereafter B03, 2005; Reddy et al. 2006), a purely chemical method (e.g. Navarro et al. 2011; Adibekyan et al. 2011), and looking at a combination of kinematics, metallicities, and stellar ages (e.g. Fuhrmann 1998; Haywood 2008a).

Although the kinematic selection is a much more common method than the chemical approach, the chemical distinction of the disks can be more useful and reliable, because chemistry is a relatively more stable property of a star than the spatial positions and kinematics. In this section we present the adopted methods to separate stars into different stellar populations on the basis of their chemistry and kinematics.

#### 4.1. Kinematical separation

To separate different stellar population by their kinematics, we computed Galactic space velocities for the stars. The space velocity components (UVW) were derived with respect to the local standard of rest, adopting the standard solar motion ( $U_S$ ,  $V_S$ ,  $W_S$ ) = (11.1, 12.24, 7.25) km s<sup>-1</sup> of Schönrich et al. (2010). The main source of the parallaxes and proper motions were the

**Table 5.** Sample table of the Galactic space velocity components and the probabilities to assign the stellar population to which each star belongs.

Star	$U_{\text{LSR}}$	$V_{\text{LSR}}$	$W_{\text{LSR}}$	B03				R03			
				$P_{\text{thick}}$	$P_{\text{thin}}$	$P_{\text{halo}}$	group	$P_{\text{thick}}$	$P_{\text{thin}}$	$P_{\text{halo}}$	group
...	...	...	...	...	...	...	...	...	...	...	...
HD 104800	122	-131	-57	0.91	0.00	0.09	thick	0.84	0.00	0.16	thick
HD 104982	72	-10	-38	0.28	0.72	0.00	thin	0.13	0.87	0.00	thin
HD 105004	-34	-225	-76	0.00	0.00	1.00	halo	0.06	0.00	0.94	halo
HD 105671	-21	0	-8	0.01	0.99	0.00	thin	0.01	0.99	0.00	thin
HD 105779	-29	-37	7	0.03	0.97	0.00	thin	0.02	0.98	0.00	thin
HD 105837	23	13	41	0.15	0.85	0.00	thin	0.08	0.92	0.00	thin
HD 105938	35	0	13	0.02	0.98	0.00	thin	0.01	0.99	0.00	thin
HD 106116	-107	8	39	0.76	0.24	0.00	thick	0.29	0.71	0.00	thin
HD 106275	18	-69	11	0.38	0.62	0.00	trans	0.11	0.89	0.00	thin
HD 104006	-21	-188	1	0.48	0.00	0.52	trans	0.69	0.00	0.31	trans
...	...	...	...	...	...	...	...	...	...	...	...

**Notes.** The full table is available at the CDS.

updated version of the Hipparcos catalog (van Leeuwen 2007). Data for eight stars with unavailable Hipparcos information were taken from the TYCHO Reference Catalog (Hog et al. 1998). The parallaxes with errors larger than 10%, (which is true for less than 5% of the stars in the sample) were redetermined following the procedure described in Sousa et al. (2011b). The percentage of stars with inaccurate proper motions (errors larger than 10%) is less than 8%. We did not perform a quality selection of them, because these errors in general do not change their membership to a certain population. The radial velocities were obtained from the HARPS spectra (courtesy of the HARPS GTO team). Combining the measurement errors in the parallaxes, proper motions, and radial velocities, the resulting average errors in the  $U$ ,  $V$ , and  $W$  velocities are of about 1 km s $^{-1}$ .

The selection of the thin disk, thick disk, and halo stars was completed using the method described in Reddy et al. (2006). This assumes that the sample is a mixture of the three populations and each population follows a Gaussian distribution of random velocities in each component (Schwarzschild 1907). Here, we adopted the mean values (asymmetric drift) and dispersion in the Gaussian distribution (characteristic velocity dispersion), and the population fractions were taken from B03 and Robin et al. (2003, hereafter R03; see also Ojha et al. 1996; Soubiran et al. 2003). We considered that a probability in excess of 70% suffices to assign a star to the concrete population. All remaining stars with a probability of less than 70% were included in a transition population. A sample of the probabilities calculated for each star according to B03 and R03, as well as Galactic space velocity components used in their calculation, are presented in Table 5. The complete results are available at the CDS.

According to the B03 criteria, among the 1111 stars in our sample, we have 964 stars from the thin disk, 78 from the thick disk, 58 are considered to be transition stars that do not belong to any group, and only 11 star belong to the halo. Adopting the criteria from R03 gives 1016 thin disk stars, 49 thick disk stars, 36 transition stars, and 10 stars belonging to the halo. We note that the B03 criteria approximately translate into the R03 criteria if  $P_{\text{thick}} > 50\%$  for a star to belong to the thick disk (Reddy et al. 2006). The distribution of stars of our sample in the Toomre diagram is shown in Fig. 6 using both the R03 and B03 criteria.

#### 4.2. Chemical separation

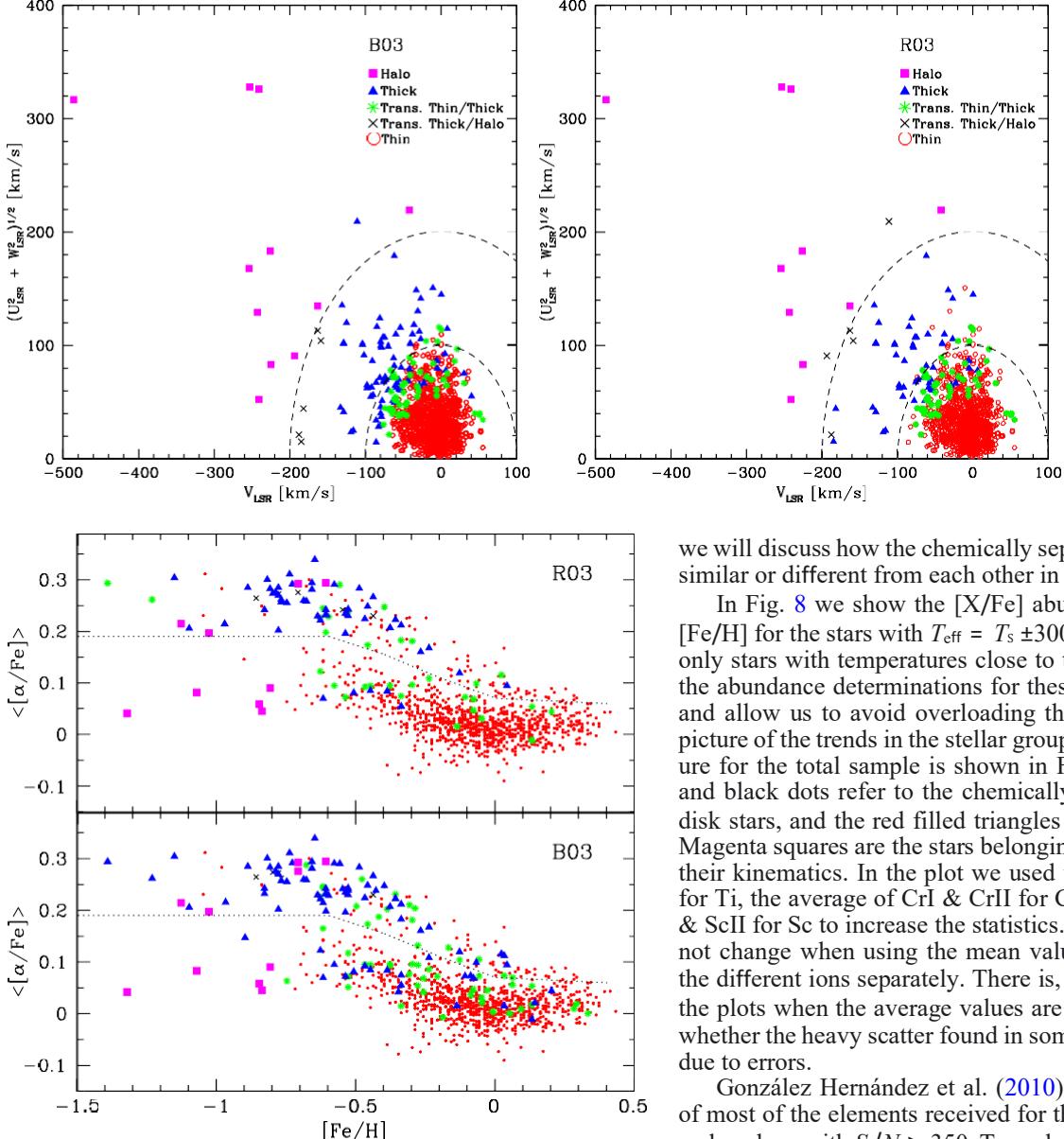
As mentioned above, in addition to the difference in their kinematics and ages, the thin- and thick disk stars are also different

in their  $\alpha$  content at a given metallicity ([Fe/H]). This dichotomy in the chemical evolution allows one to separate different stellar populations.

Adibekyan et al. (2011) showed that the stars of our sample fall into two populations, clearly separated in terms of  $[\alpha/\text{Fe}]$  (“ $\alpha$ ” refers to the average abundance of Mg, Si, and Ti) up to super-solar metallicities. We recall that Ca was not included in the  $\alpha$  index, because at solar metallicities the [Ca/Fe] trend differs from that of other  $\alpha$ -elements. In turn, high- $\alpha$  stars were also separated into two families with a gap in both  $[\alpha/\text{Fe}]$  ( $[\alpha/\text{Fe}] \approx 0.17$ ) and metallicity ( $[\text{Fe}/\text{H}] \approx -0.2$ ) distributions. This showed that the metal-rich high- $\alpha$  stars (h $\alpha$ mr) and metal-poor high- $\alpha$  (thick disk) stars are on average older than chemically defined thin disk stars (low- $\alpha$  stars). At the same time h $\alpha$ mr stars have kinematics and orbits similar to the thin disk stars.

Although in Adibekyan et al. (2011) we established a cutoff temperature for TiII because of the observed trend with  $T_{\text{eff}}$  for the [Ti/Fe] ratio (here we removed these trends, which are also observed for some other elements, see Sect. 3.2), the chemical separation of the stellar population was based on the stars with effective temperatures close to the Sun by  $\pm 300$  K. In this paper we used the chemical separation described in Adibekyan et al. (2011), i.e., thin disk, thick disk, and h $\alpha$ mr stars.

The  $[\alpha/\text{Fe}]$  versus  $[\text{Fe}/\text{H}]$  plot for the sample stars is depicted in Fig. 7. The blue triangles refer to the thick disk, red circles to the thin disk. The green asterisks and the black crosses refer to the transition stars between thin-thick and thick-halo, respectively. Magenta squares represent the stars belonging to the halo. For the kinematical separation in the top panel we used the criteria from R03, and in the bottom panel the stars are separated according to the B03 criteria. The black dashed curve separates the stars with high- and low- $\alpha$  content. Clearly, the kinematically selected samples of thick- and thin disk stars are both well mixed, judging by their  $[\alpha/\text{Fe}]$ . The chemically separated thin disk contains several kinematically hot stars that are classified as thick disk stars. Using the R03 criteria almost “cleans” the thin disk from the kinematically selected thick disk stars, but produces a high “contamination” of the chemically selected thick disk by stars with thin disk kinematics. This mixing and contamination must in part result from the fact that the assignment to the thin or the thick disk is based on probability, but the main reason could be that the stars in the local neighborhood have different birth radii and reached the solar neighborhood because of their eccentric orbits or via radial migration (e.g. Haywood 2008b; Schönrich & Binney 2009).



**Fig. 6.** Toomre diagram for the entire sample. The left and right panels show the separation of the stellar groups according to the B03 and R03 criteria, respectively. The symbols are explained in the figure.

we will discuss how the chemically separated stellar families are similar or different from each other in terms of their  $[\text{X}/\text{Fe}]$ .

In Fig. 8 we show the  $[\text{X}/\text{Fe}]$  abundance trends relative to  $[\text{Fe}/\text{H}]$  for the stars with  $T_{\text{eff}} = T_s \pm 300$  K. In the plot we present only stars with temperatures close to those of the Sun, because the abundance determinations for these stars are more accurate and allow us to avoid overloading the plot to obtain a clearer picture of the trends in the stellar groups. The corresponding figure for the total sample is shown in Fig. A.1. The blue circles and black dots refer to the chemically selected thick- and thin disk stars, and the red filled triangles represent the h $\alpha$ mr stars. Magenta squares are the stars belonging to the halo according to their kinematics. In the plot we used the average of TiI & TiII for Ti, the average of CrI & CrII for Cr, and the average of ScI & ScII for Sc to increase the statistics. The abundance trends do not change when using the mean values as compared to using the different ions separately. There is, however, lower scatter in the plots when the average values are used. It is difficult to say whether the heavy scatter found in some plots is astrophysical or due to errors.

González Hernández et al. (2010) noted the low dispersion of most of the elements received for their sample of solar twins and analogs with  $S/N > 350$ . To understand if the higher scatter found in this work is due to the quality of the data, we created a sample of solar analogs with the same stellar parameters as described in González Hernández et al. (2010). Then we devided the sample into two subsamples with  $S/N > 400$  and  $S/N < 150$ . In general, we found similar dispersions for the two subsamples, comparable with those found in González Hernández et al. (2010).

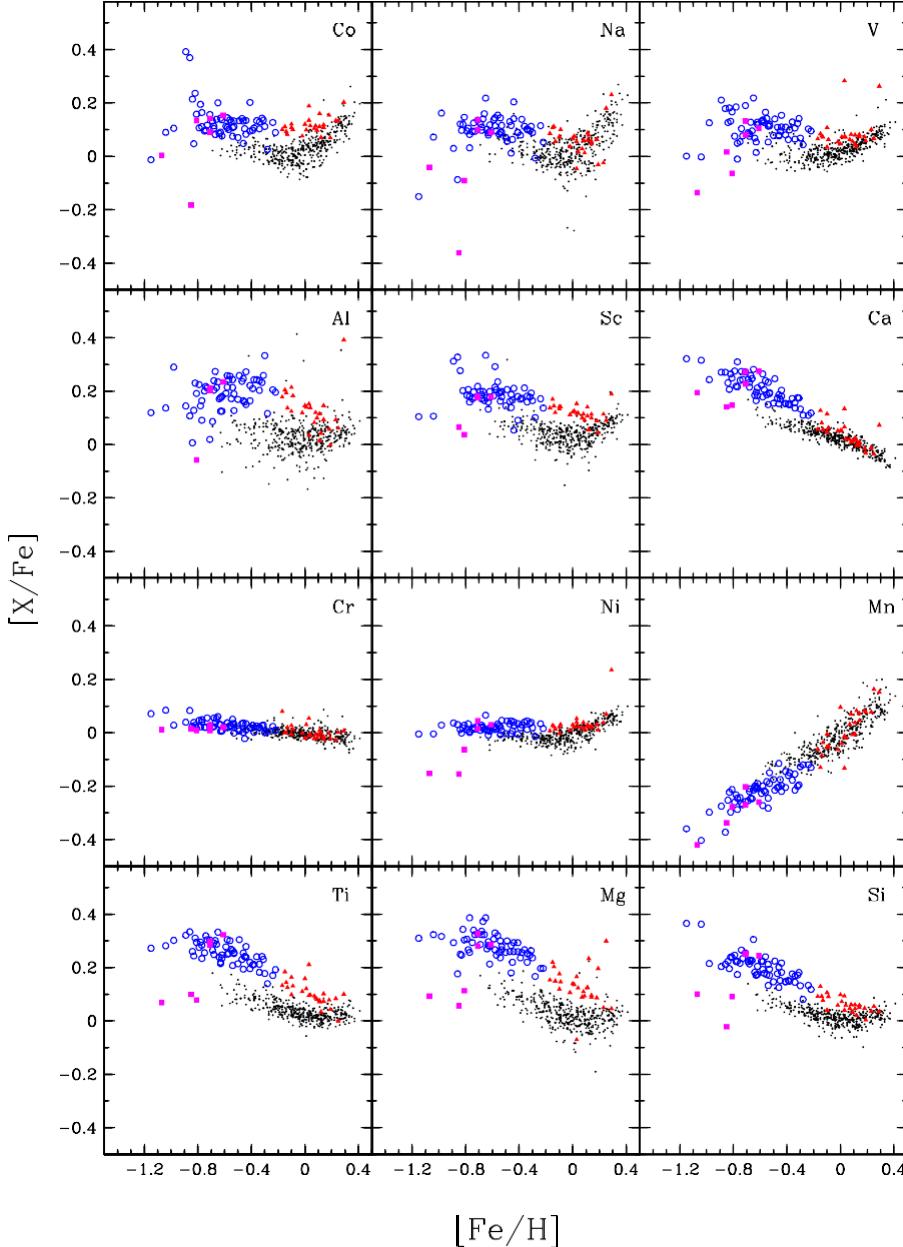
Figure 8 shows that in addition to the Mg, Si, and Ti (on which our chemical separation is based), the thin- and thick disks are chemically different for Al, Sc, Co, and Ca. There are some hints that the two disks have different Na, V, Ni, and Mn ratios, but there is no clear boundary of their  $[\text{X}/\text{Fe}]$  ratios. The only element for which the thin and the thick disks have the same  $[\text{X}/\text{Fe}]$  values is Cr. A similar result was obtained in Neves et al. (2009), who separated the thin- and thick disks according to the kinematical features of the stars.

Inspection of Fig. 8 shows that the  $\alpha$ -enhanced families, separated from the thick disk by the  $\alpha$ -element and Fe content, show different  $[\text{X}/\text{Fe}]$  trends with metallicity for different elements. As can be seen, at metallicities above solar the thin disk stars show a rise in the  $[\text{Al}/\text{Fe}]$ ,  $[\text{Sc}/\text{Fe}]$ ,  $[\text{V}/\text{Fe}]$ ,  $[\text{Ni}/\text{Fe}]$ ,  $[\text{Co}/\text{Fe}]$ , and

**Fig. 7.** Abundance ratios  $[\alpha/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  for the total sample. The blue triangles refer to the thick disk, red circles to the thin disk. The green asterisks and the black crosses refer to the transition stars between thin-thick and thick-halo, respectively. Magenta squares represent the stars belonging to the halo. The black dashed curve separates the stars with high- and low- $\alpha$  content.

#### 4.3. The $[\text{X}/\text{Fe}]$ vs. $[\text{Fe}/\text{H}]$ : the thin- and thick disks

Low-mass stars have long lifetimes and their envelopes have preserved much of their original chemical composition. Studying FGK dwarfs is very useful because they contain information about the history of the evolution of chemical abundances in the Galaxy. The  $[\text{X}/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  is traditionally used to study the Galactic chemical evolution because iron is a good chronological indicator of nucleosynthesis. In this paper we will not describe the  $[\text{X}/\text{Fe}]$  abundance trends relative to Fe because they are discussed in Neves et al. (2009), whose sample consisted of about half the number of our stars. Neves et al. (2009) also performed a detailed analysis of the  $[\text{X}/\text{Fe}]$  distributions of the kinematically separated stellar populations. In this subsection



**Fig. 8.** Abundance ratios  $[X/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  for the stars with  $T_{\text{eff}} = T_S \pm 300$  K. The blue circles and black dots refer to the chemically selected thick- and thin disk stars, and the red filled triangles are the  $\text{h}\alpha\text{mr}$  stars. Each element is identified in the *upper right corner* of the respective plot. Magenta squares represent the stars belonging to the halo according to their kinematics. The total sample is shown in Fig. A.1.

$[\text{Na}/\text{Fe}]$  (for the last two elements the rise is more pronounced and steeper), while the  $[\text{Sc}/\text{Fe}]$ ,  $[\text{Co}/\text{Fe}]$ ,  $[\text{Ni}/\text{Fe}]$ , and  $[\text{V}/\text{Fe}]$  trends for the  $\text{h}\alpha\text{mr}$  stars are essentially flat; moreover, for the  $[\text{Na}/\text{Fe}]$  and  $[\text{Al}/\text{Fe}]$  we observe a downward trend. It is interesting to see that the  $\text{h}\alpha\text{mr}$  group stars are mixed with the thin disk stars in the  $[\text{Ca}/\text{Fe}]$  vs.  $[\text{Fe}/\text{H}]$  plot, while the thick- and thin disks are separated well.

Adibekyan et al. (2011), studying the orbital properties and  $\alpha$ -element abundances of these stars, have put forward the idea that this group of stars may have originated from the inner Galactic disk. Nevertheless, their origin and exact nature still remains to be clarified.

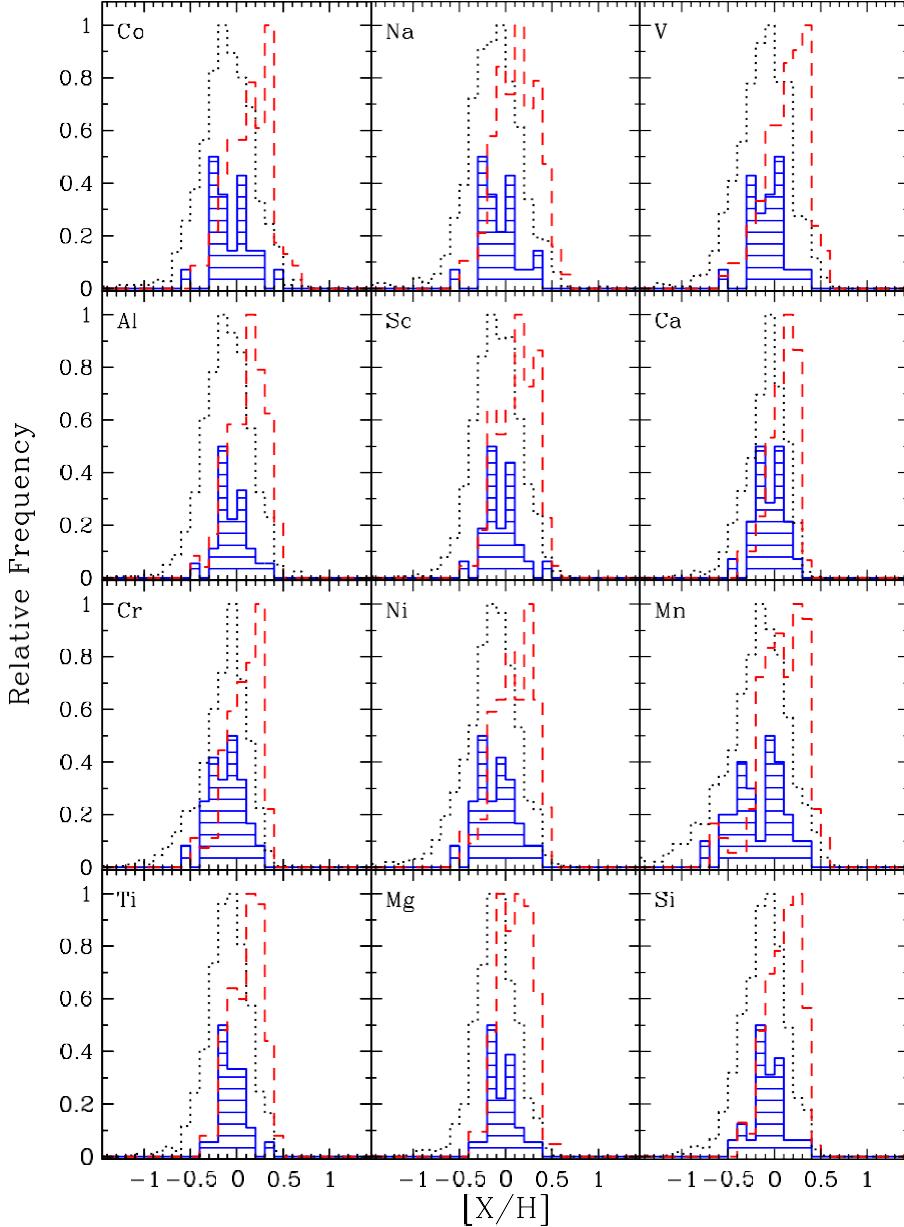
## 5. $[\text{X}/\text{H}]$ of planet-host stars

As stated before, in a separate paper we will focus on the abundance differences between the stars with and without

planets. In this section we will briefly describe the sample of planet-host and non-host stars in terms of their  $[\text{X}/\text{H}]$ .

The strong correlation is now well established between the rate of giant planets and host star metallicity. In turn, as noted before, recent studies showed that Neptune and super-Earth class planet hosts have a different metallicity distribution compared to those with giant gaseous planets. Although in this study we used relatively few PHSs (109 hosts of giants, and 26 hosts of only Neptune masses and below), this number is sufficient to observe whether there are any discernible differences in the abundances of stars without planets and planets with different masses. The  $[\text{X}/\text{H}]$  distribution histograms for planet- and non-planet hosts are depicted in Fig. 9. The stars with giant planets, without planets, and the stars hosting exclusively Neptunes and super-Earths are represented by a dashed red, dotted black, and shaded blue line, respectively.

As expected, we observe a clear metallicity excess for Jovian hosts (JH) in all spaces, which agrees well with previ-



**Fig. 9.**  $[X/H]$  distribution of the different elements. The stars with giant planets and without planets are represented by a red dashed and black dotted lines, respectively. The stars that exclusively host Neptunian and super-Earth planets are represented by a shaded blue. The Neptunian and super-Earth distribution was set smaller for clarity. The element label is located at the *upper left corner* of each plot.

ous similar studies for refractory elements (e.g. Bodaghee et al. 2003; Gilli et al. 2006; Takeda et al. 2007; Neves et al. 2009; Kang et al. 2011) and for iron (e.g. Gonzalez et al. 2001; Santos et al. 2001, 2003, 2004a, 2005; Fischer & Valenti 2005; Bond et al. 2006, 2008; Johnson et al. 2010). As already noted in the literature (e.g. Gilli et al. 2006; Neves et al. 2009), in most histograms the distributions of the abundances in JJs are not symmetrical: the distribution increases with  $[X/H]$  to a maximum value and afterward abruptly drops. The observed cutoff might suggest that this is the metallicity limit of solar neighborhood stars (e.g. Santos et al. 2003), since most of the planet hosts are at the high-metallicity end of the sample.

As can be seen from Fig. 9, the  $[X/H]$  distribution of 26 NHs in general repeats the distribution of stars without planets for all elements we studied (except that the distributions start very abruptly from the metal-poor side, probably indicating the minimum amount of some metals required to form them). This result confirms the “metal-poor” nature (e.g. Udry & Santos 2007; Sousa et al. 2008, 2011a; Mayor et al. 2011) of low-mass planet

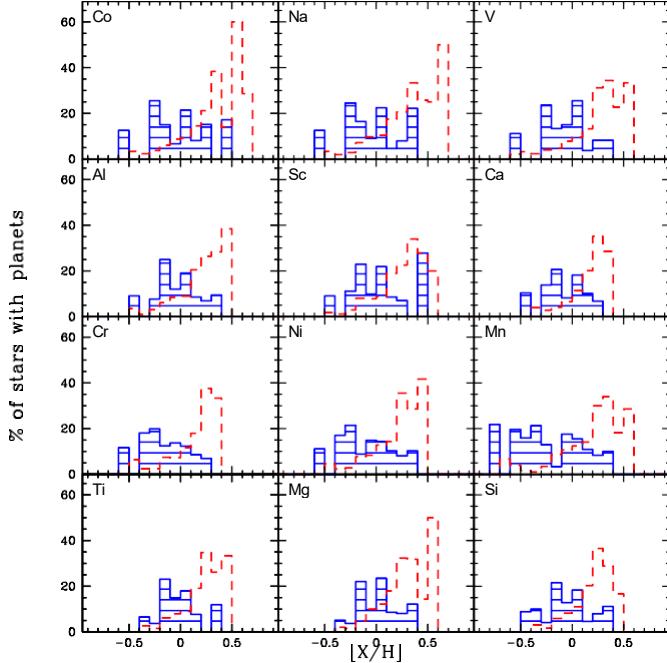
hosts, when extended to elements other than iron. The average values of  $[X/H]$  for three groups of stars, along with their rms dispersion, the number of stars used in their determination, and the difference of averages between Neptunian and Jovian hosts and stars without planets are listed in Table 6. These differences range from 0.17 (CaI) to 0.28 (MnI) for JJs and from 0.01 (CrII) to 0.09 (MgI) for NHs. These values agree well with those obtained by Neves et al. (2009) for the sample of 451 FGK stars.

Figure 10 illustrates the fraction of stars with Neptune-like and gaseous giant planets as a function of  $[X/H]$ . For each bin (the size of each bin is 0.1 dex), we divided the number of planet-bearing stars by the total number of stars in the bin. For all elements studied, we observe a continuous increase in the percentage of JJs as a function of increasing  $[X/H]$ . This result agrees with the previous findings of other authors e.g. Santos et al. (2001), Fischer & Valenti (2005), and Neves et al. (2009) for  $[Fe/H]$  and Petigura & Marcy (2011) for  $[O/H]$ ,  $[C/H]$  and  $[Fe/H]$ . Petigura & Marcy (2011), noting the small-number statistics, reported a hint of possible plateau or turnover at the

**Table 6.** Average abundances [X/H] for stars without planets, with giant planets, and stars that exclusively host Neptunians, along with their rms dispersion, the number of stars used in their determination, and the difference of averages between Neptunian and Jovian hosts and stars without planets.

Species X	Jovian hosts			Neptunian hosts			Non-planet hosts			Difference of Averages	
	([X/H])	$\sigma$	N	([X/H])	$\sigma$	N	([X/H])	$\sigma$	N	Jovian – Non-hosts	Neptunian – Non-hosts
NaI	0.12	0.23	109	-0.06	0.21	26	-0.12	0.30	975	0.24	0.06
MgI	0.10	0.18	109	-0.02	0.14	26	-0.11	0.23	976	0.21	0.09
AlI	0.10	0.19	109	-0.04	0.16	26	-0.11	0.25	969	0.21	0.07
SiI	0.10	0.18	109	-0.07	0.17	26	-0.12	0.24	976	0.22	0.05
CaI	0.08	0.15	109	-0.05	0.15	26	-0.09	0.21	976	0.17	0.04
ScI	0.14	0.21	109	-0.01	0.18	26	-0.08	0.25	947	0.22	0.07
ScII	0.11	0.20	109	-0.06	0.19	26	-0.12	0.27	976	0.23	0.06
TiI	0.11	0.17	109	-0.03	0.14	26	-0.08	0.22	976	0.19	0.05
TiII	0.10	0.18	109	-0.07	0.16	26	-0.11	0.23	976	0.21	0.04
VI	0.13	0.22	109	-0.06	0.19	26	-0.11	0.28	973	0.24	0.05
CrI	0.08	0.19	109	-0.09	0.18	26	-0.13	0.26	976	0.21	0.04
CrII	0.05	0.18	109	-0.14	0.18	26	-0.15	0.26	976	0.20	0.01
MnI	0.08	0.25	109	-0.17	0.27	26	-0.20	0.35	976	0.28	0.03
CoI	0.13	0.23	109	-0.06	0.20	26	-0.11	0.28	975	0.24	0.05
NiI	0.09	0.21	109	-0.10	0.21	26	-0.16	0.30	976	0.25	0.06
FeI	0.07	0.19	109	-0.12	0.2	26	-0.16	0.27	976	0.23	0.04
$\log g$	4.37	0.14	109	4.39	0.08	26	4.41	0.15	976	-0.04	-0.02
$\xi$	1.01	0.25	109	0.81	0.23	26	0.88	0.37	976	0.13	-0.07
$T_{\text{eff}}$	5656	412	109	5442	359	26	5490	502	976	166	-48

**Notes.** The four bottom rows list the average stellar parameters of the three aforementioned groups, taken from Sousa et al. (2008, 2011a,b).



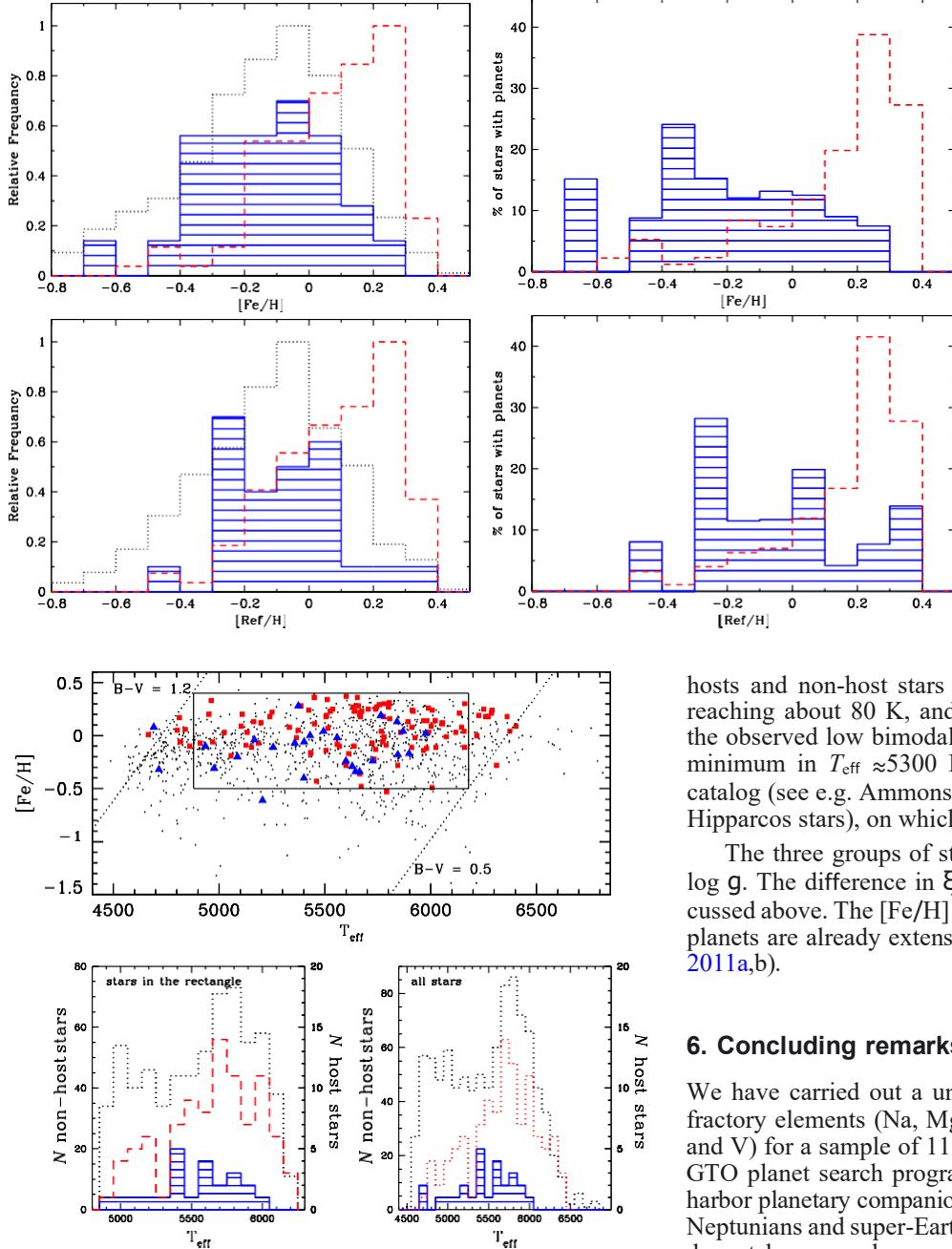
**Fig. 10.** Percentage of stars with giant (dashed red) and exclusively Neptunian and super-Earth (shaded blue) planets as a function of [X/H]. The Neptunian and super-Earth distribution was multiplied by 5 for clarity. Each element is identified in the upper left corner of the respective plot.

highest abundance bins for [C/H] and [Fe/H]. For our sample stars it is also possible to observe a small plateau or even turnover for some elements (Si, Ca, Sc, V, Cr, and Mn), but we also should note that at the highest abundance bins the number of stars sometimes does not exceed 4–5 stars.

For the fraction of low-mass planets hosts we do not observe any increasing or decreasing trends with [X/H] abundances. The distributions of the percentage of NHs are in general symmetric around the mean values listed in Table 6, which are on average less than solar abundance values by about 0.05 dex. These observations agree with the previous results for [Fe/H] (e.g. Sousa et al. 2008; Ghezzi et al. 2010; Mayor et al. 2011).

When we consider the possible dependence of planet formation on chemical composition, Gonzalez (2009) recommended to use a so-called refractory index “Ref”, which quantifies the mass abundances of refractory elements (Mg, Si and Fe) important for planet formation, rather than [Fe/H]. The importance of this index increases in the Fe-poor region, when one compares statistics of planets around the thin disk and thick disk stars. The left panel of Fig. 11 illustrates the [Fe/H] and [Ref/H] distribution histograms for planet and non-planet host stars. The fraction of stars with planets of different mass as a function of [Fe/H] and [Ref/H] are presented in the right panels. Clearly, the distributions of the three subsamples are shifted toward the higher “metallicities” in the [Ref/H] histograms, compared to their distributions in the [Fe/H]. This shift in the redistribution for planet-host stars is higher at lower metallicities, indicating their high [a/Fe] values. We again observe turnover at the highest abundance bins for [Fe/H] and [Ref/H].

The four bottom rows in Table 6 list the average stellar parameters of the three groups. It shows that hosts of low-mass planets on average have the same effective temperature as non-host stars. Interestingly, JHs are hotter by about 170 K than their non-host counterparts. The planet-search surveys are usually based on volume-limited samples, but the criteria to “cut” the sample were usually also based on the  $B - V$  color. Our sample stars mostly have  $B - V$  colors from 0.5 to 1.2. The top panel in Fig. 12 shows our sample stars in the [Fe/H] against  $T_{\text{eff}}$  (note that one star with  $T_{\text{eff}} = 7212$  K is not presented in the plot). The dotted lines represent the approximate lower and upper limits in  $B - V$  ( $B - V = 0.5$  and 1.2). The lines were constructed using



**Fig. 12.** Metallicity as a function of the effective temperature for stars with Jupiters (red circles), with Neptunes (blue triangles) and comparison sample stars (black crosses). The dotted line represents the approximate lower and upper limits in  $B-V$  ( $B-V = 0.5$  and  $1.2$ ). The  $T_{\text{eff}}$  distributions of all sample stars without planets (black dotted) and stars hosting Jovians (red dashed) and only Neptunians (shaded blue) are presented in the right bottom panel, and the distributions of the stars in the rectangle are shown in the left bottom.

the calibration equation from Sousa et al. (2008). Evidently, we missed stars with “low”  $[\text{Fe}/\text{H}]$  and “high”  $T_{\text{eff}}$  in our sample, as well as “high”  $[\text{Fe}/\text{H}]$  objects with “low”  $T_{\text{eff}}$ . To avoid these biases in  $[\text{Fe}/\text{H}]$  and  $T_{\text{eff}}$ , we cut our sample in  $[\text{Fe}/\text{H}]$  and in  $T_{\text{eff}}$ , as shown in Fig. 12. The  $T_{\text{eff}}$  distributions of all sample stars without planets and stars hosting Jovians and only Neptunians are presented in the right bottom panel of Fig. 12, and the distributions of the same groups of stars lying in the “cut rectangle” are shown in the left. The difference of average  $T_{\text{eff}}$ s of Jupiter

**Fig. 11.** Left –  $[\text{Fe}/\text{H}]$  and  $[\text{Ref}/\text{H}]$  distribution of the sample stars. The distribution lines for Jovian, Neptunian/super-Earth host and non-host stars are the same as in Fig. 9. The Neptunian and super-Earth distribution was set smaller for clarity. Right – Percentage of stars with giant (red dashed) and exclusively Neptunian and super-Earth (shaded blue) planets as a function of  $[\text{Fe}/\text{H}]$  and  $[\text{Ref}/\text{H}]$ . The Neptunian and super-Earth distribution was multiplied by 5 for better visibility.

hosts and non-host stars in the rectangle has now decreased, reaching about 80 K, and for NHs about 50 K. We note that the observed low bimodality in  $T_{\text{eff}}$  for all three groups (with a minimum in  $T_{\text{eff}} \approx 5300$  K) are inherited from the Hipparcos catalog (see e.g. Ammons et al. 2006, for the  $T_{\text{eff}}$  distribution of Hipparcos stars), on which the HARPS sample is based.

The three groups of stars have on average almost the same log  $\mathbf{g}$ . The difference in  $\xi$  reflects the difference in  $T_{\text{eff}}$ , as discussed above. The  $[\text{Fe}/\text{H}]$  distributions of stars with and without planets are already extensively discussed in Sousa et al. (2008, 2011a,b).

## 6. Concluding remarks

We have carried out a uniform abundance analysis for 12 refractory elements (Na, Mg, Al, Si, Ca, Ti, Cr, Ni, Co, Sc, Mn, and V) for a sample of 1111 FGK dwarf stars from the HARPS GTO planet search program. Of these stars, 135 are known to harbor planetary companions (26 of them are exclusively hosting Neptunians and super-Earth planets) and the remaining 976 stars do not have any known orbiting planet. The precise spectroscopic parameters for the entire sample were derived by Sousa et al. (2008, 2011a,b) in the same manner and from the same spectra as were used in the present study.

We discussed the possible sources of uncertainties and errors in our methodology in detail, and also we compared our results with those presented in other works to ensure consistency and reliability in our analysis. The large size of our sample allowed us to characterize and remove systematic abundance trends for some elements with  $T_{\text{eff}}$ .

To separate Galactic stellar populations, we applied both purely kinematical approach and chemical method. We showed that both kinematically selected thin- and thick disks are “contaminated”. The main reason of this “contamination” could be the fact that the stars in the local neighborhood have different birth radii and reached the Solar Neighborhood due to their eccentric orbits or via radial migration (e.g. Schonrich & Binney 2009).

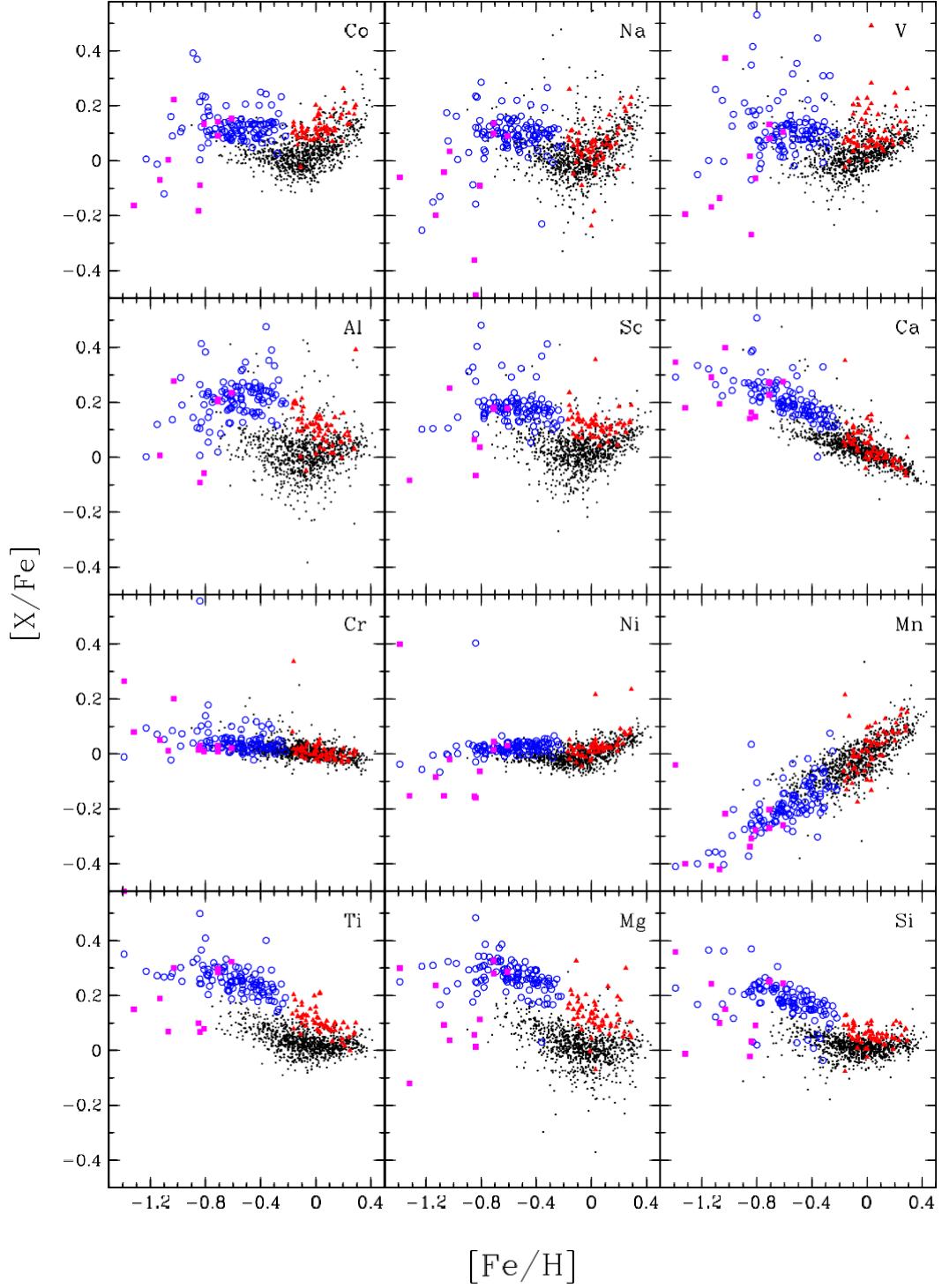
Inspection of [X/Fe] against [Fe/H] plots suggests us that chemically separated thin- and thick disks, in addition to the Mg, Si, and Ti, are also different for Al, Sc, Co, and Ca. Some bifurcation might also exist for Na, V, Ni, and Mn, but there is no clear boundary of their [X/Fe] ratios. We observed no abundance difference between the thin- and thick disks for chromium. We found that the metal-poor α-enhanced stars and their metal-rich counterparts show different [X/Fe] trends with metallicity for different elements.

We confirmed that an overabundance in giant-planet host stars is clear for all studied elements, which lends strong support to the core-accretion model of planet formation (e.g. Pollack et al. 1996). We also confirmed that stars hosting only Neptunian-like planets may be easier to detect around stars with similar metallicity than non-planet hosts, although for some elements (particularly α-elements) we observed an abrupt lower limit of [X/H], which may indicate that these elements are important in their formation. The maximum abundance difference between Neptunian-like planet hosts and non-host stars is observed for Mg ( $[\text{Mg}/\text{H}] \approx 0.09$  dex).

**Acknowledgements.** This work was supported by the European Research Council/European Community under the FP7 through Starting Grant agreement number 239953. N.C.S. also acknowledges the support from Fundação para a Ciência e a Tecnologia (FCT) through program Ciência 2007 funded by FCT/MCTES (Portugal) and POPH/FSE (EC), and in the form of grant reference PTDC/CTE-AST/098528/2008. V.Zh.A., S.G.S. and E.D.M are supported by grants SFRH/BPD/70574/2010, SFRH/BPD/47611/2008 and SFRH/BPD/76606/2011 from FCT (Portugal), respectively. J.I.G.H., and G.I. acknowledge financial support from the Spanish Ministry project MICINN AYA2011-29060 and J.I.G.H. also from the Spanish Ministry of Science and Innovation (MICINN) under the 2009 Juan de la Cierva Programme. G.Kh. would like to acknowledge the support from the CAUP-11/2011-BI fellowship. We thank the anonymous referee for its useful comments and Astrid Peter for the help concerning English.

## References

- Adibekyan, V. Zh., Santos, N. C., Sousa, S. G., & Israeli, G. 2011, A&A, 535, L11  
 Adibekyan, V. Zh., Santos, N. C., Sousa, S. G., et al. 2012, A&A, 543, A89  
 Ammons, S. M., Robinson, S. E., Strader, J., et al. 2006, ApJ, 638, 1004  
 Anders, E., & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197  
 Bensby, T., Feltzing, S., & Lundström, I. 2003, A&A, 410, 527  
 Bensby, T., Feltzing, S., Lundström, I., & Ilyin, I. 2005, A&A, 433, 185  
 Bodaghee, A., Santos, N. C., Israeli, G., & Mayor, M. 2003, A&A, 404, 715  
 Bond, J. C., Tinney, C. G., Butler, R. P., et al. 2006, MNRAS, 370, 163  
 Bond, J. C., Lauretta, D. S., Tinney, C. G., et al. 2008, ApJ, 682, 1234  
 Boss, A. P. 1997, Science, 276, 1836  
 Bovy, J., Rix, H.-W., Liu, C., et al. 2012a, ApJ, 753, 148  
 Bovy, J., Rix, H.-W., & Hogg, D. W. 2012b, ApJ, 751, 131  
 Buchhave, L., Latham, D. W., Johansen, A., et al. 2012, Nature, 486, 375  
 Buser, R., Rong, J., & Karaali, S. 1999, A&A, 348, 98  
 Cayrel, R., Depagne, E., Spite, M., et al. 2004, A&A, 416, 1117  
 Delgado Mena, E., Israeli, G., González Hernández, J. I., et al. 2010, ApJ, 725, 2349  
 Ecuvillon, A., Israeli, G., Pont, F., Santos, N. C., & Mayor, M. 2007, A&A, 461, 171  
 Feltzing, S., Bensby, T., & Lundström, I. 2003, A&A, 397, 1  
 Fischer, D. A., & Valenti, J. 2005, ApJ, 622, 1102  
 Fuhrmann, K. 1998, A&A, 338, 161  
 Gazzano, J. 2011, Ph.D. Thesis, Marseille  
 Gazzano, J., de Laverny, P., Deleuil, M., et al. 2010, A&A, 523, A91  
 Ghezzi, L., Cunha, K., Smith, V. V., et al. 2010, ApJ, 720, 1290  
 Gilli, G., Israeli, G., Ecuvillon, A., Santos, N. C., & Mayor, M. 2006, A&A, 449, 723  
 Gonzalez, G. 1998, A&A, 334, 221  
 Gonzalez, G. 2009, MNRAS, 399, L103  
 Gonzalez, G., Laws, C., Tyagi, S., & Reddy, B. E. 2001, AJ, 121, 432  
 González Hernández, J. I., Israeli, G., Santos, N. C., et al. 2010, ApJ, 720, 1592  
 Haywood, M. 2008a, A&A, 482, 673  
 Haywood, M. 2008b, MNRAS, 388, 1175  
 Hog, E., Kuzmin, A., Bastian, U., et al. 1998, A&A, 335, 65  
 Ida, S., & Lin, D. N. C. 2004, ApJ, 616, 567  
 Johnson, J. A. 2002, ApJS, 139, 219  
 Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, PASJ, 122, 905  
 Juric', M., Ivezić, Ž., Brooks, A., et al. 2008, ApJ, 673, 864  
 Kang, W., Lee, S. G., & Kim, K. M. 2011, ApJ, 736, 87  
 Kurucz, R. 1993, ATLAS9 Stellar Atmosphere Programs and 2 km s<sup>-1</sup> grid, Kurucz CD-ROM No. 13, Cambridge, Mass.: Smithsonian Astrophysical Observatory, 13  
 Lai, K. D., Bolte, M., Johnson, J. A., et al. 2008, ApJ, 681, 1524  
 Laws, C., Gonzalez, G., Walker, K. M., et al. 2003, AJ, 125, 2664  
 Lee, Y. S., Beers, T. C., An, D., et al. 2011, ApJ, 738, 187  
 Liu, C., & van de Ven, G. 2012, MNRAS, submitted [arXiv:1201.1635]  
 Lo Curto, G., Mayor, M., Benz, W., et al. 2010, A&A, 512, A48  
 Lovis, C., Mayor, M., Pepe, F., et al. 2006, Nature, 441, 305  
 Mayor, M., & Queloz, D. 1995, Nature, 378, 355  
 Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20  
 Mayor, M., Bonfils, X., Forveille, T., et al. 2009, A&A, 507, 487  
 Mayor, M., Marmier, M., Lovis, C., et al. 2011, A&A, submitted [arXiv:1109.2497]  
 Mishenina, T. V., Soubiran, C., Kovtyukh, V. V., & Korotin, S. A. 2004, A&A, 418, 551  
 Mordasini, C., Alibert, Y., & Benz, W. 2009, A&A, 501, 1139  
 Navarro, J. F., Abadi, M. G., Venn, K. A., et al. 2011, MNRAS, 412, 1203  
 Neves, V., Santos, N. C., Sousa, S. G., Correia, A. C. M., & Israeli, G. 2009, A&A, 497, 563  
 Ojha, D. K., Bienaymé, O., Robin, A. C., Creze, M., & Mohan, V. 1996, A&A, 311, 456  
 Petigura, E. A., & Marcy, G. W. 2011, ApJ, 735, 41  
 Pollack, J. B., Hubickyj, O., Bodenheimer, P., et al. 1996, Icarus, 124, 62  
 Preston, G. W., Sneden, C., Thompson, I. B., Shectman, S. A., & Burley, G. S. 2006, AJ, 132, 85  
 Prochaska, J. X., Naumov, S. O., Carney, B. W., McWilliam, A., & Wolfe, A. M. 2000, AJ, 120, 2513  
 Reddy, B. E., Lambert, D. L., & Allende Prieto, C. 2006, MNRAS, 367, 1329  
 Robin, A. C., Haywood, M., Crézé, M., Ojha, D. K., & Bienaymé, O. 1996, A&A, 305, 125  
 Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, A&A, 409, 523  
 Santos, N. C., Israeli, G., & Mayor, M. 2001, A&A, 373, 1019  
 Santos, N. C., Israeli, G., Mayor, M., Rebolo, R., & Udry, S. 2003, A&A, 398, 363  
 Santos, N. C., Israeli, G., & Mayor, M. 2004a, A&A, 415, 1153  
 Santos, N. C., Bouchy, F., Mayor, M., et al. 2004b, A&A, 426, L19  
 Santos, N. C., Israeli, G., Mayor, M., et al. 2005, A&A, 437, 1127  
 Santos, N. C., Mayor, M., Bonfils, X., et al. 2011, A&A, 526, A112  
 Schwarzschild, K. 1907, Göttingen Nachr., 614  
 Schönrich, R., & Binney, J. 2009, MNRAS, 396, 203  
 Schönrich, R., Binney, J., & Dehnen, W. 2010, MNRAS, 403, 1829  
 Sneden, C. 1973, Ph.D. Thesis, Univ. of Texas  
 Soubiran, C., Bienaymé, O., & Siebert, A. 2003, A&A, 398, 141  
 Sousa, S. G., Santos, N. C., Israeli, G., et al. 2007, A&A, 469, 783  
 Sousa, S. G., Santos, N. C., Mayor, M., et al. 2008, A&A, 487, 373  
 Sousa, S. G., Santos, N. C., Israeli, G., et al. 2011a, A&A, 533, A141  
 Sousa, S. G., Santos, N. C., Israeli, G., et al. 2011b, A&A, 526, A99  
 Suda, T., Yamada, S., Katsuta, Y., et al. 2011, MNRAS, 412, 843  
 Takeda, Y. 2007, PASJ, 59, 335  
 Udry, S., Mayor, M., Queloz, D., Naef, D., & Santos, N. 2000, Conf. Proc.: The VLT opening symposium, 571  
 Udry, S., Mayor, M., Benz, W., et al. 2006, A&A, 447, 361  
 Valenti, J. A., & Fischer, D. A. 2005, VizieR Online Data Catalog, 215, 90141  
 van Leeuwen, F. 2007, A&A, 474, 653



**Fig. A.1.** Same as Fig. 8 but for the whole sample.

# Quantum Field Theory

University of Cambridge Part III Mathematical Tripos

---

**Dr David Tong**

*Department of Applied Mathematics and Theoretical Physics,  
Centre for Mathematical Sciences,  
Wilberforce Road,  
Cambridge, CB3 OWA, UK*

<http://www.damtp.cam.ac.uk/user/tong/qft.html>  
[d.tong@damtp.cam.ac.uk](mailto:d.tong@damtp.cam.ac.uk)

## **Recommended Books and Resources**

- M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*

This is a very clear and comprehensive book, covering everything in this course at the right level. It will also cover everything in the “Advanced Quantum Field Theory” course, much of the “Standard Model” course, and will serve you well if you go on to do research. To a large extent, our course will follow the first section of this book.

There is a vast array of further Quantum Field Theory texts, many of them with redeeming features. Here I mention a few very different ones.

- S. Weinberg, *The Quantum Theory of Fields, Vol 1*

This is the first in a three volume series by one of the masters of quantum field theory. It takes a unique route to through the subject, focussing initially on particles rather than fields. The second volume covers material lectured in “AQFT”.

- L. Ryder, *Quantum Field Theory*

This elementary text has a nice discussion of much of the material in this course.

- A. Zee, *Quantum Field Theory in a Nutshell*

This is charming book, where emphasis is placed on physical understanding and the author isn’t afraid to hide the ugly truth when necessary. It contains many gems.

- M Srednicki, *Quantum Field Theory*

A very clear and well written introduction to the subject. Both this book and Zee’s focus on the path integral approach, rather than canonical quantization that we develop in this course.

There are also resources available on the web. Some particularly good ones are listed on the course webpage: <http://www.damtp.cam.ac.uk/user/tong/qft.html>

## Acknowledgements

These lecture notes are far from original. My primary contribution has been to borrow, steal and assimilate the best discussions and explanations I could find from the vast literature on the subject. I inherited the course from Nick Manton, whose notes form the backbone of the lectures. I have also relied heavily on the sources listed at the beginning, most notably the book by Peskin and Schroeder. In several places, for example the discussion of scalar Yukawa theory, I followed the lectures of Sidney Coleman, using the notes written by Brian Hill and a beautiful abridged version of these notes due to Michael Luke.

My thanks to the many who helped in various ways during the preparation of this course, including Joe Conlon, Nick Dorey, Marie Ericsson, Eyo Ita, Ian Drummond, Jerome Gauntlett, Matt Headrick, Ron Horgan, Nick Manton, Hugh Osborn and Jenni Smillie. My thanks also to the students for their sharp questions and sharp eyes in spotting typos. I am supported by the Royal Society.

## 0. Introduction

"There are no real one-particle systems in nature, not even few-particle systems. The existence of virtual pairs and of pair fluctuations shows that the days of fixed particle numbers are over."

*Viki Weisskopf*

The concept of wave-particle duality tells us that the properties of electrons and photons are fundamentally very similar. Despite obvious differences in their mass and charge, under the right circumstances both suffer wave-like diffraction and both can pack a particle-like punch.

Yet the appearance of these objects in classical physics is very different. Electrons and other matter particles are postulated to be elementary constituents of Nature. In contrast, light is a derived concept: it arises as a ripple of the electromagnetic field. If photons and particles are truly to be placed on equal footing, how should we reconcile this difference in the quantum world? Should we view the particle as fundamental, with the electromagnetic field arising only in some classical limit from a collection of quantum photons? Or should we instead view the field as fundamental, with the photon appearing only when we correctly treat the field in a manner consistent with quantum theory? And, if this latter view is correct, should we also introduce an "electron field", whose ripples give rise to particles with mass and charge? But why then didn't Faraday, Maxwell and other classical physicists find it useful to introduce the concept of matter fields, analogous to the electromagnetic field?

The purpose of this course is to answer these questions. We shall see that the second viewpoint above is the most useful: the field is primary and particles are derived concepts, appearing only after quantization. We will show how photons arise from the quantization of the electromagnetic field and how massive, charged particles such as electrons arise from the quantization of matter fields. We will learn that in order to describe the fundamental laws of Nature, we must not only introduce electron fields, but also quark fields, neutrino fields, gluon fields, W and Z-boson fields, Higgs fields and a whole slew of others. There is a field associated to each type of fundamental particle that appears in Nature.

### Why Quantum Field Theory?

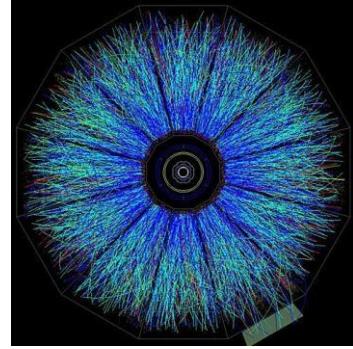
In classical physics, the primary reason for introducing the concept of the field is to construct laws of Nature that are *local*. The old laws of Coulomb and Newton involve "action at a distance". This means that the force felt by an electron (or planet) changes

immediately if a distant proton (or star) moves. This situation is philosophically unsatisfactory. More importantly, it is also experimentally wrong. The field theories of Maxwell and Einstein remedy the situation, with all interactions mediated in a local fashion by the field.

The requirement of locality remains a strong motivation for studying field theories in the quantum world. However, there are further reasons for treating the quantum field as fundamental<sup>1</sup>. Here I'll give two answers to the question: Why quantum field theory?

**Answer 1:** Because the combination of quantum mechanics and special relativity implies that particle number is not conserved.

Particles are not indestructible objects, made at the beginning of the universe and here for good. They can be created and destroyed. They are, in fact, mostly ephemeral and fleeting. This experimentally verified fact was first predicted by Dirac who understood how relativity implies the necessity of anti-particles. An extreme demonstration of particle creation is shown in the picture, which comes from the Relativistic Heavy Ion Collider (RHIC) at Brookhaven, Long Island. This machine crashes gold nuclei together, each containing 197 nucleons. The resulting explosion contains up to 10,000 particles, captured here in all their beauty by the STAR detector.



**Figure 1:**

We will review Dirac's argument for anti-particles later in this course, together with the better understanding that we get from viewing particles in the framework of quantum field theory. For now, we'll quickly sketch the circumstances in which we expect the number of particles to change. Consider a particle of mass  $m$  trapped in a box of size  $L$ . Heisenberg tells us that the uncertainty in the momentum is  $\Delta p \geq k/L$ . In a relativistic setting, momentum and energy are on an equivalent footing, so we should also have an uncertainty in the energy of order  $\Delta E \geq kc/L$ . However, when the uncertainty in the energy exceeds  $\Delta E = 2mc^2$ , then we cross the barrier to pop particle anti-particle pairs out of the vacuum. We learn that particle-anti-particle pairs are expected to be important when a particle of mass  $m$  is localized within a distance of order

$$\lambda = \frac{k}{mc}$$

<sup>1</sup>A concise review of the underlying principles and major successes of quantum field theory can be found in the article by Frank Wilczek, <http://arxiv.org/abs/hep-th/9803075>

At distances shorter than this, there is a high probability that we will detect particle-anti-particle pairs swarming around the original particle that we put in. The distance  $\lambda$  is called the *Compton wavelength*. It is always smaller than the de Broglie wavelength  $\lambda_{dB} = h/|\mathbf{p}|$ . If you like, the de Broglie wavelength is the distance at which the wavelike nature of particles becomes apparent; the Compton wavelength is the distance at which the concept of a single pointlike particle breaks down completely.

The presence of a multitude of particles and antiparticles at short distances tells us that any attempt to write down a relativistic version of the one-particle Schrödinger equation (or, indeed, an equation for any *fixed* number of particles) is doomed to failure. There is no mechanism in standard non-relativistic quantum mechanics to deal with changes in the particle number. Indeed, any attempt to naively construct a relativistic version of the one-particle Schrödinger equation meets with serious problems. (Negative probabilities, infinite towers of negative energy states, or a breakdown in causality are the common issues that arise). In each case, this failure is telling us that once we enter the relativistic regime we need a new formalism in order to treat states with an unspecified number of particles. This formalism is quantum field theory (QFT).

**Answer 2:** Because all particles of the same type are the same

This sound rather dumb. But it's not! What I mean by this is that two electrons are identical in every way, regardless of where they came from and what they've been through. The same is true of every other fundamental particle. Let me illustrate this through a rather prosaic story. Suppose we capture a proton from a cosmic ray which we identify as coming from a supernova lying 8 billion lightyears away. We compare this proton with one freshly minted in a particle accelerator here on Earth. And the two are exactly the same! How is this possible? Why aren't there errors in proton production? How can two objects, manufactured so far apart in space and time, be identical in all respects? One explanation that might be offered is that there's a sea of proton "stuff" filling the universe and when we make a proton we somehow dip our hand into this stuff and from it mould a proton. Then it's not surprising that protons produced in different parts of the universe are identical: they're made of the same stuff. It turns out that this is roughly what happens. The "stuff" is the proton field or, if you look closely enough, the quark and gluon fields.

In fact, there's more to this tale. Being the "same" in the quantum world is not like being the "same" in the classical world: quantum particles that are the same are truly indistinguishable. Swapping two particles around leaves the state completely unchanged — apart from a possible minus sign. This minus sign determines the statistics of the particle. In quantum mechanics you have to put these statistics in by hand and,

to agree with experiment, should choose Bose statistics (no minus sign) for integer spin particles, and Fermi statistics (yes minus sign) for half-integer spin particles. In quantum field theory, this relationship between spin and statistics is not something that you have to put in by hand. Rather, it is a consequence of the framework.

## What is Quantum Field Theory?

Having told you why QFT is necessary, I should really tell you what it is. The clue is in the name: it is the quantization of a classical field, the most familiar example of which is the electromagnetic field. In standard quantum mechanics, we're taught to take the classical degrees of freedom and promote them to operators acting on a Hilbert space. The rules for quantizing a field are no different. Thus the basic degrees of freedom in quantum field theory are *operator valued functions of space and time*. This means that we are dealing with an infinite number of degrees of freedom — at least one for every point in space. This infinity will come back to bite on several occasions.

It will turn out that the possible interactions in quantum field theory are governed by a few basic principles: locality, symmetry and renormalization group flow (the decoupling of short distance phenomena from physics at larger scales). These ideas make QFT a very robust framework: given a set of fields there is very often an almost unique way to couple them together.

## What is Quantum Field Theory Good For?

The answer is: almost everything. As I have stressed above, for any relativistic system it is a necessity. But it is also a very useful tool in non-relativistic systems with many particles. Quantum field theory has had a major impact in condensed matter, high-energy physics, cosmology, quantum gravity and pure mathematics. It is literally the language in which the laws of Nature are written.

### 0.1 Units and Scales

Nature presents us with three fundamental dimensionful constants; the speed of light  $c$ , Planck's constant (divided by  $2\pi$ )  $k$  and Newton's constant  $G$ . They have dimensions

$$\begin{aligned}[c] & [c] = LT^{-1} \\ & [k] = L^2 MT^{-1} \\ & [G] = L^3 M^{-1} T^{-2} \end{aligned}$$

Throughout this course we will work with “natural” units, defined by

$$c = k = 1 \tag{0.1}$$

which allows us to express all dimensionful quantities in terms of a single scale which we choose to be mass or, equivalently, energy (since  $E = mc^2$  has become  $E = m$ ). The usual choice of energy unit is  $eV$ , the electron volt or, more often  $GeV = 10^9 eV$  or  $TeV = 10^{12} eV$ . To convert the unit of energy back to a unit of length or time, we need to insert the relevant powers of  $c$  and  $k$ . For example, the length scale  $\lambda$  associated to a mass  $m$  is the Compton wavelength

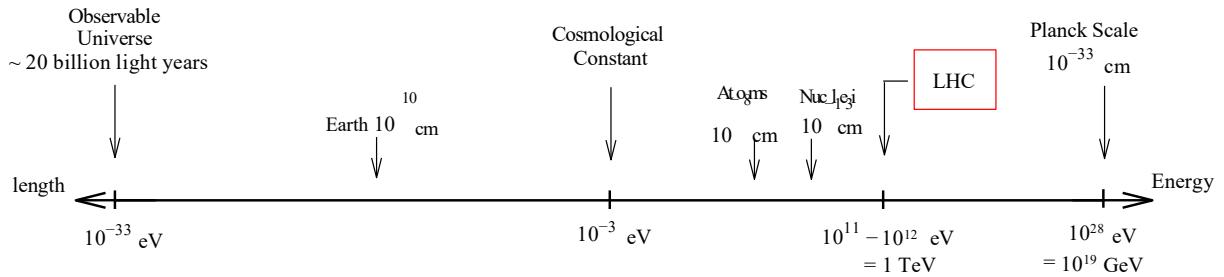
$$\lambda = \frac{k}{mc}$$

With this conversion factor, the electron mass  $m_e = 10^6 eV$  translates to a length scale  $\lambda_e \sim 10^{-12} m$ . (The Compton wavelength is also defined with an extra factor of  $2\pi$ :  $\lambda = 2\pi k/mc$ .)

Throughout this course we will refer to the *dimension* of a quantity, meaning the mass dimension. If  $X$  has dimensions of  $(\text{mass})^d$  we will write  $[X] = d$ . In particular, the surviving natural quantity  $G$  has dimensions  $[G] = -2$  and defines a mass scale,

$$G = \frac{kc}{M_p^2} = \frac{1}{M_p^2} \quad (0.2)$$

where  $M_p \approx 10^{19} GeV$  is the *Planck scale*. It corresponds to a length  $l_p \approx 10^{-33} cm$ . The Planck scale is thought to be the smallest length scale that makes sense: beyond this quantum gravity effects become important and it's no longer clear that the concept of spacetime makes sense. The largest length scale we can talk of is the size of the cosmological horizon, roughly  $10^{60} l_p$ .



**Figure 2:** Energy and Distance Scales in the Universe

Some useful scales in the universe are shown in the figure. This is a logarithmic plot, with energy increasing to the right and, correspondingly, length increasing to the left. The smallest and largest scales known are shown on the figure, together with other

relevant energy scales. The standard model of particle physics is expected to hold up to about the  $TeV$ . This is precisely the regime that is currently being probed by the Large Hadron Collider (LHC) at CERN. There is a general belief that the framework of quantum field theory will continue to hold to energy scales only slightly below the Planck scale — for example, there are experimental hints that the coupling constants of electromagnetism, and the weak and strong forces unify at around  $10^{18} \text{ GeV}$ .

For comparison, the rough masses of some elementary (and not so elementary) particles are shown in the table,

Particle	Mass
Neutrinos	$\sim 10^{-2} \text{ eV}$
Electron	0.5 MeV
Muon	100 MeV
Pions	140 MeV
Proton, Neutron	1 GeV
Tau	2 GeV
W,Z Bosons	80-90 GeV
Higgs Boson	125 GeV

## 1. Classical Field Theory

In this first section we will discuss various aspects of classical fields. We will cover only the bare minimum ground necessary before turning to the quantum theory, and will return to classical field theory at several later stages in the course when we need to introduce new ideas.

### 1.1 The Dynamics of Fields

A *field* is a quantity defined at every point of space and time ( $\mathbf{x}$ ,  $t$ ). While classical particle mechanics deals with a finite number of generalized coordinates  $q_a(t)$ , indexed by a label  $a$ , in field theory we are interested in the dynamics of fields

$$\varphi_a(\mathbf{x}, t) \quad (1.1)$$

where both  $a$  and  $\mathbf{x}$  are considered as labels. Thus we are dealing with a system with an infinite number of degrees of freedom — at least one for each point  $\mathbf{x}$  in space. Notice that the concept of position has been relegated from a dynamical variable in particle mechanics to a mere label in field theory.

#### An Example: The Electromagnetic Field

The most familiar examples of fields from classical physics are the electric and magnetic fields,  $\vec{E}(\mathbf{x}, t)$  and  $\vec{B}(\mathbf{x}, t)$ . Both of these fields are spatial 3-vectors. In a more sophisticated treatment of electromagnetism, we derive these two 3-vectors from a single 4-component field  $A^\mu(\mathbf{x}, t) = (\varphi, \vec{A})$  where  $\mu = 0, 1, 2, 3$  shows that this field is a vector in *spacetime*. The electric and magnetic fields are given by

$$\vec{E} = -\nabla\varphi - \frac{\partial\vec{A}}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A} \quad (1.2)$$

which ensure that two of Maxwell's equations,  $\nabla \cdot \vec{B} = 0$  and  $d\vec{B}/dt = -\nabla \times \vec{E}$ , hold immediately as identities.

#### The Lagrangian

The dynamics of the field is governed by a Lagrangian which is a function of  $\varphi(\mathbf{x}, t)$ ,  $\dot{\varphi}(\mathbf{x}, t)$  and  $\nabla\varphi(\mathbf{x}, t)$ . In all the systems we study in this course, the Lagrangian is of the form,

$$L(t) = \int d^3x L(\varphi_a, \partial_\mu\varphi_a) \quad (1.3)$$

where the official name for  $L$  is the *Lagrangian density*, although everyone simply calls it the Lagrangian. The action is,

$$S = \int_{t_1}^{t_2} dt \int d^3x L = \int d^4x L \quad (1.4)$$

Recall that in particle mechanics  $L$  depends on  $q$  and  $\dot{q}$ , but not  $\ddot{q}$ . In field theory we similarly restrict to Lagrangians  $L$  depending on  $\varphi$  and  $\dot{\varphi}$ , and not  $\ddot{\varphi}$ . In principle, there's nothing to stop  $L$  depending on  $\nabla\varphi$ ,  $\nabla^2\varphi$ ,  $\nabla^3\varphi$ , etc. However, with an eye to later Lorentz invariance, we will only consider Lagrangians depending on  $\nabla\varphi$  and not higher derivatives. Also we will not consider Lagrangians with explicit dependence on  $x^\mu$ ; all such dependence only comes through  $\varphi$  and its derivatives.

We can determine the equations of motion by the principle of least action. We vary the path, keeping the end points fixed and require  $\delta S = 0$ ,

$$\begin{aligned} \delta S &= \int d^4x \frac{\partial L}{\partial \varphi_a} \delta \varphi_a + \frac{\partial L}{\partial (\partial_\mu \varphi_a)} \delta (\partial_\mu \varphi_a) \\ &= \int d^4x \frac{\partial \varphi_a}{\partial \varphi_a} - \partial_\mu \frac{\partial L}{\partial (\partial_\mu \varphi_a)} \delta \varphi_a + \partial_\mu \frac{\partial L}{\partial (\partial_\mu \varphi_a)} \delta \varphi_a \end{aligned} \quad (1.5)$$

The last term is a total derivative and vanishes for any  $\delta \varphi_a(\mathbf{x}, t)$  that decays at spatial infinity and obeys  $\delta \varphi_a(\mathbf{x}, t_1) = \delta \varphi_a(\mathbf{x}, t_2) = 0$ . Requiring  $\delta S = 0$  for all such paths yields the Euler-Lagrange equations of motion for the fields  $\varphi_a$ ,

$$\partial_\mu \frac{\partial L}{\partial (\partial_\mu \varphi_a)} - \frac{\partial L}{\partial \varphi_a} = 0 \quad (1.6)$$

### 1.1.1 An Example: The Klein-Gordon Equation

Consider the Lagrangian for a real scalar field  $\varphi(\mathbf{x}, t)$ ,

$$\begin{aligned} L &= \frac{1}{2} \eta^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi - \frac{1}{2} m^2 \varphi^2 \\ &= \frac{1}{2} \dot{\varphi}^2 - \frac{1}{2} (\nabla \varphi)^2 - \frac{1}{2} m^2 \varphi^2 \end{aligned} \quad (1.7)$$

where we are using the Minkowski space metric

$$\eta^{\mu\nu} = \eta_{\mu\nu} = \begin{matrix} +1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{matrix} \quad (1.8)$$

Comparing (1.7) to the usual expression for the Lagrangian  $L = T - V$ , we identify the kinetic energy of the field as

$$T = \int d^3x \frac{1}{2} \dot{\varphi}^2 \quad (1.9)$$

and the potential energy of the field as

$$V = \int d^3x \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m^2\varphi^2 \quad (1.10)$$

The first term in this expression is called the gradient energy, while the phrase “potential energy”, or just “potential”, is usually reserved for the last term.

To determine the equations of motion arising from (1.7), we compute

$$\frac{\partial L}{\partial\varphi} = -m^2\varphi \quad \text{and} \quad \frac{\partial L}{\partial(\partial_\mu\varphi)} = \partial^\mu\varphi \equiv (\dot{\varphi}, -\nabla\varphi) \quad (1.11)$$

The Euler-Lagrange equation is then

$$\ddot{\varphi} - \nabla^2\varphi + m^2\varphi = 0 \quad (1.12)$$

which we can write in relativistic form as

$$\partial_\mu\partial^\mu\varphi + m^2\varphi = 0 \quad (1.13)$$

This is the *Klein-Gordon Equation*. The Laplacian in Minkowski space is sometimes denoted by  $\square$ . In this notation, the Klein-Gordon equation reads  $\square\varphi + m^2\varphi = 0$ .

An obvious generalization of the Klein-Gordon equation comes from considering the Lagrangian with arbitrary potential  $V(\varphi)$ ,

$$L = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - V(\varphi) \Rightarrow \partial_\mu\partial^\mu\varphi + \frac{\partial V}{\partial\varphi} = 0 \quad (1.14)$$

### 1.1.2 Another Example: First Order Lagrangians

We could also consider a Lagrangian that is linear in time derivatives, rather than quadratic. Take a complex scalar field  $\psi$  whose dynamics is defined by the real Lagrangian

$$L = \frac{i}{2}(\psi^\wedge\dot{\psi} - \dot{\psi}^\wedge\psi) - \nabla\psi^\wedge \cdot \nabla\psi - m\psi^\wedge\psi \quad (1.15)$$

We can determine the equations of motion by treating  $\psi$  and  $\psi^\wedge$  as independent objects, so that

$$\frac{\partial L}{\partial\psi^\wedge} = \frac{i}{2}\dot{\psi} - m\psi \quad \text{and} \quad \frac{\partial L}{\partial\dot{\psi}^\wedge} = -\frac{i}{2}\psi \quad \text{and} \quad \frac{\partial L}{\partial\nabla\psi^\wedge} = -\nabla\psi \quad (1.16)$$

This gives us the equation of motion

$$i\frac{\partial\psi}{\partial t} = -\nabla^2\psi + m\psi \quad (1.17)$$

This looks very much like the Schrödinger equation. Except it isn't! Or, at least, the interpretation of this equation is very different: the field  $\psi$  is a classical field with none of the probability interpretation of the wavefunction. We'll come back to this point in Section 2.8.

The initial data required on a Cauchy surface differs for the two examples above. When  $L \sim \dot{\varphi}^2$ , both  $\varphi$  and  $\dot{\varphi}$  must be specified to determine the future evolution; however when  $L \sim \psi^\wedge \dot{\psi}$ , only  $\psi$  and  $\dot{\psi}$  are needed.

### 1.1.3 A Final Example: Maxwell's Equations

We may derive Maxwell's equations in the vacuum from the Lagrangian,

$$L = -\frac{1}{2} (\partial_\mu A_\nu) \eta^{\mu\nu} + \frac{1}{2} (\partial_\mu A_0)^2 \quad (1.18)$$

Notice the funny minus signs! This is to ensure that the kinetic terms for  $A_i$  are positive using the Minkowski space metric (1.8), so  $L \sim \frac{1}{2} \dot{A}_i^2$ . The Lagrangian (1.18) has no kinetic term  $A_0^2$  for  $A_0$ . We will see the consequences of this in Section 6. To see that Maxwell's equations indeed follow from (1.18), we compute

$$\frac{\partial L}{\partial(\partial_\mu A_\nu)} = -\partial^\mu A^\nu + (\partial_\rho A^\rho) \eta^{\mu\nu} \quad (1.19)$$

from which we may derive the equations of motion,

$$\partial_\mu \frac{\partial L}{\partial(\partial_\mu A_\nu)} = -\partial^2 A^\nu + \partial^\nu (\partial_\rho A^\rho) = -\partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) \equiv -\partial_\mu F^{\mu\nu} \quad (1.20)$$

where the *field strength* is defined by  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ . You can check using (1.2) that this reproduces the remaining two Maxwell's equations in a vacuum:  $\nabla \cdot \vec{E} = 0$  and  $\partial \vec{E} / \partial t = \nabla \times \vec{B}$ . Using the notation of the field strength, we may rewrite the Maxwell Lagrangian (up to an integration by parts) in the compact form

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (1.21)$$

### 1.1.4 Locality, Locality, Locality

In each of the examples above, the Lagrangian is *local*. This means that there are no terms in the Lagrangian coupling  $\varphi(\mathbf{x}, t)$  directly to  $\varphi(\mathbf{y}, t)$  with  $\mathbf{x} \neq \mathbf{y}$ . For example, there are no terms that look like

$$L = \int d^3x d^3y \varphi(\mathbf{x}) \varphi(\mathbf{y}) \quad (1.22)$$

A priori, there's no reason for this. After all,  $\mathbf{x}$  is merely a label, and we're quite happy to couple other labels together (for example, the term  $\partial_3 A_0 \partial_0 A_3$  in the Maxwell Lagrangian couples the  $\mu = 0$  field to the  $\mu = 3$  field). But the closest we get for the  $\mathbf{x}$  label is a coupling between  $\varphi(\mathbf{x})$  and  $\varphi(\mathbf{x} + \delta \rightarrow \mathbf{x})$  through the gradient term  $(\nabla \varphi)^2$ . This property of locality is, as far as we know, a key feature of *all* theories of Nature. Indeed, one of the main reasons for introducing field theories in classical physics is to implement locality. In this course, we will only consider local Lagrangians.

## 1.2 Lorentz Invariance

The laws of Nature are relativistic, and one of the main motivations to develop quantum field theory is to reconcile quantum mechanics with special relativity. To this end, we want to construct field theories in which space and time are placed on an equal footing and the theory is invariant under Lorentz transformations,

$$x^\mu \longrightarrow (x^r)^\mu = \Lambda^\mu_\nu x^\nu \quad (1.23)$$

where  $\Lambda^\mu_\nu$  satisfies

$$\sum_{\sigma} \eta^{\sigma\tau} \Lambda^\nu_\tau = \eta^{\mu\nu} \quad (1.24)$$

For example, a rotation by  $\vartheta$  about the  $x^3$ -axis, and a boost by  $v < 1$  along the  $x^1$ -axis are respectively described by the Lorentz transformations

$$\Lambda^\mu_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \vartheta & -\sin \vartheta & 0 \\ 0 & \sin \vartheta & \cos \vartheta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \Lambda^\mu_\nu = \begin{pmatrix} \gamma & -\gamma v & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1.25)$$

with  $\gamma = \sqrt{1 - v^2}$ . The Lorentz transformations form a Lie group under matrix multiplication. You'll learn more about this in the "Symmetries and Particle Physics" course.

The Lorentz transformations have a *representation* on the fields. The simplest example is the scalar field which, under the Lorentz transformation  $x \rightarrow \Lambda x$ , transforms as

$$\varphi(x) \rightarrow \varphi^r(x) = \varphi(\Lambda^{-1}x) \quad (1.26)$$

The inverse  $\Lambda^{-1}$  appears in the argument because we are dealing with an *active* transformation in which the field is truly shifted. To see why this means that the inverse appears, it will suffice to consider a non-relativistic example such as a temperature field. Suppose we start with an initial field  $\varphi(\mathbf{x})$  which has a hotspot at, say,  $\mathbf{x} = (1, 0, 0)$ . After a rotation  $\mathbf{x} \rightarrow R \mathbf{x}$  about the  $z$ -axis, the new field  $\varphi^r(\mathbf{x})$  will have the hotspot at  $\mathbf{x} = (0, 1, 0)$ . If we want to express  $\varphi^r(\mathbf{x})$  in terms of the old field  $\varphi$ , we need to place ourselves at  $\mathbf{x} = (0, 1, 0)$  and ask what the old field looked like where we've come from at  $R^{-1}(0, 1, 0) = (1, 0, 0)$ . This  $R^{-1}$  is the origin of the inverse transformation. (If we were instead dealing with a passive transformation in which we relabel our choice of coordinates, we would have instead  $\varphi(x) \rightarrow \varphi^r(x) = \varphi(\Lambda x)$ ).

The definition of a Lorentz invariant theory is that if  $\varphi(x)$  solves the equations of motion then  $\varphi(\Lambda^{-1}x)$  also solves the equations of motion. We can ensure that this property holds by requiring that the action is Lorentz invariant. Let's look at our examples:

### Example 1: The Klein-Gordon Equation

For a real scalar field we have  $\varphi(x) \rightarrow \varphi^r(x) = \varphi(\Lambda^{-1}x)$ . The derivative of the scalar field transforms as a vector, meaning

$$(\partial_\mu \varphi)(x) \rightarrow (\Lambda^{-1})^\nu_\mu (\partial_\nu \varphi)(y)$$

where  $y = \Lambda^{-1}x$ . This means that the derivative terms in the Lagrangian density transform as

$$\begin{aligned} L_{\text{deriv}}(x) &= \partial_\mu \varphi(x) \partial_\nu \varphi(x) \eta^{\mu\nu} \longrightarrow (\Lambda^{-1})^\rho_\mu (\partial_\rho \varphi)(y) (\Lambda^{-1})^\sigma_\nu (\partial_\sigma \varphi)(y) \eta^{\mu\nu} \\ &= (\partial_\rho \varphi)(y) (\partial_\sigma \varphi)(y) \eta^{\rho\sigma} \\ &= L_{\text{deriv}}(y) \end{aligned} \quad (1.27)$$

The potential terms transform in the same way, with  $\varphi^2(x) \rightarrow \varphi^2(y)$ . Putting this all together, we find that the action is indeed invariant under Lorentz transformations,

$$S = \int d^4x L(x) \longrightarrow \int d^4x L(y) = \int d^4y L(y) = S \quad (1.28)$$

where, in the last step, we need the fact that we don't pick up a Jacobian factor when we change integration variables from  $d^4x$  to  $d^4y$ . This follows because  $\det \Lambda = 1$ . (At least for Lorentz transformations connected to the identity which, for now, is all we deal with).

### Example 2: First Order Dynamics

In the first-order Lagrangian (1.15), space and time are not on the same footing. ( $L$  is linear in time derivatives, but quadratic in spatial derivatives). The theory is not Lorentz invariant.

In practice, it's easy to see if the action is Lorentz invariant: just make sure all the Lorentz indices  $\mu = 0, 1, 2, 3$  are contracted with Lorentz invariant objects, such as the metric  $\eta_{\mu\nu}$ . Other Lorentz invariant objects you can use include the totally antisymmetric tensor  $\epsilon_{\mu\nu\rho\sigma}$  and the matrices  $\gamma_\mu$  that we will introduce when we come to discuss spinors in Section 4.

### Example 3: Maxwell's Equations

Under a Lorentz transformation  $A^\mu(x) \rightarrow \Lambda_v^\mu A^v(\Lambda^{-1}x)$ . You can check that Maxwell's Lagrangian (1.21) is indeed invariant. Of course, historically electrodynamics was the first Lorentz invariant theory to be discovered: it was found even before the concept of Lorentz invariance.

### 1.3 Symmetries

The role of symmetries in field theory is possibly even more important than in particle mechanics. There are Lorentz symmetries, internal symmetries, gauge symmetries, supersymmetries.... We start here by recasting Noether's theorem in a field theoretic framework.

#### 1.3.1 Noether's Theorem

Every continuous symmetry of the Lagrangian gives rise to a conserved current  $j^\mu(x)$  such that the equations of motion imply

$$\partial_\mu j^\mu = 0 \quad (1.29)$$

or, in other words,  $\partial j^0 / \partial t + \nabla \cdot \vec{j} = 0$ .

**A Comment:** A conserved current implies a conserved charge  $Q$ , defined as

$$Q = \int_{\mathbf{R}^3} d^3x j^0 \quad (1.30)$$

which one can immediately see by taking the time derivative,

$$\frac{dQ}{dt} = \int_{\mathbf{R}^3} d^3x \frac{\partial j^0}{\partial t} = - \int_{\mathbf{R}^3} d^3x \nabla \cdot \vec{j} = 0 \quad (1.31)$$

assuming that  $\vec{j} \rightarrow 0$  sufficiently quickly as  $|x| \rightarrow \infty$ . However, the existence of a current is a much stronger statement than the existence of a conserved charge because it implies that charge is conserved *locally*. To see this, we can define the charge in a finite volume  $V$ ,

$$Q_V = \int_V d^3x j^0 \quad (1.32)$$

Repeating the analysis above, we find that

$$\frac{dQ_V}{dt} = - \int_V d^3x \nabla \cdot \vec{j} = - \int_A \vec{j} \cdot d\vec{s} \quad (1.33)$$

where  $A$  is the area bounding  $V$  and we have used Stokes' theorem. This equation means that any charge leaving  $V$  must be accounted for by a flow of the current 3-vector  $\vec{j}$  out of the volume. This kind of local conservation of charge holds in any local field theory.

**Proof of Noether's Theorem:** We'll prove the theorem by working infinitesimally. We may always do this if we have a continuous symmetry. We say that the transformation

$$\delta\varphi_a(x) = X_a(\varphi) \quad (1.34)$$

is a symmetry if the Lagrangian changes by a total derivative,

$$\delta L = \partial_\mu F^\mu \quad (1.35)$$

for some set of functions  $F^\mu(\varphi)$ . To derive Noether's theorem, we first consider making an *arbitrary* transformation of the fields  $\delta\varphi_a$ . Then

$$\begin{aligned} \delta L &= \frac{\partial L}{\partial\varphi_a} \delta\varphi_a + \frac{\partial L}{\partial(\partial_\mu\varphi_a)} \partial_\mu (\delta\varphi_a) \\ &= \frac{\partial L}{\partial\varphi_a} - \partial_\mu \frac{\partial L}{\partial(\partial_\mu\varphi_a)} \delta\varphi_a + \partial_\mu \left[ \frac{\partial L}{\partial(\partial_\mu\varphi_a)} \right] \delta\varphi_a \end{aligned} \quad (1.36)$$

When the equations of motion are satisfied, the term in square brackets vanishes. So we're left with

$$\delta L = \partial_\mu \left[ \frac{\partial L}{\partial(\partial_\mu\varphi_a)} \right] \delta\varphi_a \quad (1.37)$$

But for the symmetry transformation  $\delta\varphi_a = X_a(\varphi)$ , we have by definition  $\delta L = \partial_\mu F^\mu$ . Equating this expression with (1.37) gives us the result

$$\partial_\mu j^\mu = 0 \quad \text{with} \quad j^\mu = \frac{\partial L}{\partial(\partial_\mu\varphi_a)} X_a(\varphi) - F^\mu(\varphi) \quad (1.38)$$

### 1.3.2 An Example: Translations and the Energy-Momentum Tensor

Recall that in classical particle mechanics, invariance under spatial translations gives rise to the conservation of momentum, while invariance under time translations is responsible for the conservation of energy. We will now see something similar in field theories. Consider the infinitesimal translation

$$x^\nu \rightarrow x^\nu - \epsilon^\nu \quad \Rightarrow \quad \varphi_a(x) \rightarrow \varphi_a(x) + \epsilon^\nu \partial_\nu \varphi_a(x) \quad (1.39)$$

(where the sign in the field transformation is plus, instead of minus, because we're doing an active, as opposed to passive, transformation). Similarly, once we substitute a specific field configuration  $\varphi(x)$  into the Lagrangian, the Lagrangian itself also transforms as

$$L(x) \rightarrow L(x) + \epsilon^v \partial_v L(x) \quad (1.40)$$

Since the change in the Lagrangian is a total derivative, we may invoke Noether's theorem which gives us four conserved currents  $(j^\mu)_v$ , one for each of the translations  $\epsilon^v$  with  $v = 0, 1, 2, 3$ ,

$$(j^\mu)_v = \frac{\partial L}{\partial(\partial_\mu \varphi_a)} \partial_v \varphi_a - \delta_v^\mu L \equiv T_\nu^\mu \quad (1.41)$$

$T_\nu^\mu$  is called the *energy-momentum tensor*. It satisfies

$$\partial_\mu T_\nu^\mu = 0 \quad (1.42)$$

The four conserved quantities are given by

$$E = \int d^3x T^{00} \quad \text{and} \quad P^i = \int d^3x T^{0i} \quad (1.43)$$

where  $E$  is the total energy of the field configuration, while  $P^i$  is the total momentum of the field configuration.

### An Example of the Energy-Momentum Tensor

Consider the simplest scalar field theory with Lagrangian (1.7). From the above discussion, we can compute

$$T^{\mu\nu} = \partial^\mu \varphi \partial^\nu \varphi - \eta^{\mu\nu} L \quad (1.44)$$

One can verify using the equation of motion for  $\varphi$  that this expression indeed satisfies  $\partial_\mu T^{\mu\nu} = 0$ . For this example, the conserved energy and momentum are given by

$$E = \int d^3x \frac{1}{2} \dot{\varphi}^2 + \frac{1}{2} (\nabla \varphi)^2 + \frac{1}{2} m^2 \varphi^2 \quad (1.45)$$

$$P^i = \int d^3x \varphi^i \partial^i \varphi \quad (1.46)$$

Notice that for this example,  $T^{\mu\nu}$  came out symmetric, so that  $T^{\mu\nu} = T^{\nu\mu}$ . This won't always be the case. Nevertheless, there is typically a way to massage the energy-momentum tensor of any theory into a symmetric form by adding an extra term

$$\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\rho \Gamma^{\rho\mu\nu} \quad (1.47)$$

where  $\Gamma^{\rho\mu\nu}$  is some function of the fields that is anti-symmetric in the first two indices so  $\Gamma^{\rho\mu\nu} = -\Gamma^{\mu\rho\nu}$ . This guarantees that  $\partial_\mu \partial_\rho \Gamma^{\rho\mu\nu} = 0$  so that the new energy-momentum tensor is also a conserved current.

## A Cute Trick

One reason that you may want a symmetric energy-momentum tensor is to make contact with general relativity: such an object sits on the right-hand side of Einstein's field equations. In fact this observation provides a quick and easy way to determine a symmetric energy-momentum tensor. Firstly consider coupling the theory to a curved background spacetime, introducing an arbitrary metric  $g_{\mu\nu}(x)$  in place of  $\eta_{\mu\nu}$ , and replacing the kinetic terms with suitable covariant derivatives using "minimal coupling". Then a symmetric energy momentum tensor in the flat space theory is given by

$$\Theta^{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\partial(\sqrt{-g}L)}{\partial g_{\mu\nu}} \quad g_{\mu\nu} = \eta_{\mu\nu} \quad (1.48)$$

It should be noted however that this trick requires a little more care when working with spinors.

### 1.3.3 Another Example: Lorentz Transformations and Angular Momentum

In classical particle mechanics, rotational invariance gave rise to conservation of angular momentum. What is the analogy in field theory? Moreover, we now have further Lorentz transformations, namely boosts. What conserved quantity do they correspond to? To answer these questions, we first need the infinitesimal form of the Lorentz transformations

$$\Lambda^\mu_v = \delta^\mu_v + \omega^\mu_v \quad (1.49)$$

where  $\omega^\mu_v$  is infinitesimal. The condition (1.24) for  $\Lambda$  to be a Lorentz transformation becomes

$$\begin{aligned} (\delta^\mu_\sigma + \omega^\mu_\sigma)(\delta^\nu_\tau + \omega^\nu_\tau) \eta^{\sigma\tau} &= \eta^{\mu\nu} \\ \Rightarrow \quad \omega^{\mu\nu} + \omega^{\nu\mu} &= 0 \end{aligned} \quad (1.50)$$

So the infinitesimal form  $\omega^{\mu\nu}$  of the Lorentz transformation must be an anti-symmetric matrix. As a check, the number of different  $4 \times 4$  anti-symmetric matrices is  $4 \times 3/2 = 6$ , which agrees with the number of different Lorentz transformations (3 rotations + 3 boosts). Now the transformation on a scalar field is given by

$$\begin{aligned} \varphi(x) \rightarrow \varphi'(x) &= \varphi(\Lambda^{-1}x) \\ &= \varphi(x^\mu - \omega^\mu_v x^v) \\ &= \varphi(x^\mu) - \omega^\mu_v x^v \partial_\mu \varphi(x) \end{aligned} \quad (1.51)$$

from which we see that

$$\delta\varphi = -\omega^\mu_v x^\nu \partial_\mu \varphi \quad (1.52)$$

By the same argument, the Lagrangian density transforms as

$$\delta L = -\omega^\mu_v x^\nu \partial_\mu L = -\partial_\mu (\omega^\mu_v x^\nu L) \quad (1.53)$$

where the last equality follows because  $\omega^\mu_\mu = 0$  due to anti-symmetry. Once again, the Lagrangian changes by a total derivative so we may apply Noether's theorem (now with  $F^\mu = -\omega^\mu_v x^\nu L$ ) to find the conserved current

$$\begin{aligned} j^\mu &= -\frac{\partial L}{\partial(\partial_\mu \varphi)} \omega^\rho_v x^\nu \partial_\rho \varphi + \omega^\mu_v x^\nu L \\ &= -\omega^\rho_v \frac{\partial L}{\partial(\partial_\mu \varphi)} x^\nu \partial_\rho \varphi - \delta^\mu_\rho x^\nu L = -\omega^\rho_v T^\mu_\rho x^\nu \end{aligned} \quad (1.54)$$

Unlike in the previous example, I've left the infinitesimal choice of  $\omega^\mu_v$  in the expression for this current. But really, we should strip it out to give six different currents, i.e. one for each choice of  $\omega^\mu_v$ . We can write them as

$$(J^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} \quad (1.55)$$

which satisfy  $\partial_\mu (J^\mu)^{\rho\sigma} = 0$  and give rise to 6 conserved charges. For  $\rho, \sigma = 1, 2, 3$ , the Lorentz transformation is a rotation and the three conserved charges give the total angular momentum of the field.

$$Q^{ij} = \int d^3x (x^i T^{0j} - x^j T^{0i}) \quad (1.56)$$

But what about the boosts? In this case, the conserved charges are

$$Q^{0i} = \int d^3x (x^0 T^{0i} - x^i T^{00}) \quad (1.57)$$

The fact that these are conserved tells us that

$$\begin{aligned} 0 &= \frac{dQ^{0i}}{dt} = \int d^3x \frac{\partial T^{0i}}{\partial t} + t \int d^3x \frac{d}{dt} \frac{d}{dt} \int d^3x x^i T^{00} \\ &= P^i + t \frac{dP^i}{dt} - \int d^3x x^i T^{00} \end{aligned} \quad (1.58)$$

But we know that  $P^i$  is conserved, so  $dP^i/dt = 0$ , leaving us with the following consequence of invariance under boosts:

$$\frac{d}{dt} \int d^3x x^i T^{00} = \text{constant} \quad (1.59)$$

This is the statement that the center of energy of the field travels with a constant velocity. It's kind of like a field theoretic version of Newton's first law but, rather surprisingly, appearing here as a conservation law.

### 1.3.4 Internal Symmetries

The above two examples involved transformations of spacetime, as well as transformations of the field. An *internal symmetry* is one that only involves a transformation of the fields and acts the same at every point in spacetime. The simplest example occurs for a complex scalar field  $\psi(x) = (\varphi_1(x) + i\varphi_2(x))/\sqrt{2}$ . We can build a real Lagrangian by

$$L = \partial_\mu \psi^\lambda \partial^\mu \psi - V(|\psi|^2) \quad (1.60)$$

where the potential is a general polynomial in  $|\psi|^2 = \psi^\lambda \psi$ . To find the equations of motion, we could expand  $\psi$  in terms of  $\varphi_1$  and  $\varphi_2$  and work as before. However, it's easier (and equivalent) to treat  $\psi$  and  $\psi^\lambda$  as independent variables and vary the action with respect to both of them. For example, varying with respect to  $\psi^\lambda$  leads to the equation of motion

$$\partial_\mu \partial^\mu \psi + \frac{\partial V(\psi^\lambda \psi)}{\partial \psi^\lambda} = 0 \quad (1.61)$$

The Lagrangian has a continuous symmetry which rotates  $\varphi_1$  and  $\varphi_2$  or, equivalently, rotates the phase of  $\psi$ :

$$\psi \rightarrow e^{i\alpha} \psi \quad \text{or} \quad \delta\psi = i\alpha \psi \quad (1.62)$$

where the latter equation holds with  $\alpha$  infinitesimal. The Lagrangian remains invariant under this change:  $\delta L = 0$ . The associated conserved current is

$$j^\mu = i(\partial^\mu \psi^\lambda) \psi - i\psi^\lambda (\partial^\mu \psi) \quad (1.63)$$

We will later see that the conserved charges arising from currents of this type have the interpretation of electric charge or particle number (for example, baryon or lepton number).

### Non-Abelian Internal Symmetries

Consider a theory involving  $N$  scalar fields  $\varphi_a$ , all with the same mass and the Lagrangian

$$L = \frac{1}{2} \sum_{a=1}^N \partial_\mu \varphi_a \partial^\mu \varphi_a - \frac{1}{2} \sum_{a=1}^N m^2 \varphi_a^2 - g \sum_{a=1}^N \varphi_a^2 \quad (1.64)$$

In this case the Lagrangian is invariant under the non-Abelian symmetry group  $G = SO(N)$ . (Actually  $O(N)$  in this case). One can construct theories from complex fields in a similar manner that are invariant under an  $SU(N)$  symmetry group. Non-Abelian symmetries of this type are often referred to as *global* symmetries to distinguish them from the “local gauge” symmetries that you will meet later. Isospin is an example of such a symmetry, albeit realized only approximately in Nature.

## Another Cute Trick

There is a quick method to determine the conserved current associated to an internal symmetry  $\delta\varphi = \alpha\varphi$  for which the Lagrangian is invariant. Here,  $\alpha$  is a constant real number. (We may generalize the discussion easily to a non-Abelian internal symmetry for which  $\alpha$  becomes a matrix). Now consider performing the transformation but where  $\alpha$  depends on spacetime:  $\alpha = \alpha(x)$ . The action is no longer invariant. However, the change must be of the form

$$\delta L = (\partial_\mu \alpha) h^\mu(\varphi) \quad (1.65)$$

since we know that  $\delta L = 0$  when  $\alpha$  is constant. The change in the action is therefore

$$\delta S = \int d^4x \delta L = - \int d^4x \alpha(x) \partial_\mu h^\mu \quad (1.66)$$

which means that when the equations of motion are satisfied (so  $\delta S = 0$  for all variations, including  $\delta\varphi = \alpha(x)\varphi$ ) we have

$$\partial_\mu h^\mu = 0 \quad (1.67)$$

We see that we can identify the function  $h^\mu = j^\mu$  as the conserved current. This way of viewing things emphasizes that it is the derivative terms, not the potential terms, in the action that contribute to the current. (The potential terms are invariant even when  $\alpha = \alpha(x)$ ).

## 1.4 The Hamiltonian Formalism

The link between the Lagrangian formalism and the quantum theory goes via the path integral. In this course we will not discuss path integral methods, and focus instead on canonical quantization. For this we need the Hamiltonian formalism of field theory. We start by defining the *momentum*  $\pi^a(x)$  conjugate to  $\varphi_a(x)$ ,

$$\pi^a(x) = \frac{\partial L}{\partial \dot{\varphi}_a} \quad (1.68)$$

The conjugate momentum  $\pi^a(x)$  is a function of  $x$ , just like the field  $\varphi_a(x)$  itself. It is not to be confused with the total momentum  $P^i$  defined in (1.43) which is a single number characterizing the whole field configuration. The *Hamiltonian density* is given by

$$H = \pi^a(x) \dot{\varphi}_a(x) - L(x) \quad (1.69)$$

where, as in classical mechanics, we eliminate  $\dot{\varphi}_a(x)$  in favour of  $\pi^a(x)$  everywhere in  $H$ . The Hamiltonian is then simply

$$H = \int d^3x H \quad (1.70)$$

## An Example: A Real Scalar Field

For the Lagrangian

$$L = \frac{1}{2}\dot{\varphi}^2 - \frac{1}{2}(\nabla\varphi)^2 - V(\varphi) \quad (1.71)$$

the momentum is given by  $\pi = \dot{\varphi}$ , which gives us the Hamiltonian,

$$H = \int d^3x \left[ \frac{1}{2}\pi^2 + \frac{1}{2}(\nabla\varphi)^2 + V(\varphi) \right] \quad (1.72)$$

Notice that the Hamiltonian agrees with the definition of the total energy (1.45) that we get from applying Noether's theorem for time translation invariance.

In the Lagrangian formalism, Lorentz invariance is clear for all to see since the action is invariant under Lorentz transformations. In contrast, the Hamiltonian formalism is *not* manifestly Lorentz invariant: we have picked a preferred time. For example, the equations of motion for  $\varphi(x) = \varphi(\mathbf{x}, t)$  arise from Hamilton's equations,

$$\dot{\varphi}(\mathbf{x}, t) = \frac{\partial H}{\partial \pi(\mathbf{x}, t)} \quad \text{and} \quad \dot{\pi}(\mathbf{x}, t) = -\frac{\partial H}{\partial \varphi(\mathbf{x}, t)} \quad (1.73)$$

which, unlike the Euler-Lagrange equations (1.6), do not look Lorentz invariant. Nevertheless, even though the Hamiltonian framework doesn't *look* Lorentz invariant, the physics must remain unchanged. If we start from a relativistic theory, all final answers must be Lorentz invariant even if it's not manifest at intermediate steps. We will pause at several points along the quantum route to check that this is indeed the case.

## 2. Free Fields

“The career of a young theoretical physicist consists of treating the harmonic oscillator in ever-increasing levels of abstraction.”

Sidney Coleman

### 2.1 Canonical Quantization

In quantum mechanics, canonical quantization is a recipe that takes us from the Hamiltonian formalism of classical dynamics to the quantum theory. The recipe tells us to take the generalized coordinates  $q_a$  and their conjugate momenta  $p^a$  and promote them to operators. The Poisson bracket structure of classical mechanics morphs into the structure of commutation relations between operators, so that, in units with  $\hbar = 1$ ,

$$\begin{aligned}[q_a, q_b] &= [p^a, p^b] = 0 \\ [q_a, p^b] &= i \delta_a^b\end{aligned}\tag{2.1}$$

In field theory we do the same, now for the field  $\varphi_a(\mathbf{x})$  and its momentum conjugate  $\pi^b(\mathbf{x})$ . Thus a *quantum field* is an operator valued function of space obeying the commutation relations

$$\begin{aligned}[\varphi_a(\mathbf{x}), \varphi_b(\mathbf{y})] &= [\pi^a(\mathbf{x}), \pi^b(\mathbf{y})] = 0 \\ [\varphi_a(\mathbf{x}), \pi^b(\mathbf{y})] &= i \delta^{(3)}(\mathbf{x} - \mathbf{y}) \delta_a^b\end{aligned}\tag{2.2}$$

Note that we've lost all track of Lorentz invariance since we have separated space  $\mathbf{x}$  and time  $t$ . We are working in the Schrödinger picture so that the operators  $\varphi_a(\mathbf{x})$  and  $\pi^a(\mathbf{x})$  do not depend on time at all — only on space. All time dependence sits in the states  $|\psi\rangle$  which evolve by the usual Schrödinger equation

$$i \frac{d|\psi\rangle}{dt} = H |\psi\rangle\tag{2.3}$$

We aren't doing anything different from usual quantum mechanics; we're merely applying the old formalism to fields. Be warned however that the notation  $|\psi\rangle$  for the state is deceptively simple: if you were to write the wavefunction in quantum field theory, it would be a *functional*, that is a function of every possible configuration of the field  $\varphi$ .

The typical information we want to know about a quantum theory is the spectrum of the Hamiltonian  $H$ . In quantum field theories, this is usually *very hard*. One reason for this is that we have an infinite number of degrees of freedom — at least one for every point  $\mathbf{x}$  in space. However, for certain theories — known as *free theories* — we can find a way to write the dynamics such that each degree of freedom evolves independently

from all the others. Free field theories typically have Lagrangians which are quadratic in the fields, so that the equations of motion are linear. For example, the simplest relativistic free theory is the classical Klein-Gordon (KG) equation for a real scalar field  $\varphi(\mathbf{x}, t)$ ,

$$\partial_\mu \partial^\mu \varphi + m^2 \varphi = 0 \quad (2.4)$$

To exhibit the coordinates in which the degrees of freedom decouple from each other, we need only take the Fourier transform,

$$\varphi(\mathbf{x}, t) = \int \frac{d^3 p}{(2\pi)^3} e^{i\mathbf{p}\cdot\mathbf{x}} \varphi(\mathbf{p}, t) \quad (2.5)$$

Then  $\varphi(\mathbf{p}, t)$  satisfies

$$\frac{\partial^2}{\partial t^2} + (\mathbf{p}^2 + m^2) \varphi(\mathbf{p}, t) = 0 \quad (2.6)$$

Thus, for each value of  $\mathbf{p}$ ,  $\varphi(\mathbf{p}, t)$  solves the equation of a harmonic oscillator vibrating at frequency

$$\omega_{\mathbf{p}} = \sqrt{\mathbf{p}^2 + m^2} \quad (2.7)$$

We learn that the most general solution to the KG equation is a linear superposition of simple harmonic oscillators, each vibrating at a different frequency with a different amplitude. To quantize  $\varphi(\mathbf{x}, t)$  we must simply quantize this infinite number of harmonic oscillators. Let's recall how to do this.

### 2.1.1 The Simple Harmonic Oscillator

Consider the quantum mechanical Hamiltonian

$$H = \frac{1}{2} \mathbf{p}^2 + \frac{1}{2} \omega^2 \mathbf{q}^2 \quad (2.8)$$

with the canonical commutation relations  $[\mathbf{q}, \mathbf{p}] = i$ . To find the spectrum we define the creation and annihilation operators (also known as raising/lowering operators, or sometimes ladder operators)

$$a = \frac{i}{2\omega} \mathbf{q} + \frac{1}{2\omega} \mathbf{p} \quad , \quad a^\dagger = \frac{i}{2\omega} \mathbf{q} - \frac{1}{2\omega} \mathbf{p} \quad (2.9)$$

which can be easily inverted to give

$$\mathbf{q} = \frac{1}{2\omega} (a + a^\dagger) \quad , \quad \mathbf{p} = -i \frac{\omega}{2} (a - a^\dagger) \quad (2.10)$$

Substituting into the above expressions we find

$$[a, a^\dagger] = 1 \quad (2.11)$$

while the Hamiltonian is given by

$$\begin{aligned} H &= \frac{1}{2}\omega(aa^\dagger + a^\dagger a) \\ &= \omega(a^\dagger a + \frac{1}{2}) \end{aligned} \quad (2.12)$$

One can easily confirm that the commutators between the Hamiltonian and the creation and annihilation operators are given by

$$[H, a^\dagger] = \omega a^\dagger \quad \text{and} \quad [H, a] = -\omega a \quad (2.13)$$

These relations ensure that  $a$  and  $a^\dagger$  take us between energy eigenstates. Let  $|E\rangle$  be an eigenstate with energy  $E$ , so that  $H|E\rangle = E|E\rangle$ . Then we can construct more eigenstates by acting with  $a$  and  $a^\dagger$ ,

$$Ha^\dagger|E\rangle = (E + \omega)a^\dagger|E\rangle, \quad Ha|E\rangle = (E - \omega)a|E\rangle \quad (2.14)$$

So we find that the system has a ladder of states with energies

$$\dots, E - \omega, E, E + \omega, E + 2\omega, \dots \quad (2.15)$$

If the energy is bounded below, there must be a *ground state*  $|0\rangle$  which satisfies  $a|0\rangle = 0$ . This has ground state energy (also known as zero point energy),

$$H|0\rangle = \frac{1}{2}\omega|0\rangle \quad (2.16)$$

Excited states then arise from repeated application of  $a^\dagger$ ,

$$|n\rangle = (a^\dagger)^n|0\rangle \quad \text{with} \quad H|n\rangle = (n + \frac{1}{2})\omega|n\rangle \quad (2.17)$$

where I've ignored the normalization of these states so,  $\langle n|n\rangle \neq 1$ .

## 2.2 The Free Scalar Field

We now apply the quantization of the harmonic oscillator to the free scalar field. We write  $\varphi$  and  $\pi$  as a linear sum of an infinite number of creation and annihilation operators  $a_{\vec{p}}^\dagger$  and  $a_{\vec{p}}$ , indexed by the 3-momentum  $\vec{p}$ ,

$$\varphi(\vec{x}) = \int \frac{d^3p}{(2\pi)^3} \frac{\sqrt{1}}{2\omega_{\vec{p}}} h \left[ a_{\vec{p}} e^{i\vec{p} \cdot \vec{x}} + a_{\vec{p}}^\dagger e^{-i\vec{p} \cdot \vec{x}} \right] \quad (2.18)$$

$$\pi(\vec{x}) = \frac{i}{(2\pi)^3} (-i) \frac{\omega_{\vec{p}}}{2} h \left[ a_{\vec{p}} e^{i\vec{p} \cdot \vec{x}} - a_{\vec{p}}^\dagger e^{-i\vec{p} \cdot \vec{x}} \right] \quad (2.19)$$

**Claim:** The commutation relations for  $\varphi$  and  $\pi$  are equivalent to the following commutation relations for  $a_{p\rightarrow}$  and  $a_{p\rightarrow}^\dagger$

$$\begin{aligned} [\varphi(\mathbf{x}), \varphi(\mathbf{y})] &= [\pi(\mathbf{x}), \pi(\mathbf{y})] = 0 & [a_{p\rightarrow} a_{q\rightarrow}] &= [a_{p\rightarrow}^\dagger a_{q\rightarrow}^\dagger] = 0 \\ &\Leftrightarrow & & \\ [\varphi(\mathbf{x}), \pi(\mathbf{y})] &= i\delta^{(3)}(\mathbf{x} \rightarrow \mathbf{y}) & [a_{p\rightarrow} a_{q\rightarrow}^\dagger] &= (2\pi)^3 \delta^{(3)}(\mathbf{p} \rightarrow \mathbf{q}) \end{aligned} \quad (2.20)$$

**Proof:** We'll show this just one way. Assume that  $[a, a_{p\rightarrow}^\dagger] = (2\pi)^3 \delta^{(3)}(\mathbf{p} \rightarrow \mathbf{q})$ . Then

$$\begin{aligned} [\varphi(\mathbf{x}), \pi(\mathbf{y})] &= \int \frac{d^3 p}{(2\pi)^6} \frac{d^3 q}{2} \frac{(-i)}{\omega_{p\rightarrow}} - [a_{p\rightarrow} a_{q\rightarrow}^\dagger] e^{ip\cdot x - iq\cdot y} + [a_{p\rightarrow}^\dagger a_{q\rightarrow}] e^{-ip\cdot x + iq\cdot y} \\ &= \frac{1}{(2\pi)^3} \frac{1}{2} - e^{ip\cdot(x-y)} - e^{ip\cdot(y-x)} \\ &= i\delta^{(3)}(\mathbf{x} \rightarrow \mathbf{y}) \end{aligned} \quad (2.21)$$

## The Hamiltonian

Let's now compute the Hamiltonian in terms of  $a_{p\rightarrow}$  and  $a_{p\rightarrow}^\dagger$ . We have

$$\begin{aligned} H &= \frac{1}{2} \int d^3 x \pi^2 + (\nabla \varphi)^2 + m^2 \varphi^2 \\ &= \frac{1}{2} \int \frac{d^3 x}{(2\pi)^6} \frac{d^3 p}{2} \frac{d^3 q}{2} \sqrt{\omega_{p\rightarrow} \omega_{q\rightarrow}} (a_{p\rightarrow} e^{ip\cdot x} - a_{p\rightarrow}^\dagger e^{-ip\cdot x})(a_{q\rightarrow} e^{iq\cdot x} - a_{q\rightarrow}^\dagger e^{-iq\cdot x}) \\ &\quad + \frac{\sqrt{1}}{2 \omega_{p\rightarrow} \omega_{q\rightarrow}} (i\mathbf{p} \cdot a_{p\rightarrow} e^{ip\cdot x} - i\mathbf{p} \cdot a_{p\rightarrow}^\dagger e^{-ip\cdot x}) \cdot (i\mathbf{q} \cdot a_{q\rightarrow} e^{iq\cdot x} - i\mathbf{q} \cdot a_{q\rightarrow}^\dagger e^{-iq\cdot x}) \\ &\quad + \frac{\sqrt{m^2}}{2 \omega_{p\rightarrow} \omega_{q\rightarrow}} (a_{p\rightarrow} e^{ip\cdot x} + a_{p\rightarrow}^\dagger e^{-ip\cdot x})(a_{q\rightarrow} e^{iq\cdot x} + a_{q\rightarrow}^\dagger e^{-iq\cdot x}) \\ &= \frac{1}{4} \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\omega_{p\rightarrow}} h(-\omega_{p\rightarrow} + \mathbf{p}^2/2 + m)(a_{p\rightarrow} a_{-p\rightarrow} + a_{p\rightarrow}^\dagger a_{-p\rightarrow}^\dagger) + (\omega_{p\rightarrow}/2 + \mathbf{p}^2/2 + m)(a_{p\rightarrow} a_{p\rightarrow}^\dagger + a_{p\rightarrow}^\dagger a_{p\rightarrow}) \end{aligned}$$

where in the second line we've used the expressions for  $\varphi$  and  $\pi$  given in (2.18) and (2.19); to get to the third line we've integrated over  $d^3 x$  to get delta-functions  $\delta^{(3)}(\mathbf{p} \pm \mathbf{q})$  which, in turn, allow us to perform the  $d^3 q$  integral. Now using the expression for the frequency  $\omega^2 = \mathbf{p}^2 + m^2$ , the first term vanishes and we're left with

$$\begin{aligned} H &= \frac{1}{2} \int \frac{d^3 p}{(2\pi)^3} \frac{h}{\omega_{p\rightarrow}} a_{p\rightarrow} a_{p\rightarrow}^\dagger + a_{p\rightarrow}^\dagger a_{p\rightarrow} \mathbf{i} \\ &= \frac{1}{(2\pi)^3} \omega_{p\rightarrow} a_{p\rightarrow}^\dagger a_{p\rightarrow} + \frac{1}{2} (2\pi)^3 \delta^{(3)}(0) \mathbf{i} \end{aligned} \quad (2.22)$$

Hmmmm. We've found a delta-function, evaluated at zero where it has its infinite spike. Moreover, the integral over  $\omega_{p\rightarrow}$  diverges at large  $p$ . What to do? Let's start by looking at the ground state where this infinity first becomes apparent.

### 2.3 The Vacuum

Following our procedure for the harmonic oscillator, let's define the vacuum  $|0\rangle$  by insisting that it is annihilated by *all*  $a_{p\rightarrow}$ ,

$$a_{p\rightarrow}|0\rangle = 0 \quad \forall p \rightarrow \quad (2.23)$$

With this definition, the energy  $E_0$  of the ground state comes from the second term in (2.22),

$$H|0\rangle \equiv E_0|0\rangle = \int d^3p \frac{1}{2} \omega_{p\rightarrow} \delta^{(3)}(0) |0\rangle = \infty |0\rangle \quad (2.24)$$

The subject of quantum field theory is rife with infinities. Each tells us something important, usually that we're doing something wrong, or asking the wrong question. Let's take some time to explore where this infinity comes from and how we should deal with it.

In fact there are two different  $\infty$ 's lurking in the expression (2.24). The first arises because space is infinitely large. (Infinities of this type are often referred to as *infra-red divergences* although in this case the  $\infty$  is so simple that it barely deserves this name). To extract out this infinity, let's consider putting the theory in a box with sides of length  $L$ . We impose periodic boundary conditions on the field. Then, taking the limit where  $L \rightarrow \infty$ , we get

$$(2\pi)^3 \delta^{(3)}(0) = \lim_{L \rightarrow \infty} \int_{-L/2}^{L/2} d^3x e^{i\vec{x}\cdot\vec{p}} = \lim_{L \rightarrow \infty} \int_{-L/2}^{L/2} d^3x = V \quad (2.25)$$

where  $V$  is the volume of the box. So the  $\delta(0)$  divergence arises because we're computing the total energy, rather than the energy density  $E_0$ . To find  $E_0$  we can simply divide by the volume,

$$E_0 = \frac{E_0}{V} = \frac{\int d^3p}{(2\pi)^3} \frac{1}{2} \omega_{p\rightarrow} \quad (2.26)$$

which is still infinite. We recognize it as the sum of ground state energies for each harmonic oscillator. But  $E_0 \rightarrow \infty$  due to the  $|p\rightarrow| \rightarrow \infty$  limit of the integral. This is a high frequency — or short distance — infinity known as an *ultra-violet divergence*. This divergence arises because of our hubris. We've assumed that our theory is valid to arbitrarily short distance scales, corresponding to arbitrarily high energies. This is clearly absurd. The integral should be cut-off at high momentum in order to reflect the fact that our theory is likely to break down in some way.

We can deal with the infinity in (2.24) in a more practical way. In physics we're only interested in energy differences. There's no way to measure  $E_0$  directly, so we can simply redefine the Hamiltonian by subtracting off this infinity,

$$H = \int \frac{d^3 p}{(2\pi)^3} \omega_{p \rightarrow} a^\dagger a \quad (2.27)$$

so that, with this new definition,  $H |0\rangle = 0$ . In fact, the difference between this Hamiltonian and the previous one is merely an ordering ambiguity in moving from the classical theory to the quantum theory. For example, if we defined the Hamiltonian of the harmonic oscillator to be  $H = (1/2)(\omega q - ip)(\omega q + ip)$ , which is classically the same as our original choice, then upon quantization it would naturally give  $H = \omega a^\dagger a$  as in (2.27). This type of ordering ambiguity arises a lot in field theories. We'll come across a number of ways of dealing with it. The method that we've used above is called *normal ordering*.

**Definition:** We write the *normal ordered* string of operators  $\varphi_1(\star_1) \dots \varphi_n(\star_n)$  as

$$:\varphi_1(\star_1) \dots \varphi_n(\star_n): \quad (2.28)$$

It is defined to be the usual product with all annihilation operators  $a_{p \rightarrow}$  placed to the right. So, for the Hamiltonian, we could write (2.27) as

$$:H := \int \frac{d^3 p}{(2\pi)^3} \omega_{p \rightarrow} a^\dagger a \quad (2.29)$$

In the remainder of this section, we will normal order all operators in this manner.

### 2.3.1 The Cosmological Constant

Above I wrote "there's no way to measure  $E_0$  directly". There is a BIG caveat here: gravity is supposed to see everything! The sum of all the zero point energies should contribute to the stress-energy tensor that appears on the right-hand side of Einstein's equations. We expect them to appear as a *cosmological constant*  $\Lambda = E_0/V$ ,

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = -8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu} \quad (2.30)$$

Current observation suggests that 70% of the energy density in the universe has the properties of a cosmological constant with  $\Lambda \sim (10^{-3} \text{ eV})^4$ . This is much smaller than other scales in particle physics. In particular, the Standard Model is valid at least up to  $10^{12} \text{ eV}$ . Why don't the zero point energies of these fields contribute to  $\Lambda$ ? Or, if they do, what cancels them to such high accuracy? This is the cosmological constant problem. No one knows the answer!

### 2.3.2 The Casimir Effect

"I mentioned my results to Niels Bohr, during a walk. That is nice, he said, that is something new... and he mumbled something about zero-point energy."

*Hendrik Casimir*

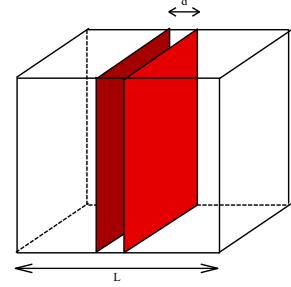
Using the normal ordering prescription we can happily set  $E_0 = 0$ , while chanting the mantra that only energy differences can be measured. But we should be careful, for there is a situation where differences in the energy of vacuum fluctuations themselves can be measured.

To regulate the infra-red divergences, we'll make the  $x^1$  direction periodic, with size  $L$ , and impose periodic boundary conditions such that

$$\varphi(\mathbf{x}) = \varphi(\mathbf{x} + L\hat{\mathbf{e}}) \quad (2.31)$$

with  $\hat{\mathbf{e}} = (1, 0, 0)$ . We'll leave  $y$  and  $z$  alone, but remember that we should compute all physical quantities per unit area  $A$ . We insert two reflecting plates, separated by a distance  $d \ll L$  in the  $x^1$  direction. The plates are such that they impose  $\varphi(x) = 0$  at the position of the plates. The presence of these plates affects the Fourier decomposition of the field and, in particular, means that the momentum of the field inside the plates is quantized as

$$\hat{p}_x = \frac{n\pi}{d}, p_y, p_z \quad n \in \mathbb{Z}^+ \quad (2.32)$$



**Figure 3:**

For a *massless* scalar field, the ground state energy between the plates is

$$\frac{E(d)}{A} = \int_{n=1}^{\infty} \frac{dp_y dp_z}{(2\pi)^2} \frac{1}{2} \sqrt{\frac{n\pi}{d}}^2 + p_y^2 + p_z^2 \quad (2.33)$$

while the energy outside the plates is  $E(L - d)$ . The total energy is therefore

$$E = E(d) + E(L - d) \quad (2.34)$$

which – at least naively – depends on  $d$ . If this naive guess is true, it would mean that there is a force on the plates due to the fluctuations of the vacuum. This is the Casimir force, first predicted in 1948 and observed 10 years later. In the real world, the effect is due to the vacuum fluctuations of the electromagnetic field, with the boundary conditions imposed by conducting plates. Here we model this effect with a scalar.

But there's a problem.  $E$  is infinite! What to do? The problem comes from the arbitrarily high momentum modes. We could regulate this in a number of different ways. Physically one could argue that any real plate cannot reflect waves of arbitrarily high frequency: at some point, things begin to leak. Mathematically, we want to find a way to neglect modes of momentum  $p \gg a^{-1}$  for some distance scale  $a \ll d$ , known as the ultra-violet (UV) cut-off. One way to do this is to change the integral (2.33) to,

$$\frac{E(d)}{A} = \sum_{n=1}^{\infty} \int \frac{dp_y dp_z}{(2\pi)^2} \frac{1}{2} \frac{n\pi^2}{d} e^{-a\sqrt{\left(\frac{n\pi^2}{d}\right)^2 + p_y^2 + p_z^2}} \quad (2.35)$$

which has the property that as  $a \rightarrow 0$ , we regain the full, infinite, expression (2.33). However (2.35) is finite, and gives us something we can easily work with. Of course, we made it finite in a rather ad-hoc manner and we better make sure that any physical quantity we calculate doesn't depend on the UV cut-off  $a$ , otherwise it's not something we can really trust.

The integral (2.35) is do-able, but a little complicated. It's a lot simpler if we look at the problem in  $d = 1 + 1$  dimensions, rather than  $d = 3 + 1$  dimensions. We'll find that all the same physics is at play. Now the energy is given by

$$E_{1+1}(d) = \frac{\pi}{2d} \sum_{n=1}^{\infty} n \quad (2.36)$$

We now regulate this sum by introducing the UV cutoff  $a$  introduced above. This renders the expression finite, allowing us to start manipulating it thus,

$$\begin{aligned} E_{1+1}(d) &\rightarrow \frac{\pi}{2d} \sum_{n=1}^{\infty} n e^{-an\pi/d} \\ &= -\frac{1}{2d} \frac{\partial}{\partial a} \sum_{n=1}^{\infty} e^{-an\pi/d} \\ &= -\frac{1}{2d} \frac{\partial}{\partial a} \frac{1}{1 - e^{-a\pi/d}} \\ &= \frac{\pi}{2d} \frac{e^{a\pi/d}}{(e^{a\pi/d} - 1)^2} \\ &= \frac{d}{2\pi a^2} - \frac{\pi}{24d} + O(a^2) \end{aligned} \quad (2.37)$$

where, in the last line, we've used the fact that  $a \ll d$ . We can now compute the full energy,

$$E_{1+1} = E_{1+1}(d) + E_{1+1}(L-d) = \frac{L}{2\pi a^2} - \frac{\pi}{24} \left( \frac{1}{d} + \frac{1}{L-d} \right) + O(a^2) \quad (2.38)$$

This is still infinite in the limit  $a \rightarrow 0$ , which is to be expected. However, the force is given by

$$\frac{\partial E_{1+1}}{\partial d} = \frac{\pi}{24d^2} + \dots \quad (2.39)$$

where the  $\dots$  include terms of size  $d/L$  and  $a/d$ . The key point is that as we remove both the regulators, and take  $a \rightarrow 0$  and  $L \rightarrow \infty$ , the force between the plates remains finite. This is the Casimir force<sup>2</sup>.

If we ploughed through the analogous calculation in  $d = 3 + 1$  dimensions, and performed the integral (2.35), we would find the result

$$\frac{1}{A} \frac{\partial E}{\partial d} = \frac{\pi^2}{480d^4} \quad (2.40)$$

The true Casimir force is twice as large as this, due to the two polarization states of the photon.

## 2.4 Particles

Having dealt with the vacuum, we can now turn to the excitations of the field. It's easy to verify that

$$[H, a_{\vec{p}}^\dagger] = \omega_{p\rightarrow} a_{\vec{p}}^\dagger \quad \text{and} \quad [H, a_{\vec{p}}] = -\omega_{p\rightarrow} a_{\vec{p}} \quad (2.41)$$

which means that, just as for the harmonic oscillator, we can construct energy eigenstates by acting on the vacuum  $|0\rangle$  with  $a_{\vec{p}}^\dagger$ . Let

$$|\vec{p}\rangle = a_{\vec{p}}^\dagger |0\rangle \quad (2.42)$$

This state has energy

$$H|\vec{p}\rangle = \omega_{p\rightarrow} |\vec{p}\rangle \quad \text{with} \quad \omega_{p\rightarrow}^2 = \vec{p}^2 + m^2 \quad (2.43)$$

But we recognize this as the relativistic dispersion relation for a particle of mass  $m$  and 3-momentum  $\vec{p}$ ,

$$E_{\vec{p}}^2 = \vec{p}^2 + m^2 \quad (2.44)$$

---

<sup>2</sup>The number 24 that appears in the denominator of the one-dimensional Casimir force plays a more famous role in string theory: the same calculation in that context is the reason the bosonic string lives in  $26 = 24 + 2$  spacetime dimensions. (The +2 comes from the fact the string itself is extended in one space and one time dimension). You will need to attend next term's "String Theory" course to see what on earth this has to do with the Casimir force.

We interpret the state  $|p\rightarrow\rangle$  as the momentum eigenstate of a single particle of mass  $m$ . To stress this, from now on we'll write  $E_{p\rightarrow}$  everywhere instead of  $\omega_{p\rightarrow}$ . Let's check this particle interpretation by studying the other quantum numbers of  $|p\rightarrow\rangle$ . We may take the classical total momentum  $\vec{P}$  given in (1.46) and turn it into an operator. After normal ordering, it becomes

$$\vec{P} = \int d^3x \pi \vec{\nabla} \varphi = \int \frac{d^3p}{(2\pi)^3} \not{p} a^\dagger_{p\rightarrow} a_{p\rightarrow} \quad (2.45)$$

Acting on our state  $|p\rightarrow\rangle$  with  $\vec{P}$ , we learn that it is indeed an eigenstate,

$$\vec{P} |p\rightarrow\rangle = p\rightarrow |p\rightarrow\rangle \quad (2.46)$$

telling us that the state  $|p\rightarrow\rangle$  has momentum  $p\rightarrow$ . Another property of  $|p\rightarrow\rangle$  that we can study is its angular momentum. Once again, we may take the classical expression for the total angular momentum of the field (1.55) and turn it into an operator,

$$J^i = \epsilon^{ijk} \int d^3x (J^0)_{jk} \quad (2.47)$$

It's not hard to show that acting on the one-particle state with zero momentum,  $J^i |p\rightarrow=0\rangle = 0$ , which we interpret as telling us that the particle carries no internal angular momentum. In other words, quantizing a scalar field gives rise to a spin 0 particle.

### Multi-Particle States, Bosonic Statistics and Fock Space

We can create multi-particle states by acting multiple times with  $a^\dagger$ 's. We interpret the state in which  $n$   $a^\dagger$ 's act on the vacuum as an  $n$ -particle state,

$$|\not{p}_1, \dots, \not{p}_n\rangle = a_{p_1}^\dagger \dots a_{p_n}^\dagger |0\rangle \quad (2.48)$$

Because all the  $a^\dagger$ 's commute among themselves, the state is symmetric under exchange of any two particles. For example,

$$|\not{p}, \not{q}\rangle = |\not{q}, \not{p}\rangle \quad (2.49)$$

This means that the particles are *bosons*.

The full Hilbert space of our theory is spanned by acting on the vacuum with all possible combinations of  $a^\dagger$ 's,

$$|0\rangle, a_{p\rightarrow}^\dagger |0\rangle, a_{p\rightarrow}^\dagger a_{q\rightarrow}^\dagger |0\rangle, a_{p\rightarrow}^\dagger a_{q\rightarrow}^\dagger a_{r\rightarrow}^\dagger |0\rangle \dots \quad (2.50)$$

This space is known as a *Fock space*. The Fock space is simply the sum of the  $n$ -particle Hilbert spaces, for all  $n \geq 0$ . There is a useful operator which counts the number of particles in a given state in the Fock space. It is called the *number operator*  $N$

$$N = \int \frac{d^3 p}{(2\pi)^3} a^\dagger a \quad (2.51)$$

and satisfies  $N | p_1, \dots, p_n \rangle = n | p_1, \dots, p_n \rangle$ . The number operator commutes with the Hamiltonian,  $[N, H] = 0$ , ensuring that particle number is conserved. This means that we can place ourselves in the  $n$ -particle sector, and stay there. This is a property of free theories, but will no longer be true when we consider interactions: interactions create and destroy particles, taking us between the different sectors in the Fock space.

### Operator Valued Distributions

Although we're referring to the states  $| p \rangle$  as "particles", they're not localized in space in any way — they are momentum eigenstates. Recall that in quantum mechanics the position and momentum eigenstates are not good elements of the Hilbert space since they are not normalizable (they normalize to delta-functions). Similarly, in quantum field theory neither the operators  $\varphi(x)$ , nor  $a_{p \rightarrow}$  are good operators acting on the Fock space. This is because they don't produce normalizable states. For example,

$$\langle 0 | a_{p \rightarrow} a_{p \rightarrow} | 0 \rangle = \langle p | p \rangle = (2\pi)^3 \delta(0) \quad \text{and} \quad \langle 0 | \varphi(x) \varphi(x) | 0 \rangle = \langle x | x \rangle = \delta(0) \quad (2.52)$$

They are operator valued distributions, rather than functions. This means that although  $\varphi(x)$  has a well defined vacuum expectation value,  $\langle 0 | \varphi(x) | 0 \rangle = 0$ , the fluctuations of the operator at a fixed point are infinite,  $\langle 0 | \varphi(x) \varphi(x) | 0 \rangle = \infty$ . We can construct well defined operators by smearing these distributions over space. For example, we can create a wavepacket

$$|\phi\rangle = \int \frac{d^3 p}{(2\pi)^3} e^{-ip \cdot x} \phi(p) |p\rangle \quad (2.53)$$

which is partially localized in both position and momentum space. (A typical state might be described by the Gaussian  $\phi(p) = \exp(-p^2/2m^2)$ ).

#### 2.4.1 Relativistic Normalization

We have defined the vacuum  $|0\rangle$  which we normalize as  $\langle 0 | 0 \rangle = 1$ . The one-particle states  $|p \rangle = a_{p \rightarrow}^\dagger |0\rangle$  then satisfy

$$\langle p | q \rangle = (2\pi)^3 \delta^{(3)}(p - q) \quad (2.54)$$

But is this Lorentz invariant? It's not obvious because we only have 3-vectors. What could go wrong? Suppose we have a Lorentz transformation

$$p^\mu \rightarrow (p^r)^\mu = \Lambda^\mu_\nu p^\nu \quad (2.55)$$

such that the 3-vector transforms as  $\vec{p} \rightarrow \vec{p}'$ . In the quantum theory, it would be preferable if the two states are related by a unitary transformation,

$$|\vec{p}\rangle \rightarrow |\vec{p}'\rangle = U(\Lambda)|\vec{p}\rangle \quad (2.56)$$

This would mean that the normalizations of  $|\vec{p}\rangle$  and  $|\vec{p}'\rangle$  are the same whenever  $\vec{p}$  and  $\vec{p}'$  are related by a Lorentz transformation. But we haven't been at all careful with normalizations. In general, we could get

$$|\vec{p}\rangle \rightarrow \lambda(\vec{p}, \vec{p}')|\vec{p}'\rangle \quad (2.57)$$

for some unknown function  $\lambda(\vec{p}, \vec{p}')$ . How do we figure this out? The trick is to look at an object which we know is Lorentz invariant. One such object is the identity operator on one-particle states (which is really the projection operator onto one-particle states). With the normalization (2.54) we know this is given by

$$1 = \frac{\int d^3p}{(2\pi)^3} |\vec{p}\rangle \langle \vec{p}| \quad (2.58)$$

This operator is Lorentz invariant, but it consists of two terms: the measure  $\int d^3p$  and the projector  $|\vec{p}\rangle \langle \vec{p}|$ . Are these individually Lorentz invariant? In fact the answer is no.

**Claim** The Lorentz invariant measure is,

$$\frac{\int d^3p}{2E_{\vec{p}}} \quad (2.59)$$

**Proof:**  $d^4p$  is obviously Lorentz invariant. And the relativistic dispersion relation for a massive particle,

$$p_\mu p^\mu = m^2 \Rightarrow p_0^2 = E_{\vec{p}}^2 = \vec{p}^2 + m^2 \quad (2.60)$$

is also Lorentz invariant. Solving for  $p_0$ , there are two branches of solutions:  $p_0 = \pm E_{\vec{p}}$ . But the choice of branch is another Lorentz invariant concept. So piecing everything together, the following combination must be Lorentz invariant,

$$\int \frac{d^4p}{(2\pi)^4} \delta(p_0^2 - \vec{p}^2 - m^2)_{p_0>0} = \int \frac{d^3p}{2p_0} \Big|_{p_0=E_{\vec{p}}} \quad (2.61)$$

which completes the proof.

From this result we can figure out everything else. For example, the Lorentz invariant  $\delta$ -function for 3-vectors is

$$2E_{\vec{p}} \delta^{(3)}(\vec{p} - \vec{q}) \quad (2.62)$$

which follows because

$$\int \frac{d^3p}{2E_{\vec{p}}} 2E_{\vec{p}} \delta^{(3)}(\vec{p} - \vec{q}) = 1 \quad (2.63)$$

So finally we learn that the relativistically normalized momentum states are given by

$$|p\rangle = \sqrt{\frac{1}{2E_{\vec{p}}}} |\vec{p}\rangle = \sqrt{\frac{1}{2E_{\vec{p}}}} a_{\vec{p}}^\dagger |0\rangle \quad (2.64)$$

Notice that our notation is rather subtle: the relativistically normalized momentum state  $|p\rangle$  differs from  $|\vec{p}\rangle$  by the factor  $\sqrt{\frac{1}{2E_{\vec{p}}}}$ . These states now satisfy

$$\langle p | q \rangle = (2\pi)^3 2E_{\vec{p}} \delta^{(3)}(\vec{p} - \vec{q}) \quad (2.65)$$

Finally, we can rewrite the identity on one-particle states as

$$1 = \frac{d^3p}{(2\pi)^3 2E_{\vec{p}}} |p\rangle \langle p| \quad (2.66)$$

Some texts also define relativistically normalized creation operators by  $a^\dagger(p) = \sqrt{\frac{1}{2E_{\vec{p}}}} \vec{a}_{\vec{p}}$ . We won't make use of this notation here.

## 2.5 Complex Scalar Fields

Consider a complex scalar field  $\psi(x)$  with Lagrangian

$$L = \partial_\mu \psi^\dagger \partial^\mu \psi - M^2 \psi^\dagger \psi \quad (2.67)$$

Notice that, in contrast to the Lagrangian (1.7) for a real scalar field, there is no factor of  $1/2$  in front of the Lagrangian for a complex scalar field. If we write  $\psi$  in terms of real scalar fields by  $\psi = (\varphi_1 + i\varphi_2)/\sqrt{2}$ , we get the factor of  $1/2$  coming from the  $\sqrt{1/2}$ 's. The equations of motion are

$$\begin{aligned} \partial_\mu \partial^\mu \psi + M^2 \psi &= 0 \\ \partial_\mu \partial^\mu \psi^\dagger + M^2 \psi^\dagger &= 0 \end{aligned} \quad (2.68)$$

where the second equation is the complex conjugate of the first. We expand the complex field operator as a sum of plane waves as

$$\begin{aligned} \psi &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} b_{\vec{p}} e^{+ip \cdot \vec{x}} + c_{\vec{p}}^\dagger e^{-ip \cdot \vec{x}} \\ \psi^\dagger &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} b_{\vec{p}}^\dagger e^{-ip \cdot \vec{x}} + c_{\vec{p}} e^{+ip \cdot \vec{x}} \end{aligned} \quad (2.69)$$

Since the classical field  $\psi$  is not real, the corresponding quantum field  $\psi$  is not hermitian. This is the reason that we have different operators  $b$  and  $c^\dagger$  appearing in the positive and negative frequency parts. The classical field momentum is  $\pi = \partial L / \partial \dot{\psi} = \dot{\psi}^\wedge$ . We also turn this into a quantum operator field which we write as,

$$\begin{aligned}\pi &= \int \frac{d^3 p}{(2\pi)^3} i \frac{\mathbf{E}_{p \rightarrow}}{2} b_{p \rightarrow} e^{-ip \cdot x} - c_{p \rightarrow} e^{+ip \cdot x} \\ \pi^\dagger &= \int \frac{d^3 p}{(2\pi)^3} \frac{i}{2} \frac{\mathbf{E}_{p \rightarrow}}{p \rightarrow} b_{p \rightarrow} e^{+ip \cdot x} - c_{p \rightarrow}^\dagger e^{-ip \cdot x}\end{aligned}\quad (2.70)$$

The commutation relations between fields and momenta are given by

$$[\psi(\mathbf{x}), \pi(\mathbf{y})] = i\delta^{(3)}(\mathbf{x} \rightarrow \mathbf{y}) \text{ and } [\psi(\mathbf{x}), \pi^\dagger(\mathbf{y})] = 0 \quad (2.71)$$

together with others related by complex conjugation, as well as the usual  $[\psi(\mathbf{x}), \psi(\mathbf{y})] = [\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] = 0$ , etc. One can easily check that these field commutation relations are equivalent to the commutation relations for the operators  $b_{p \rightarrow}$  and  $c_{p \rightarrow}$ ,

$$\begin{aligned}[b_{p \rightarrow}, b_{q \rightarrow}^\dagger] &= (2\pi)^3 \delta^{(3)}(\mathbf{p} \rightarrow \mathbf{q}) \\ [c_{p \rightarrow}, c_{q \rightarrow}^\dagger] &= (2\pi)^3 \delta^{(3)}(\mathbf{p} \rightarrow \mathbf{q})\end{aligned}\quad (2.72)$$

and

$$[b_{p \rightarrow}, b_{q \rightarrow}] = [c_{p \rightarrow}, c_{q \rightarrow}] = [b_{p \rightarrow}, c_{q \rightarrow}] = [b_{p \rightarrow}, c_{q \rightarrow}^\dagger] = 0 \quad (2.73)$$

In summary, quantizing a complex scalar field gives rise to two creation operators,  $b_{p \rightarrow}^\dagger$  and  $c_{p \rightarrow}^\dagger$ . These have the interpretation of creating two types of particle, both of mass  $M$  and both spin zero. They are interpreted as particles and anti-particles. In contrast, for a real scalar field there is only a single type of particle: for a real scalar field, the particle is its own antiparticle.

Recall that the theory (2.67) has a classical conserved charge

$$Q = i \int d^3 x (\psi^\wedge \psi - \psi^\wedge \psi^\wedge) = i \int d^3 x (\pi \psi - \psi^\wedge \pi^\wedge) \quad (2.74)$$

After normal ordering, this becomes the quantum operator

$$Q = \int \frac{d^3 p}{(2\pi)} (c_{p \rightarrow}^\dagger c_{p \rightarrow} - b_{p \rightarrow}^\dagger b_{p \rightarrow}) = N_c - N_b \quad (2.75)$$

so  $Q$  counts the number of anti-particles (created by  $c^\dagger$ ) minus the number of particles (created by  $b^\dagger$ ). We have  $[H, Q] = 0$ , ensuring that  $Q$  is conserved quantity in the quantum theory. Of course, in our free field theory this isn't such a big deal because both  $N_c$  and  $N_b$  are separately conserved. However, we'll soon see that in interacting theories  $Q$  survives as a conserved quantity, while  $N_c$  and  $N_b$  individually do not.

## 2.6 The Heisenberg Picture

Although we started with a Lorentz invariant Lagrangian, we slowly butchered it as we quantized, introducing a preferred time coordinate  $t$ . It's not at all obvious that the theory is still Lorentz invariant after quantization. For example, the operators  $\varphi(\mathbf{x})$  depend on space, but not on time. Meanwhile, the one-particle states evolve in time by Schrödinger's equation,

$$i \frac{d|\psi(t)\rangle}{dt} = H |\psi(t)\rangle \Rightarrow |\psi(t)\rangle = e^{-iE_p t} |\psi\rangle \quad (2.76)$$

Things start to look better in the Heisenberg picture where time dependence is assigned to the operators  $O$ ,

$$O_H = e^{iHt} O_S e^{-iHt} \quad (2.77)$$

so that

$$\frac{dO_H}{dt} = i[H, O_H] \quad (2.78)$$

where the subscripts  $S$  and  $H$  tell us whether the operator is in the Schrödinger or Heisenberg picture. In field theory, we drop these subscripts and we will denote the picture by specifying whether the fields depend on space  $\varphi(\mathbf{x})$  (the Schrödinger picture) or spacetime  $\varphi(\mathbf{x}, t) = \varphi(x)$  (the Heisenberg picture).

The operators in the two pictures agree at a fixed time, say,  $t = 0$ . The commutation relations (2.2) become equal time commutation relations in the Heisenberg picture,

$$\begin{aligned} [\varphi(\mathbf{x}, t), \varphi(\mathbf{y}, t)] &= [\pi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = 0 \\ [\varphi(\mathbf{x}, t), \pi(\mathbf{y}, t)] &= i\delta^{(3)}(\mathbf{x} - \mathbf{y}) \end{aligned} \quad (2.79)$$

Now that the operator  $\varphi(x) = \varphi(\mathbf{x}, t)$  depends on time, we can start to study how it evolves. For example, we have

$$\begin{aligned} \dot{\varphi} &= i[H, \varphi] = \frac{i}{2} \int d^3y \left[ \pi(y)^2 + \nabla\varphi(y)^2 + m^2\varphi(y)^2, \varphi(x) \right] \\ &= i \int d^3y \pi(y) (-i) \delta^{(3)}(\mathbf{y} - \mathbf{x}) = \pi(x) \end{aligned} \quad (2.80)$$

Meanwhile, the equation of motion for  $\pi$  reads,

$$\dot{\pi} = i[H, \pi] = \frac{i}{2} \int d^3y \left[ \pi(y)^2 + \nabla\varphi(y)^2 + m^2\varphi(y)^2, \pi(x) \right]$$

$$\begin{aligned}
&= \frac{i}{2} \int d^3y (\nabla_y [\varphi(y), \pi(x)]) \nabla \varphi(y) + \nabla \varphi(y) \nabla_y [\varphi(y), \pi(x)] \\
&\quad + 2im^2 \varphi(y) \delta^{(3)}(\mathbf{x} - \mathbf{y}) \\
&= - \int d^3y \nabla_y \delta^{(3)}(\mathbf{x} - \mathbf{y}) \nabla_y \varphi(y) - m^2 \varphi(x) \\
&= \nabla^2 \varphi - m^2 \varphi
\end{aligned} \tag{2.81}$$

where we've included the subscript  $y$  on  $\nabla_y$  when there may be some confusion about which argument the derivative is acting on. To reach the last line, we've simply integrated by parts. Putting (2.80) and (2.81) together we find that the field operator  $\varphi$  satisfies the Klein-Gordon equation

$$\partial_\mu \partial^\mu \varphi + m^2 \varphi = 0 \tag{2.82}$$

Things are beginning to look more relativistic. We can write the Fourier expansion of  $\varphi(x)$  by using the definition (2.77) and noting,

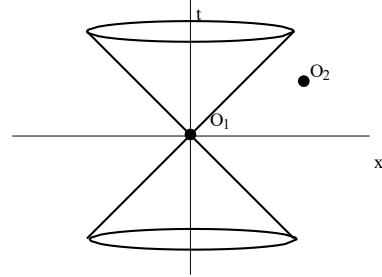
$$e^{iHt} a_{\vec{p}} e^{-iHt} = e^{-iE_{\vec{p}} t} a_{\vec{p}} \text{ and } e^{iHt} a_{\vec{p}}^\dagger e^{-iHt} = e^{+iE_{\vec{p}} t} a_{\vec{p}}^\dagger \tag{2.83}$$

which follows from the commutation relations  $[H, a_{\vec{p}}] = -E_{\vec{p}} a_{\vec{p}}$  and  $[H, a_{\vec{p}}^\dagger] = +E_{\vec{p}} a_{\vec{p}}^\dagger$ .

This then gives,

$$\varphi(\mathbf{x}, t) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{\frac{2E}{p^2}}} a_{\vec{p}} e^{-ip \cdot x} + a_{\vec{p}}^\dagger e^{+ip \cdot x} \tag{2.84}$$

which looks very similar to the previous expansion (2.18) except that the exponent is now written in terms of 4-vectors,  $p \cdot x = E_{\vec{p}} t - \vec{p} \cdot \vec{x}$ . (Note also that a sign has flipped in the exponent due to our Minkowski metric contraction). It's simple to check that (2.84) indeed satisfies the Klein-Gordon equation (2.82).



### 2.6.1 Causality

We're approaching something Lorentz invariant in the Heisenberg picture, where  $\varphi(x)$  now satisfies the Klein-Gordon equation. But there's still a hint of non-Lorentz invariance because  $\varphi$  and  $\pi$  satisfy *equal time* commutation relations,

$$[\varphi(\mathbf{x}, t), \pi(\mathbf{y}, t)] = i\delta^{(3)}(\mathbf{x} - \mathbf{y}) \tag{2.85}$$

**Figure 4:**

But what about arbitrary spacetime separations? In particular, for our theory to be *causal*, we must require that all spacelike separated operators commute,

$$[O_1(x), O_2(y)] = 0 \quad \forall (x - y)^2 < 0 \quad (2.86)$$

This ensures that a measurement at  $x$  cannot affect a measurement at  $y$  when  $x$  and  $y$  are not causally connected. Does our theory satisfy this crucial property? Let's define

$$\Delta(x - y) = [\varphi(x), \varphi(y)] \quad (2.87)$$

The objects on the right-hand side of this expression are operators. However, it's easy to check by direct substitution that the left-hand side is simply a c-number function with the integral expression

$$\Delta(x - y) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_p} e^{-ip \cdot (x-y)} - e^{ip \cdot (x-y)} \quad (2.88)$$

What do we know about this function?

- It's Lorentz invariant, thanks to the appearance of the Lorentz invariant measure  $d^3 p / 2E_p$  that we introduced in (2.59).
- It doesn't vanish for timelike separation. For example, taking  $x - y = (t, 0, 0, 0)$  gives  $[\varphi(\mathbf{x}, 0), \varphi(\mathbf{x}, t)] \sim e^{-imt} - e^{+imt}$ .
- It vanishes for space-like separations. This follows by noting that  $\Delta(x - y) = 0$  at equal times for all  $(x - y)^2 = -(x^0 - y^0)^2 < 0$ , which we can see explicitly by writing

$$[\varphi(\mathbf{x}, t), \varphi(\mathbf{y}, t)] = \frac{1}{(2\pi)^3} \int \frac{d^3 p}{p^2 + m^2} e^{ip \cdot (\mathbf{x}-\mathbf{y})} - e^{-ip \cdot (\mathbf{x}-\mathbf{y})} \quad (2.89)$$

and noticing that we can flip the sign of  $p^0$  in the last exponent as it is an integration variable. But since  $\Delta(x - y)$  is Lorentz invariant, it can only depend on  $(x - y)^2$  and must therefore vanish for all  $(x - y)^2 < 0$ .

We therefore learn that our theory is indeed causal with commutators vanishing outside the lightcone. This property will continue to hold in interacting theories; indeed, it is usually given as one of the axioms of local quantum field theories. I should mention however that the fact that  $[\varphi(x), \varphi(y)]$  is a c-number function, rather than an operator, is a property of free fields only.

## 2.7 Propagators

We could ask a different question to probe the causal structure of the theory. Prepare a particle at spacetime point  $y$ . What is the amplitude to find it at point  $x$ ? We can calculate this:

$$\begin{aligned} \langle 0 | \varphi(x) \varphi(y) | 0 \rangle &= \int \frac{d^3 p d^3 p'}{(2\pi)^6} \sqrt{\frac{1}{4E_p E_{p'}}} \langle 0 | a_{p \rightarrow} a_{p' \rightarrow}^\dagger | 0 \rangle e^{-ip \cdot x + ip' \cdot y} \\ &= \frac{1}{(2\pi)^3 2E_p} e^{-ip \cdot (x-y)} \equiv D(x-y) \end{aligned} \quad (2.90)$$

The function  $D(x-y)$  is called the *propagator*. For spacelike separations,  $(x-y)^2 < 0$ , one can show that  $D(x-y)$  decays like

$$D(x-y) \sim e^{-m|x-y|} \quad (2.91)$$

So it decays exponentially quickly outside the lightcone but, nonetheless, is non-vanishing! The quantum field appears to leak out of the lightcone. Yet we've just seen that spacelike measurements commute and the theory is causal. How do we reconcile these two facts? We can rewrite the calculation (2.89) as

$$[\varphi(x), \varphi(y)] = D(x-y) - D(y-x) = 0 \text{ if } (x-y)^2 < 0 \quad (2.92)$$

There are words you can drape around this calculation. When  $(x-y)^2 < 0$ , there is no Lorentz invariant way to order events. If a particle can travel in a spacelike direction from  $x \rightarrow y$ , it can just as easily travel from  $y \rightarrow x$ . In any measurement, the amplitudes for these two events cancel.

With a complex scalar field, it is more interesting. We can look at the equation  $[\psi(x), \psi^\dagger(y)] = 0$  outside the lightcone. The interpretation now is that the amplitude for the particle to propagate from  $x \rightarrow y$  cancels the amplitude for the *antiparticle* to travel from  $y \rightarrow x$ . In fact, this interpretation is also there for a real scalar field because the particle is its own antiparticle.

### 2.7.1 The Feynman Propagator

As we will see shortly, one of the most important quantities in interacting field theory is the *Feynman propagator*,

$$\Delta_F(x-y) = \langle 0 | T\varphi(x)\varphi(y) | 0 \rangle = \begin{cases} D(x-y) & x^0 > y^0 \\ D(y-x) & y^0 > x^0 \end{cases} \quad (2.93)$$

where  $T$  stands for time ordering, placing all operators evaluated at later times to the left so,

$$T\varphi(x)\varphi(y) = \begin{cases} \varphi(x)\varphi(y) & x^0 > y^0 \\ \varphi(y)\varphi(x) & y^0 > x^0 \end{cases} \quad (2.94)$$

**Claim:** There is a useful way of writing the Feynman propagator in terms of a 4-momentum integral that shows that it is explicitly Lorentz invariant

$$\Delta_F(x - y) = \frac{d^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2} e^{-ip \cdot (x-y)} \quad (2.95)$$

Notice that this is the first time in this course that we've integrated over 4-momentum. Until now, we integrated only over 3-momentum, with  $p^0$  fixed by the mass-shell condition to be  $p^0 = E_{p\rightarrow}$ . In the expression (2.95) for  $\Delta_F$ , we have no such condition on  $p^0$ . However, as it stands this integral is ill-defined because, for each value of  $p\rightarrow$ , the denominator  $p^2 - m^2 = (p^0)^2 - \vec{p}^2 - m^2$  produces a pole when  $p^0 = \pm E_{p\rightarrow} = \pm \sqrt{\vec{p}^2 + m^2}$ . We need a prescription for avoiding these singularities in the  $p_0$  integral. To get the Feynman propagator, we must choose the contour to be

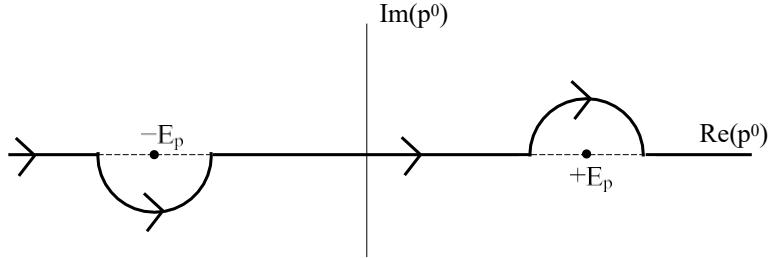


Figure 5: The contour for the Feynman propagator.

**Proof:**

$$\frac{1}{p^2 - m^2} = \frac{1}{(p^0)^2 - E_{p\rightarrow}^2} = \frac{1}{(p^0 - E_{p\rightarrow})(p^0 + E_{p\rightarrow})} \quad (2.96)$$

so the residue of the pole at  $p^0 = \pm E_{p\rightarrow}$  is  $\pm 1/2E_{p\rightarrow}$ . When  $x^0 > y^0$ , we close the contour in the lower half plane, where  $p^0 \rightarrow -i\infty$ , ensuring that the integrand vanishes since  $e^{-ip^0(x^0-y^0)} \rightarrow 0$ . The integral over  $p^0$  then picks up the residue at  $p^0 = +E_{p\rightarrow}$  which is  $-2\pi i/2E_{p\rightarrow}$  where the minus sign arises because we took a clockwise contour. Hence when  $x^0 > y^0$  we have

$$\Delta_F(x - y) = \int \frac{d^3 p}{(2\pi)^4} \frac{-2\pi i}{2E_{p\rightarrow}} i e^{-iE_{p\rightarrow}(x^0-y^0)+ip\cdot(x-y)}$$

$$= \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{p\rightarrow}} e^{-ip \cdot (x-y)} = D(x-y) \quad (2.97)$$

which is indeed the Feynman propagator for  $x^0 > y^0$ . In contrast, when  $y^0 > x^0$ , we close the contour in an anti-clockwise direction in the upper half plane to get,

$$\begin{aligned} \Delta_F(x-y) &= \int \frac{d^3 p}{(2\pi)^3} \frac{2\pi i}{-2E_{p\rightarrow}} i e^{+iE_{p\rightarrow}(x^0-y^0)+ip\cdot(\rightarrow x \rightarrow y)} \\ &= \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{p\rightarrow}} e^{-iE_{p\rightarrow}(y^0-x^0)-ip\cdot(\rightarrow y \rightarrow x)} \\ &= \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{p\rightarrow}} e^{-ip\cdot(y-x)} = D(y-x) \end{aligned} \quad (2.98)$$

where to go to from the second line to the third, we have flipped the sign of  $p\rightarrow$  which is valid since we integrate over  $d^3 p$  and all other quantities depend only on  $p\rightarrow^2$ . Once again we reproduce the Feynman propagator.

Instead of specifying the contour, it is standard to write the Feynman propagator as

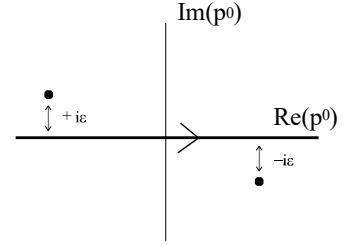
$$\Delta_F(x-y) = \int \frac{d^4 p}{(2\pi)^4} \frac{ie^{-ip\cdot(x-y)}}{p^2 - m^2 + i\epsilon} \quad (2.99)$$

with  $\epsilon > 0$ , and infinitesimal. This has the effect of shifting the poles slightly off the real axis, so the integral along the real  $p^0$  axis is equivalent to the contour shown in Figure 5. This way of writing the propagator is, for obvious reasons, called the “ $i\epsilon$  prescription”.

### 2.7.2 Green's Functions

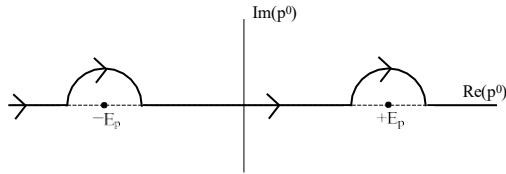
There is another avatar of the propagator: it is a Green's function for the Klein-Gordon operator. If we stay away from the singularities, we have

$$\begin{aligned} (\partial_t^2 - \nabla^2 + m^2) \Delta_F(x-y) &= \int \frac{d^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2} (-p^2 + m^2) e^{-ip\cdot(x-y)} \\ &= -i \int \frac{d^4 p}{(2\pi)^4} e^{-ip\cdot(x-y)} \\ &= -i\delta^{(4)}(x-y) \end{aligned} \quad (2.100)$$

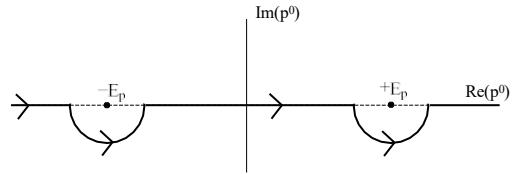


**Figure 6:**

Note that we didn't make use of the contour anywhere in this derivation. For some purposes it is also useful to pick other contours which also give rise to Green's functions.



**Figure 7:** The retarded contour



**Figure 8:** The advanced contour

For example, the *retarded* Green's function  $\Delta_R(x - y)$  is defined by the contour shown in Figure 7 which has the property

$$\Delta_R(x - y) = \begin{cases} D(x - y) - D(y - x) & x^0 > y^0 \\ 0 & y^0 > x^0 \end{cases} \quad (2.101)$$

The retarded Green's function is useful in classical field theory if we know the initial value of some field configuration and want to figure out what it evolves into in the presence of a source, meaning that we want to know the solution to the inhomogeneous Klein-Gordon equation,

$$\partial_\mu \partial^\mu \varphi + m^2 \varphi = J(x) \quad (2.102)$$

for some fixed background function  $J(x)$ . Similarly, one can define the *advanced* Green's function  $\Delta_A(x - y)$  which vanishes when  $y^0 < x^0$ , which is useful if we know the end point of a field configuration and want to figure out where it came from. Given that next term's course is called "Advanced Quantum Field Theory", there is an obvious name for the current course. But it got shot down in the staff meeting. In the quantum theory, we will see that the Feynman Green's function is most relevant.

## 2.8 Non-Relativistic Fields

Let's return to our classical complex scalar field obeying the Klein-Gordon equation. We'll decompose the field as

$$\psi(\mathbf{x}, t) = e^{-i\omega t} \tilde{\psi}(\mathbf{x}, t) \quad (2.103)$$

Then the KG-equation reads

$$\partial_t^2 \psi - \nabla^2 \psi + m^2 \psi = e^{-i\omega t} (\tilde{\psi}'' - 2im\tilde{\psi}' - \nabla^2 \tilde{\psi}) = 0 \quad (2.104)$$

with the  $m^2$  term cancelled by the time derivatives. The non-relativistic limit of a particle is  $|p| \ll m$ . Let's look at what this does to our field. After a Fourier transform,

this is equivalent to saying that  $|\tilde{\psi}| \sim m|\dot{\tilde{\psi}}|$ . In this limit, we drop the term with two time derivatives and the KG equation becomes,

$$i\frac{\partial \tilde{\psi}}{\partial t} = \frac{1}{2m}\nabla^2 \tilde{\psi} \quad (2.105)$$

This looks very similar to the Schrödinger equation for a non-relativistic free particle of mass  $m$ . Except it doesn't have any probability interpretation — it's simply a classical field evolving through an equation that's first order in time derivatives.

We wrote down a Lagrangian in section 1.1.2 which gives rise to field equations which are first order in time derivatives. In fact, we can derive this from the relativistic Lagrangian for a scalar field by again taking the limit  $\partial_t \tilde{\psi} \rightarrow m\dot{\psi}$ . After losing the tilde, so  $\tilde{\psi} \rightarrow \psi$ , the non-relativistic Lagrangian becomes

$$L = +i\psi \dot{\psi} - \frac{1}{2m}\nabla^\lambda \psi^\lambda \nabla_\lambda \psi \quad (2.106)$$

where we've divided by  $1/2m$ . This Lagrangian has a conserved current arising from the internal symmetry  $\psi \rightarrow e^{i\alpha} \psi$ . The current has time and space components

$$j^\mu = -\psi^\lambda \psi_\lambda - \frac{i}{2m}(\psi^\lambda \nabla_\lambda \psi - \psi \nabla_\lambda \psi^\lambda) \quad (2.107)$$

To move to the Hamiltonian formalism we compute the momentum

$$\pi = \frac{\partial L}{\partial \dot{\psi}} = i\psi^\lambda \quad (2.108)$$

This means that the momentum conjugate to  $\psi$  is  $i\psi^\lambda$ . The momentum does not depend on time derivatives at all! This looks a little disconcerting but it's fully consistent for a theory which is first order in time derivatives. In order to determine the full trajectory of the field, we need only specify  $\psi$  and  $\psi^\lambda$  at time  $t = 0$ : no time derivatives on the initial slice are required.

Since the Lagrangian already contains a “ $pq$ ” term (instead of the more familiar  $\frac{1}{2}pq$  term), the time derivatives drop out when we compute the Hamiltonian. We get,

$$H = \frac{1}{2m}\nabla_\lambda \psi^\lambda \nabla^\lambda \psi \quad (2.109)$$

To quantize we impose (in the Schrödinger picture) the canonical commutation relations

$$\begin{aligned} [\psi(\mathbf{x}), \psi(\mathbf{y})] &= [\psi^\dagger(\mathbf{x}), \psi^\dagger(\mathbf{y})] = 0 \\ [\psi(\mathbf{x}), \psi^\dagger(\mathbf{y})] &= \delta^{(3)}(\mathbf{x} - \mathbf{y}) \end{aligned} \quad (2.110)$$

We may expand  $\psi(\mathbf{x})$  as a Fourier transform

$$\psi(\mathbf{x}) = \int \frac{d^3 p}{(2\pi)^3} a_{p \rightarrow} e^{ip \cdot x} \quad (2.111)$$

where the commutation relations (2.110) require

$$[a_{p \rightarrow}, a_{q \rightarrow}^\dagger] = (2\pi)^3 \delta^{(3)}(p - q) \quad (2.112)$$

The vacuum satisfies  $a_{p \rightarrow} |0\rangle = 0$ , and the excitations are  $a_{p \rightarrow_1}^\dagger \dots a_{p \rightarrow_n}^\dagger |0\rangle$ . The one-particle states have energy

$$H|p \rightarrow\rangle = \frac{\vec{p}^2}{2m} |p \rightarrow\rangle \quad (2.113)$$

which is the non-relativistic dispersion relation. We conclude that quantizing the first order Lagrangian (2.106) gives rise to non-relativistic particles of mass  $m$ . Some comments:

- We have a complex field but only a single type of particle. The anti-particle is not in the spectrum. The existence of anti-particles is a consequence of relativity.
- A related fact is that the conserved charge  $Q = \int d^3x : \psi^\dagger \psi :$  is the particle number. This remains conserved even if we include interactions in the Lagrangian of the form

$$\Delta L = V(\psi^\dagger \psi) \quad (2.114)$$

So in non-relativistic theories, particle number is conserved. It is only with relativity, and the appearance of anti-particles, that particle number can change.

- There is no non-relativistic limit of a real scalar field. In the relativistic theory, the particles are their own anti-particles, and there can be no way to construct a multi-particle theory that conserves particle number.

### 2.8.1 Recovering Quantum Mechanics

In quantum mechanics, we talk about the position and momentum operators  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{P}}$ . In quantum field theory, position is relegated to a label. How do we get back to quantum mechanics? We already have the operator for the total momentum of the field

$$\hat{\mathbf{P}} = \int \frac{d^3 p}{(2\pi)^3} \hat{a}_{p \rightarrow}^\dagger a_{p \rightarrow} \quad (2.115)$$

which, on one-particle states, gives  $\hat{P} |p\rangle = p |p\rangle$ . It's also easy to construct the position operator. Let's work in the non-relativistic limit. Then the operator

$$\psi^\dagger(\mathbf{x}) = \frac{d^3 p}{(2\pi)^3} a_{p\rightarrow}^\dagger e^{-ip\cdot x} \quad (2.116)$$

creates a particle with  $\delta$ -function localization at  $\rightarrow x$ . We write  $|\mathbf{x}\rangle = \psi^\dagger(\mathbf{x}) |0\rangle$ . A natural position operator is then

$$\hat{\mathbf{x}} = \int d^3x \mathbf{x} \psi^\dagger(\mathbf{x}) \psi(\mathbf{x}) \quad (2.117)$$

so that  $\hat{\mathbf{x}} |\rightarrow x\rangle = \mathbf{x} |\mathbf{x}\rangle$ .

Let's now construct a state  $|\phi\rangle$  by taking superpositions of one-particle states  $|\mathbf{x}\rangle$ ,

$$|\phi\rangle = \int d^3x \phi(\mathbf{x}) |\mathbf{x}\rangle \quad (2.118)$$

The function  $\phi(\mathbf{x})$  is what we would usually call the Schrödinger wavefunction (in the position representation). Let's make sure that it indeed satisfies all the right properties.

Firstly, it's clear that acting with the position operator  $\hat{\mathbf{x}}$  has the right action of  $\phi(\mathbf{x})$ ,

$$\mathbf{x}^i |\phi\rangle = \int d^3x x^i \phi(\mathbf{x}) |\rightarrow x\rangle \quad (2.119)$$

but what about the momentum operator  $\hat{P}$ ? We will now show that

$$P^i |\phi\rangle = \int d^3x -i \frac{\partial \phi}{\partial x^i} |\rightarrow x\rangle \quad (2.120)$$

which tells us that  $P^i$  acts as the familiar derivative on wavefunctions  $|\phi\rangle$ . To see that this is the case, we write

$$\begin{aligned} P^i |\phi\rangle &= \int \frac{d^3x d^3p}{(2\pi)^3} p^i a_{p\rightarrow}^\dagger a_{p\rightarrow} \phi(\mathbf{x}) \psi^\dagger(\mathbf{x}) |0\rangle \\ &= \int \frac{d^3x d^3p}{(2\pi)^3} p^i a^\dagger p \frac{\partial}{\partial x^i} e^{-ip\cdot x} \phi(\mathbf{x}) |0\rangle \end{aligned} \quad (2.121)$$

where we've used the relationship  $[a_{p\rightarrow}, \psi^\dagger(\mathbf{x})] = e^{-ip\cdot x}$  which can be easily checked.

Proceeding with our calculation, we have

$$\begin{aligned} P^i |\phi\rangle &= \int \frac{d^3x d^3p}{(2\pi)^3} p^i a^\dagger \frac{i}{\partial x^i} \frac{\partial}{\partial x^i} e^{-ip\cdot x} \phi(\mathbf{x}) |0\rangle \\ &= \int \frac{d^3x d^3p}{(2\pi)^3} e^{-ip\cdot x} -i \frac{\partial}{\partial x^i} a^\dagger \frac{\partial \phi}{\partial x^i} |0\rangle \\ &= \int d^3x -i \frac{\partial \phi}{\partial x^i} |\mathbf{x}\rangle \end{aligned} \quad (2.122)$$

which confirms (2.120). So we learn that when acting on one-particle states, the operators  $\hat{X}$  and  $\hat{P}$  act as position and momentum operators in quantum mechanics, with  $[X^i, P^j] |\phi\rangle = i\delta^{ij} |\phi\rangle$ . But what about dynamics? How does the wavefunction  $\phi(\hat{x}, t)$  change in time? The Hamiltonian (2.109) can be rewritten as

$$H = \int_{\mathbb{R}^3} \frac{1}{2m} \nabla \psi^\dagger \nabla \psi = \int \frac{d^3p}{(2\pi)^3} \frac{\hat{p}^2 a^\dagger a}{2m} \quad (2.123)$$

so we find that

$$i \frac{\partial \phi}{\partial t} = - \frac{1}{2m} \nabla^2 \phi \quad (2.124)$$

But this is the same equation obeyed by the original field (2.105)! Except this time, it really is the Schrödinger equation, complete with the usual probabilistic interpretation for the wavefunction  $\phi$ . Note in particular that the conserved charge arising from the Noether current (2.107) is  $Q = \int d^3x |\phi(\hat{x})|^2$  which is the total probability.

Historically, the fact that the equation for the classical field (2.105) and the one-particle wavefunction (2.124) coincide caused some confusion. It was thought that perhaps we are quantizing the wavefunction itself and the resulting name “second quantization” is still sometimes used today to mean quantum field theory. It’s important to stress that, despite the name, we’re not quantizing anything twice! We simply quantize a classical field once. Nonetheless, in practice it’s useful to know that if we treat the one-particle Schrödinger equation as the equation for a quantum field then it will give the correct generalization to multi-particle states.

## Interactions

Often in quantum mechanics, we’re interested in particles moving in some fixed background potential  $V(\hat{x})$ . This can be easily incorporated into field theory by working with a Lagrangian with explicit  $\hat{x}$  dependence,

$$L = i\psi^\dagger \dot{\psi} - \frac{1}{2m} \nabla \psi^\dagger \nabla \psi - V(\hat{x}) \psi^\dagger \psi \quad (2.125)$$

Note that this Lagrangian doesn’t respect translational symmetry and we won’t have the associated energy-momentum tensor. While such Lagrangians are useful in condensed matter physics, we rarely (or never) come across them in high-energy physics, where all equations obey translational (and Lorentz) invariance.

One can also consider interactions *between* particles. Obviously these are only important for  $n$  particle states with  $n \geq 2$ . We therefore expect them to arise from additions to the Lagrangian of the form

$$\Delta L = \psi^\wedge(\rightarrow x) \psi^\wedge(\rightarrow x) \psi(\star) \psi(\star) \quad (2.126)$$

which, in the quantum theory, is an operator which destroys two particles before creating two new ones. Such terms in the Lagrangian will indeed lead to inter-particle forces, both in the non-relativistic and relativistic setting. In the next section we explore these types of interaction in detail for relativistic theories.

### 3. Interacting Fields

The free field theories that we've discussed so far are very special: we can determine their spectrum, but nothing interesting then happens. They have particle excitations, but these particles don't interact with each other.

Here we'll start to examine more complicated theories that include interaction terms. These will take the form of higher order terms in the Lagrangian. We'll start by asking what kind of *small* perturbations we can add to the theory. For example, consider the Lagrangian for a real scalar field,

$$L = \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - \frac{1}{2} m^2 \varphi^2 - \sum_{n \geq 3} \frac{\lambda_n}{n!} \varphi^n \quad (3.1)$$

The coefficients  $\lambda_n$  are called *coupling constants*. What restrictions do we have on  $\lambda_n$  to ensure that the additional terms are small perturbations? You might think that we need simply make " $\lambda_n = 1$ ". But this isn't quite right. To see why this is the case, let's do some dimensional analysis. Firstly, note that the action has dimensions of angular momentum or, equivalently, the same dimensions as  $k$ . Since we've set  $k = 1$ , using the convention described in the introduction, we have  $[S] = 0$ . With  $S = \int d^4x L$ , and  $[d^4x] = 4$ , the Lagrangian density must therefore have

$$[L] = 4 \quad (3.2)$$

What does this mean for the Lagrangian (3.1)? Since  $[\partial_\mu] = 1$ , we can read off the mass dimensions of all the factors to find,

$$[\varphi] = 1 , \quad [m] = 1 , \quad [\lambda_n] = 4 - n \quad (3.3)$$

So now we see why we can't simply say we need  $\lambda_n = 1$ , because this statement only makes sense for dimensionless quantities. The various terms, parameterized by  $\lambda_n$ , fall into three different categories

- $[\lambda_3] = 1$ : For this term, the dimensionless parameter is  $\lambda_3/E$ , where  $E$  has dimensions of mass. Typically in quantum field theory,  $E$  is the energy scale of the process of interest. This means that  $\lambda_3 \varphi^3/3!$  is a small perturbation at high energies  $E \gg \lambda_3$ , but a large perturbation at low energies  $E \ll \lambda_3$ . Terms that we add to the Lagrangian with this behavior are called *relevant* because they're most relevant at low energies (which, after all, is where most of the physics we see lies). In a relativistic theory,  $E > m$ , so we can always make this perturbation small by taking  $\lambda_3 \ll m$ .

- $[\lambda_4] = 0$ : this term is small if  $\lambda_4 \ll 1$ . Such perturbations are called *marginal*.
- $[\lambda_n] < 0$  for  $n \geq 5$ : The dimensionless parameter is  $(\lambda_n E)^{\frac{1}{n-4}}$ , which is small at low-energies and large at high energies. Such perturbations are called *irrelevant*.

As you'll see later, it is typically impossible to avoid high energy processes in quantum field theory. (We've already seen a glimpse of this in computing the vacuum energy). This means that we might expect problems with irrelevant operators. Indeed, these lead to "non-renormalizable" field theories in which one cannot make sense of the infinities at arbitrarily high energies. This doesn't necessarily mean that the theory is useless; just that it is incomplete at some energy scale.

Let me note however that the naive assignment of relevant, marginal and irrelevant is not always fixed in stone: quantum corrections can sometimes change the character of an operator.

### An Important Aside: Why QFT is Simple

Typically in a quantum field theory, only the relevant and marginal couplings are important. This is basically because, as we've seen above, the irrelevant couplings become small at low-energies. This is a huge help: of the infinite number of interaction terms that we could write down, only a handful are actually needed (just two in the case of the real scalar field described above).

Let's look at this a little more. Suppose that we some day discover the true superduper "theory of everything unimportant" that describes the world at very high energy scales, say the GUT scale, or the Planck scale. Whatever this scale is, let's call it  $\Lambda$ . It is an energy scale, so  $[\Lambda] = 1$ . Now we want to understand the laws of physics down at our puny energy scale  $E \ll \Lambda$ . Let's further suppose that down at the energy scale  $E$ , the laws of physics are described by a real scalar field. (They're not of course: they're described by non-Abelian gauge fields and fermions, but the same argument applies in that case so bear with me). This scalar field will have some complicated interaction terms (3.1), where the precise form is dictated by all the stuff that's going on in the high energy superduper theory. What are these interactions? Well, we could write our dimensionful coupling constants  $\lambda_n$  in terms of dimensionless couplings  $g_n$ , multiplied by a suitable power of the relevant scale  $\Lambda$ ,

$$\lambda_n = \frac{g_n}{\Lambda^{n-4}} \quad (3.4)$$

The exact values of dimensionless couplings  $g_n$  depend on the details of the high-energy superduper theory, but typically one expects them to be of order 1:  $g_n \sim O(1)$ . This

means that for experiments at small energies  $E \ll \Lambda$ , the interaction terms of the form  $\varphi^n$  with  $n > 4$  will be suppressed by powers of  $(E/\Lambda)^{n-4}$ . This is usually a suppression by many orders of magnitude. (e.g. for the energies  $E$  explored at the LHC,  $E/M_{\text{Pl}} \sim 10^{-16}$ ). It is this simple argument, based on dimensional analysis, that ensures that we need only focus on the first few terms in the interaction: those which are relevant and marginal. It also means that if we only have access to low-energy experiments (which we do!), it's going to be very difficult to figure out the high energy theory (which it is!), because its effects are highly diluted except for the relevant and marginal interactions. The discussion given above is a poor man's version of the ideas of *effective field theory* and *Wilson's renormalization group*, about which you can learn more in the "Statistical Field Theory" course.

### Examples of Weakly Coupled Theories

In this course we'll study only weakly coupled field theories i.e. ones that can truly be considered as small perturbations of the free field theory at all energies. In this section, we'll look at two types of interactions

#### 1) $\varphi^4$ theory:

$$L = \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - \frac{1}{2} m^2 \varphi^2 - \frac{\lambda}{4!} \varphi^4 \quad (3.5)$$

with  $\lambda < 1$ . We can get a hint for what the effects of this extra term will be. Expanding out  $\varphi^4$  in terms of  $a_{p \rightarrow}^\dagger$  and  $a_{p \rightarrow}$ , we see a sum of interactions that look like

$$a_{p \rightarrow}^\dagger a_{p \rightarrow}^\dagger a_{p \rightarrow}^\dagger a_{p \rightarrow} \quad \text{and} \quad a_{p \rightarrow}^\dagger a_{p \rightarrow}^\dagger a_{p \rightarrow}^\dagger a_{p \rightarrow} \quad \text{etc.} \quad (3.6)$$

These will create and destroy particles. This suggests that the  $\varphi^4$  Lagrangian describes a theory in which particle number is not conserved. Indeed, we could check that the number operator  $N$  now satisfies  $[H, N] \neq 0$ .

#### 2) Scalar Yukawa Theory

$$L = \partial_\mu \psi^\dagger \partial^\mu \psi + \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi - M^2 \psi^\dagger \psi - \frac{1}{2} m^2 \varphi^2 - g \psi^\dagger \psi \varphi \quad (3.7)$$

with  $g \ll M, m$ . This theory couples a complex scalar  $\psi$  to a real scalar  $\varphi$ . While the individual particle numbers of  $\psi$  and  $\varphi$  are no longer conserved, we do still have a symmetry rotating the phase of  $\psi$ , ensuring the existence of the charge  $Q$  defined in (2.75) such that  $[Q, H] = 0$ . This means that the number of  $\psi$  particles minus the number of  $\psi^-$  anti-particles is conserved. It is common practice to denote the anti-particle as  $\psi^-$ .

The scalar Yukawa theory has a slightly worrying aspect: the potential has a stable local minimum at  $\varphi = \psi = 0$ , but is unbounded below for large enough  $-g\varphi$ . This means we shouldn't try to push this theory too far.

### A Comment on Strongly Coupled Field Theories

In this course we restrict attention to weakly coupled field theories where we can use perturbative techniques. The study of strongly coupled field theories is much more difficult, and one of the major research areas in theoretical physics. For example, some of the amazing things that can happen include

- **Charge Fractionalization:** Although electrons have electric charge 1, under the right conditions the elementary excitations in a solid have fractional charge  $1/N$  (where  $N \in 2\mathbb{Z} + 1$ ). For example, this occurs in the fractional quantum Hall effect.
- **Confinement:** The elementary excitations of quantum chromodynamics (QCD) are quarks. But they *never* appear on their own, only in groups of three (in a baryon) or with an anti-quark (in a meson). They are confined.
- **Emergent Space:** There are field theories in four dimensions which at strong coupling become quantum gravity theories in ten dimensions! The strong coupling effects cause the excitations to act as if they're gravitons moving in higher dimensions. This is quite extraordinary and still poorly understood. It's called the AdS/CFT correspondence.

### 3.1 The Interaction Picture

There's a useful viewpoint in quantum mechanics to describe situations where we have small perturbations to a well-understood Hamiltonian. Let's return to the familiar ground of quantum mechanics with a finite number of degrees of freedom for a moment. In the Schrödinger picture, the states evolve as

$$i \frac{d|\psi\rangle_s}{dt} = H |\psi\rangle_s \quad (3.8)$$

while the operators  $O_S$  are independent of time.

In contrast, in the Heisenberg picture the states are fixed and the operators change in time

$$\begin{aligned} O_H(t) &= e^{iHt} O_S e^{-iHt} \\ |\psi\rangle_H &= e^{iHt} |\psi\rangle_S \end{aligned} \quad (3.9)$$

The *interaction picture* is a hybrid of the two. We split the Hamiltonian up as

$$H = H_0 + H_{\text{int}} \quad (3.10)$$

The time dependence of operators is governed by  $H_0$ , while the time dependence of states is governed by  $H_{\text{int}}$ . Although the split into  $H_0$  and  $H_{\text{int}}$  is arbitrary, it's useful when  $H_0$  is soluble (for example, when  $H_0$  is the Hamiltonian for a free field theory). The states and operators in the interaction picture will be denoted by a subscript  $I$  and are given by,

$$\begin{aligned} |\psi(t)\rangle_I &= e^{iH_0 t} |\psi(t)\rangle_S \\ O_I(t) &= e^{iH_0 t} O_S e^{-iH_0 t} \end{aligned} \quad (3.11)$$

This last equation also applies to  $H_{\text{int}}$ , which is time dependent. The interaction Hamiltonian in the interaction picture is,

$$H_I \equiv (H_{\text{int}})_I = e^{iH_0 t} (H_{\text{int}})_S e^{-iH_0 t} \quad (3.12)$$

The Schrödinger equation for states in the interaction picture can be derived starting from the Schrödinger picture

$$\begin{aligned} i \frac{d|\psi\rangle_S}{dt} &= H_S |\psi\rangle_S \quad \Rightarrow \quad i \frac{d}{dt} e^{-iH_0 t} |\psi\rangle_I = (H_0 + H_{\text{int}})_S e^{-iH_0 t} |\psi\rangle_I \\ &\Rightarrow \quad i \frac{d}{dt} |\psi\rangle_I = e^{iH_0 t} (H_{\text{int}})_S e^{-iH_0 t} |\psi\rangle_I \end{aligned} \quad (3.13)$$

So we learn that

$$i \frac{d|\psi\rangle_I}{dt} = H_I(t) |\psi\rangle_I \quad (3.14)$$

### 3.1.1 Dyson's Formula

"Well, Birmingham has much the best theoretical physicist to work with, Peierls; Bristol has much the best experimental physicist, Powell; Cambridge has some excellent architecture. You can make your choice."

*Oppenheimer's advice to Dyson on which university position to accept.*

We want to solve (3.14). Let's write the solution as

$$|\psi(t)\rangle_I = U(t, t_0) |\psi(t_0)\rangle_I \quad (3.15)$$

where  $U(t, t_0)$  is a unitary time evolution operator such that  $U(t_1, t_2)U(t_2, t_3) = U(t_1, t_3)$  and  $U(t, t) = 1$ . Then the interaction picture Schrödinger equation (3.14) requires that

$$i \frac{dU}{dt} = H_I(t) U \quad (3.16)$$

If  $H_I$  were a function, then we could simply solve this by

$$U(t, t_0) \stackrel{?}{=} \exp \int_{t_0}^t H_I(t') dt' \quad (3.17)$$

But there's a problem. Our Hamiltonian  $H_I$  is an operator, and we have ordering issues. Let's see why this causes trouble. The exponential of an operator is defined in terms of the expansion,

$$\exp \int_{t_0}^t H_I(t') dt' = 1 - i \int_{t_0}^t H_I(t') dt' + \frac{(-i)^2}{2} \int_{t_0}^t \int_{t'}^t H_I(t') H_I(t'') dt' dt'' + \dots \quad (3.18)$$

But when we try to differentiate this with respect to  $t$ , we find that the quadratic term gives us

$$-\frac{1}{2} \int_{t_0}^t H_I(t') dt' H_I(t) - \frac{1}{2} H_I(t) \int_{t_0}^t H_I(t') dt' \quad (3.19)$$

Now the second term here looks good, since it will give part of the  $H_I(t)U$  that we need on the right-hand side of (3.16). But the first term is no good since the  $H_I(t)$  sits on the wrong side of the integral term, and we can't commute it through because  $[H_I(t'), H_I(t)] \neq 0$  when  $t' \neq t$ . So what's the way around this?

**Claim:** The solution to (3.16) is given by *Dyson's Formula*. (Essentially first figured out by Dirac, although the compact notation is due to Dyson).

$$U(t, t_0) = T \exp \int_{t_0}^t H_I(t') dt' \quad (3.20)$$

where  $T$  stands for *time ordering* where operators evaluated at later times are placed to the left

$$T(O_1(t_1) O_2(t_2)) = \begin{cases} O_1(t_1) O_2(t_2) & t_1 > t_2 \\ O_2(t_2) O_1(t_1) & t_2 > t_1 \end{cases} \quad (3.21)$$

Expanding out the expression (3.20), we now have

$$\begin{aligned} U(t, t_0) = 1 - i \int_{t_0}^t dt' H_I(t') + \frac{(-i)^2}{2} \int_{t_0}^t \int_{t'}^t dt' dt'' H_I(t') H_I(t'') \\ + \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' H_I(t') H_I(t'') + \dots \end{aligned}$$

Actually these last two terms double up since

$$\begin{aligned} \int_{t_0}^t \int_{t'} dt^{rr} H_l(t^{rr}) H_l(t^r) &= \int_{t_0}^t dt^{rr} \int_{t_0}^{t''} dt^r H_l(t^{rr}) H_l(t^r) \\ &= \int_{t_0}^t dt^r \int_{t_0}^{t'} dt^{rr} H_l(t^r) H_l(t^{rr}) \end{aligned} \quad (3.22)$$

where the range of integration in the first expression is over  $t^{rr} \geq t^r$ , while in the second expression it is  $t^r \leq t^{rr}$  which is, of course, the same thing. The final expression is the same as the second expression by a simple relabelling. This means that we can write

$$U(t, t_0) = 1 - i \int_{t_0}^t dt^r H_l(t^r) + (-i)^2 \int_{t_0}^t dt^r \int_{t_0}^{t'} dt^{rr} H_l(t^r) H_l(t^{rr}) + \dots \quad (3.23)$$

**Proof:** The proof of Dyson's formula is simpler than explaining what all the notation means! Firstly observe that under the  $T$  sign, all operators commute (since their order is already fixed by the  $T$  sign). Thus

$$\begin{aligned} i \frac{\partial}{\partial t} T \exp \left[ -i \int_{t_0}^t dt^r H_l(t^r) \right] &= T \exp \left[ -i \int_{t_0}^t dt^r H_l(t^r) \right] \int_{t_0}^t dt^r \\ &= H_l(t) T \exp \left[ -i \int_{t_0}^t dt^r H_l(t^r) \right] \end{aligned} \quad (3.24)$$

since  $t$ , being the upper limit of the integral, is the latest time so  $H_l(t)$  can be pulled out to the left.

Before moving on, I should confess that Dyson's formula is rather formal. It is typically very hard to compute time ordered exponentials in practice. The power of the formula comes from the expansion which is valid when  $H_l$  is small and is very easily computed.

### 3.2 A First Look at Scattering

Let us now apply the interaction picture to field theory, starting with the interaction Hamiltonian for our scalar Yukawa theory,

$$H_{\text{int}} = g \int d^3x \psi^\dagger \psi \varphi \quad (3.25)$$

Unlike the free theories discussed in Section 2, this interaction doesn't conserve particle number, allowing particles of one type to morph into others. To see why this is, we use

the interaction picture and follow the evolution of the state:  $|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle$ , where  $U(t, t_0)$  is given by Dyson's formula (3.20) which is an expansion in powers of  $H_{\text{int}}$ . But  $H_{\text{int}}$  contains creation and annihilation operators for each type of particle. In particular,

- $\varphi \sim a + a^\dagger$ : This operator can create or destroy  $\varphi$  particles. Let's call them *mesons*.
- $\psi \sim b + c^\dagger$ : This operator can destroy  $\psi$  particles through  $b$ , and create anti-particles through  $c^\dagger$ . Let's call these particles *nucleons*. Of course, in reality nucleons are spin 1/2 particles, and don't arise from the quantization of a scalar field. But we'll treat our scalar Yukawa theory as a toy model for nucleons interacting with mesons.
- $\psi^\dagger \sim b^\dagger + c$ : This operator can create nucleons through  $b^\dagger$ , and destroy anti-nucleons through  $c$ .

Importantly,  $Q = N_c - N_b$  remains conserved in the presence of  $H_{\text{int}}$ . At first order in perturbation theory, we find terms in  $H_{\text{int}}$  like  $c^\dagger b^\dagger a$ . This kills a meson, producing a nucleon-anti-nucleon pair. It will contribute to meson decay  $\varphi \rightarrow \psi \bar{\psi}$ .

At second order in perturbation theory, we'll have more complicated terms in  $(H_{\text{int}})^2$ , for example  $(c^\dagger b^\dagger a)(c b a^\dagger)$ . This term will give contributions to scattering processes  $\psi \bar{\psi} \rightarrow \varphi \rightarrow \psi \bar{\psi}$ . The rest of this section is devoted to computing the quantum amplitudes for these processes to occur.

To calculate amplitudes we make an important, and slightly dodgy, assumption:

Initial and final states are eigenstates of the free theory

This means that we take the initial state  $|i\rangle$  at  $t \rightarrow -\infty$ , and the final state  $|f\rangle$  at  $t \rightarrow +\infty$ , to be eigenstates of the free Hamiltonian  $H_0$ . At some level, this sounds plausible: at  $t \rightarrow -\infty$ , the particles in a scattering process are far separated and don't feel the effects of each other. Furthermore, we intuitively expect these states to be eigenstates of the individual number operators  $N$ , which commute with  $H_0$ , but not  $H_{\text{int}}$ . As the particles approach each other, they interact briefly, before departing again, each going on its own merry way. The amplitude to go from  $|i\rangle$  to  $|f\rangle$  is

$$\lim_{t_\pm \rightarrow \pm\infty} \langle f | U(t_+, t_-) | i \rangle \equiv \langle f | S | i \rangle \quad (3.26)$$

where the unitary operator  $S$  is known as the *S-matrix*. ( $S$  is for scattering). There are a number of reasons why the assumption of non-interacting initial and final states is shaky:

- Obviously we can't cope with bound states. For example, this formalism can't describe the scattering of an electron and proton which collide, bind, and leave as a Hydrogen atom. It's possible to circumvent this objection since it turns out that bound states show up as poles in the S-matrix.
- More importantly, a single particle, a long way from its neighbors, is never alone in field theory. This is true even in classical electrodynamics, where the electron sources the electromagnetic field from which it can never escape. In quantum electrodynamics (QED), a related fact is that there is a cloud of *virtual* photons surrounding the electron. This line of thought gets us into the issues of renormalization — more on this next term in the “AQFT” course. Nevertheless, motivated by this problem, after developing scattering theory using the assumption of non-interacting asymptotic states, we'll mention a better way.

### 3.2.1 An Example: Meson Decay

Consider the relativistically normalized initial and final states,

$$\begin{aligned} |i\rangle &= \sqrt{\frac{2E_{p\rightarrow}}{p\rightarrow}} a^\dagger |0\rangle \\ |f\rangle &= \sqrt{\frac{4E_{q_1\rightarrow_1} E_{q_2\rightarrow_2}}{q_1\rightarrow_1 q_2\rightarrow_2}} b^\dagger c^\dagger |0\rangle \end{aligned} \quad (3.27)$$

The initial state contains a single meson of momentum  $p$ ; the final state contains a nucleon-anti-nucleon pair of momentum  $q_1$  and  $q_2$ . We may compute the amplitude for the decay of a meson to a nucleon-anti-nucleon pair. To leading order in  $g$ , it is

$$\langle f | S | i \rangle = -ig \langle f | \int d^4x \psi^\dagger(x) \psi(x) \varphi(x) | i \rangle \quad (3.28)$$

Let's go slowly. We first expand out  $\varphi \sim a + a^\dagger$  using (2.84). (Remember that the  $\varphi$  in this formula is in the interaction picture, which is the same as the Heisenberg picture of the free theory). The  $a$  piece will turn  $|i\rangle$  into something proportional to  $|0\rangle$ , while the  $a^\dagger$  piece will turn  $|i\rangle$  into a two meson state. But the two meson state will have zero overlap with  $\langle f |$ , and there's nothing in the  $\psi$  and  $\psi^\dagger$  operators that lie between them to change this fact. So we have

$$\begin{aligned} \langle f | S | i \rangle &= -ig \langle f | \int d^4x \psi^\dagger(x) \psi(x) a^\dagger | i \rangle \\ &= -ig \langle f | \int d^4x \psi^\dagger(x) \psi(x) e^{-ip \cdot x} | 0 \rangle \end{aligned} \quad (3.29)$$

where, in the second line, we've commuted  $a^\dagger$  past  $a_{p\rightarrow}$ , picking up a  $\delta^{(3)}(\vec{p} - \vec{k})$  delta-function which kills the  $d^3k$  integral. We now similarly expand out  $\psi \sim b + b^\dagger$  and

$\psi^\dagger \sim b^\dagger + c$ . To get non-zero overlap with  $\langle f |$ , only the  $b^\dagger$  and  $c^\dagger$  contribute, for they create the nucleon and anti-nucleon from  $|0\rangle$ . We then have

$$\begin{aligned} \langle f | S | i \rangle &= -ig \langle 0 | \frac{d^4x d^3k_1 d^3k_2}{(2\pi)^6} \overbrace{\cancel{E_1} \cancel{E_2}}^{q \rightarrow q_1, q_2} c_{q_2}^\dagger b_{q_1}^\dagger c_{k_1}^\dagger b_{k_2}^\dagger | 0 \rangle e^{i(k_1 + k_2 - p) \cdot x} \\ &= -ig (2\pi)^4 \delta^{(4)}(q_1 + q_2 - p) \end{aligned} \quad (3.30)$$

and so we get our first quantum field theory amplitude.

Notice that the  $\delta$ -function puts constraints on the possible decays. In particular, the decay only happens at all if  $m \geq 2M$ . To see this, we may always boost ourselves to a reference frame where the meson is stationary, so  $p = (m, 0, 0, 0)$ . Then the delta function imposes momentum conservation, telling us that  $\rightarrow q_1 = -\rightarrow q_2$  and  $m = 2\sqrt{M^2 + |\vec{p}|^2}$ .

Later you will learn how to turn this quantum amplitude into something more physical, namely the lifetime of the meson. The reason this is a little tricky is that we must square the amplitude to get the probability for decay, which means we get the square of a  $\delta$ -function. We'll explain how to deal with this in Section 3.6 below, and again in next term's "Standard Model" course.

### 3.3 Wick's Theorem

From Dyson's formula, we want to compute quantities like  $\langle f | T \{H_1(x_1) \dots H_l(x_n)\} | i \rangle$ , where  $|i\rangle$  and  $|f\rangle$  are eigenstates of the free theory. The ordering of the operators is fixed by  $T$ , time ordering. However, since the  $H_i$ 's contain certain creation and annihilation operators, our life will be much simpler if we can start to move all annihilation operators to the right where they can start killing things in  $|i\rangle$ . Recall that this is the definition of normal ordering. Wick's theorem tells us how to go from time ordered products to normal ordered products.

#### 3.3.1 An Example: Recovering the Propagator

Let's start simple. Consider a real scalar field which we decompose in the Heisenberg picture as

$$\varphi(x) = \varphi^+(x) + \varphi^-(x) \quad (3.31)$$

where

$$\begin{aligned} \varphi^+(x) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{\frac{2E}{\cancel{p}}}} a_{\vec{p}} e^{-ip \cdot x} \\ \varphi^-(x) &= - \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{\frac{2E}{\cancel{p}}}} a_{\vec{p}}^\dagger e^{+ip \cdot x} \end{aligned} \quad (3.32)$$

where the  $\pm$  signs on  $\varphi^\pm$  make little sense, but apparently you have Pauli and Heisenberg to blame. (They come about because  $\varphi^+ \sim e^{-iEt}$ , which is sometimes called the positive frequency piece, while  $\varphi^- \sim e^{+iEt}$  is the negative frequency piece). Then choosing  $x^0 > y^0$ , we have

$$\begin{aligned} T\varphi(x)\varphi(y) &= \varphi(x)\varphi(y) \\ &= (\varphi^+(x) + \varphi^-(x))(\varphi^+(y) + \varphi^-(y)) \\ &= \varphi^+(x)\varphi^+(y) + \varphi^-(x)\varphi^+(y) + \varphi^-(y)\varphi^+(x) + [\varphi^+(x), \varphi^-(y)] + \varphi^-(x)\varphi^-(y) \end{aligned} \quad (3.33)$$

where the last line is normal ordered, and for our troubles we have picked up the extra term  $D(x-y) = [\varphi^+(x), \varphi^-(y)]$  which is the propagator we met in (2.90). So for  $x^0 > y^0$  we have

$$T\varphi(x)\varphi(y) =: \varphi(x)\varphi(y) : + D(x-y) \quad (3.34)$$

Meanwhile, for  $y^0 > x^0$ , we may repeat the calculation to find

$$T\varphi(x)\varphi(y) =: \varphi(x)\varphi(y) : + D(y-x) \quad (3.35)$$

So putting this together, we have the final expression

$$T\varphi(x)\varphi(y) =: \varphi(x)\varphi(y) : + \Delta_F(x-y) \quad (3.36)$$

where  $\Delta_F(x-y)$  is the Feynman propagator defined in (2.93), for which we have the integral representation

$$\Delta_F(x-y) = \int \frac{d^4k}{(2\pi)^4} \frac{ie^{ik \cdot (x-y)}}{k^2 - m^2 + i\epsilon} \quad (3.37)$$

Let me reiterate a comment from Section 2: although  $T\varphi(x)\varphi(y)$  and  $: \varphi(x)\varphi(y) :$  are both operators, the difference between them is a c-number function,  $\Delta_F(x-y)$ .

**Definition:** We define the *contraction* of a pair of fields in a string of operators  $\dots \varphi(x_1) \dots \varphi(x_2) \dots$  to mean replacing those operators with the Feynman propagator, leaving all other operators untouched. We use the notation,

$$\dots \overset{\leftrightarrow}{\varphi(x_1)} \dots \overset{\leftrightarrow}{\varphi(x_2)} \dots \quad (3.38)$$

to denote contraction. So, for example,

$$\overset{\leftrightarrow}{\varphi(x)\varphi(y)} = \Delta_F(x-y) \quad (3.39)$$

A similar discussion holds for complex scalar fields. We have

$$T\psi(x)\psi^\dagger(y) = : \psi(x)\psi^\dagger(y) : + \Delta_F(x - y) \quad (3.40)$$

prompting us to define the contraction

$$\psi(x)\psi^\dagger(y) = \Delta_F(x - y) \quad \text{and} \quad \psi(x)\psi(y) = \psi^\dagger(x)\psi^\dagger(y) = 0 \quad (3.41)$$

### 3.3.2 Wick's Theorem

For any collection of fields  $\varphi_1 = \varphi(x_1)$ ,  $\varphi_2 = \varphi(x_2)$ , etc, we have

$$T(\varphi_1 \dots \varphi_n) = : \varphi_1 \dots \varphi_n : + : \text{all possible contractions} : \quad (3.42)$$

To see what the last part of this equation means, let's look at an example. For  $n = 4$ , the equation reads

$$\begin{aligned} T(\varphi_1 \varphi_2 \varphi_3 \varphi_4) &= : \varphi_1 \varphi_2 \varphi_3 \varphi_4 : + \varphi_1 \varphi_2 : \varphi_3 \varphi_4 : + \varphi_1 \varphi_3 : \varphi_2 \varphi_4 : + \text{four similar terms} \\ &\quad + \varphi_1 \varphi_2 \varphi_3 \varphi_4 + \varphi_1 \varphi_3 \varphi_2 \varphi_4 + \varphi_1 \varphi_4 \varphi_2 \varphi_3 \end{aligned} \quad (3.43)$$

**Proof:** The proof of Wick's theorem proceeds by induction and a little thought. It's true for  $n = 2$ . Suppose it's true for  $\varphi_2 \dots \varphi_n$  and now add  $\varphi_1$ . We'll take  $x_1^0 > x_k^0$  for all  $k = 2, \dots, n$ . Then we can pull  $\varphi_1$  out to the left of the time ordered product, writing

$$T(\varphi_1 \varphi_2 \dots \varphi_n) = (\varphi_1^+ + \varphi_1^-) \left( : \varphi_2 \dots \varphi_n : + : \text{contractions} : \right) \quad (3.44)$$

The  $\varphi_1^-$  term stays where it is since it is already normal ordered. But in order to write the right-hand side as a normal ordered product, the  $\varphi_1^+$  term has to make its way past the crowd of  $\varphi_k^-$  operators. Each time it moves past  $\varphi_k^-$ , we pick up a factor of  $\varphi_1 \varphi_k = \Delta_F(x_1 - x_k)$  from the commutator. (Try it!)

### 3.3.3 An Example: Nucleon Scattering

Let's look at  $\psi\psi \rightarrow \psi\psi$  scattering. We have the initial and final states

$$\begin{aligned} |i\rangle &= \sqrt{2E_{p_1}} \sqrt{2E_{p_2}} b_1^\dagger b_2^\dagger |0\rangle \equiv |p_1, p_2\rangle \\ |f\rangle &= \sqrt{2E_{p'_1}} \sqrt{2E_{p'_2}} b_1^\dagger b_2^\dagger |0\rangle \equiv |p'_1, p'_2\rangle \end{aligned} \quad (3.45)$$

We can then look at the expansion of  $\langle f | S | i \rangle$ . In fact, we really want to calculate  $\langle f | S - 1 | i \rangle$  since we're not interested in situations where no scattering occurs. At order  $g^2$  we have the term

$$\frac{(-ig)^2}{2} \int d^4x_1 d^4x_2 T \psi^\dagger(x_1)\psi(x_1)\varphi(x_1)\psi^\dagger(x_2)\psi(x_2)\varphi(x_2) \quad (3.46)$$

Now, using Wick's theorem we see there is a piece in the string of operators which looks like

$$:\psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2): \varphi(x_1)\varphi(x_2) \quad (3.47)$$

which will contribute to the scattering because the two  $\psi$  fields annihilate the  $\psi$  particles, while the two  $\psi^\dagger$  fields create  $\psi$  particles. Any other way of ordering the  $\psi$  and  $\psi^\dagger$  fields will give zero contribution. This means that we have

$$\begin{aligned} & \langle p_1^r, p_2^r | : \psi^\dagger(x_1)\psi(x_1)\psi^\dagger(x_2)\psi(x_2) : | p_1^r, p_2^r \rangle \\ &= \langle p_1^r, p_2^r | \psi^\dagger(x_1)\psi^\dagger(x_2) | 0 \rangle \langle 0 | \psi(x_1)\psi(x_2) | p_1^r, p_2^r \rangle \\ &= e^{ip'_1 \cdot x_1 + ip'_2 \cdot x_2} + e^{ip'_1 \cdot x_2 + ip'_2 \cdot x_1} - e^{-ip_1 \cdot x_1 - ip_2 \cdot x_2} + e^{-ip_1 \cdot x_2 - ip_2 \cdot x_1} \\ &= e^{ix_1 \cdot (p'_1 - p_1) + ix_2 \cdot (p'_2 - p_2)} + e^{ix_1 \cdot (p'_2 - p_1) + ix_2 \cdot (p'_1 - p_2)} + (x_1 \leftrightarrow x_2) \end{aligned} \quad (3.48)$$

where, in going to the third line, we've used the fact that for relativistically normalized states,

$$\langle 0 | \psi(x) | p \rangle = e^{-ip \cdot x} \quad (3.49)$$

Now let's insert this into (3.46), to get the expression for  $\langle f | S | i \rangle$  at order  $g^2$ ,

$$\frac{(-ig)^2}{2} \int d^4x_1 d^4x_2 [e^{i\cdots} + e^{i\cdots} + (x_1 \leftrightarrow x_2)] \int \frac{d^4k}{(2\pi)^4} \frac{ie^{ik \cdot (x_1 - x_2)}}{k^2 - m^2 + i\epsilon} \quad (3.50)$$

where the expression in square brackets is (3.48), while the final integral is the  $\varphi$  propagator which comes from the contraction in (3.47). Now the  $(x_1 \leftrightarrow x_2)$  terms double up with the others to cancel the factor of  $1/2$  out front. Meanwhile, the  $x_1$  and  $x_2$  integrals give delta-functions. We're left with the expression

$$\begin{aligned} (-ig)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i(2\pi)^8}{k^2 - m^2 + i\epsilon} & \delta^{(4)}(p_1^r - p_1 + k) \delta^{(4)}(p_2^r - p_2 - k) \\ & + \delta^{(4)}(p_1^r - p_1 + k) \delta^{(4)}(p_2^r - p_2 - k) \end{aligned} \quad (3.51)$$

Finally, we can trivially do the  $d^4k$  integral using the delta-functions to get

$$i(-ig)^2 \frac{1}{(p_1 - p_1^r)^2 - m^2 + i\epsilon} + \frac{1}{(p_1 - p_2^r)^2 - m^2 + i\epsilon} = \frac{(2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_1^r - p_2^r)}{(p_1 - p_1^r)^2 - m^2 + i\epsilon}$$

In fact, for this process we may drop the  $+i\epsilon$  terms since the denominator is never zero. To see this, we can go to the center of mass frame, where  $p_{\rightarrow 1} = -p_{\rightarrow 2}$  and, by

momentum conservation,  $|p_1| = |p^r|$ . This ensures that the 4-momentum of the meson is  $k = (0, p^r - p^r)$ , so  $k^2 < 0$ . We therefore have the end result,

$$i(-ig)^2 \frac{1}{(p_1 - p_1)^2 - m^2} + \frac{1}{(p_1 - p_2)^2 - m^2} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_1^r - p_2^r) \quad (3.52)$$

We will see another, much simpler way to reproduce this result shortly using Feynman diagrams. This will also shed light on the physical interpretation.

This calculation is also relevant for other scattering processes, such as  $\bar{\psi}\bar{\psi} \rightarrow \psi^-\psi^-$ ,  $\psi\psi \rightarrow \psi^-\psi^-$ . Each of these comes from the term (3.48) in Wick's theorem. However, we will never find a term that contributes to scattering  $\psi\psi \rightarrow \psi^-\psi^-$ , for this would violate the conservation of  $Q$  charge.

### Another Example: Meson-Nucleon Scattering

If we want to compute  $\psi\varphi \rightarrow \psi\varphi$  scattering at order  $g^2$ , we would need to pick out the term

$$:\psi^\dagger(x_1)\varphi(x_1)\psi(x_2)\varphi(x_2): \xrightarrow{x_1} \psi(x_1) \psi^\dagger(x_2) \quad (3.53)$$

and a similar term with  $\psi$  and  $\psi^\dagger$  exchanged. Once more, this term also contributes to similar scattering processes, including  $\psi^-\varphi \rightarrow \psi^-\varphi$  and  $\varphi\varphi \rightarrow \psi^-\psi^-$ .

### 3.4 Feynman Diagrams

“Like the silicon chips of more recent years, the Feynman diagram was bringing computation to the masses.”

*Julian Schwinger*

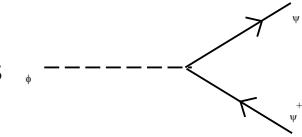
As the above example demonstrates, to actually compute scattering amplitudes using Wick's theorem is rather tedious. There's a much better way. It requires drawing pretty pictures. These pictures represent the expansion of  $\langle f | S | i \rangle$  and we will learn how to associate numbers (or at least integrals) to them. These pictures are called *Feynman diagrams*.

The object that we really want to compute is  $\langle f | S - 1 | i \rangle$ , since we're not interested in processes where no scattering occurs. The various terms in the perturbative expansion can be represented pictorially as follows

- Draw an external line for each particle in the initial state  $|i\rangle$  and each particle in the final state  $|f\rangle$ . We'll choose dotted lines for mesons, and solid lines for nucleons. Assign a directed momentum  $p$  to each line. Further, add an arrow to

solid lines to denote its charge; we'll choose an incoming (outgoing) arrow in the initial state for  $\psi$  ( $\bar{\psi}$ ). We choose the reverse convention for the final state, where an outgoing arrow denotes  $\psi$ .

- Join the external lines together with trivalent vertices



Each such diagram you can draw is in 1-1 correspondence with the terms in the expansion of  $\langle f | S - 1 | i \rangle$ .

### 3.4.1 Feynman Rules

To each diagram we associate a number, using the *Feynman rules*

- Add a momentum  $k$  to each internal line
- To each vertex, write down a factor of

$$(-ig) (2\pi)^4 \delta^{(4)}(\sum_i k_i) \quad (3.54)$$

where  $\sum_i k_i$  is the sum of all momenta flowing *into* the vertex.

- For each internal dotted line, corresponding to a  $\varphi$  particle with momentum  $k$ , we write down a factor of

$$\int \frac{d^4k}{(2\pi)^4} \frac{i}{k^2 - m^2 + i\epsilon} \quad (3.55)$$

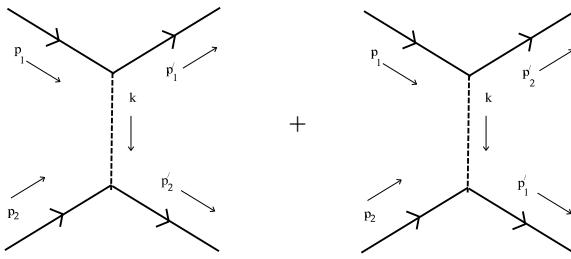
We include the same factor for solid internal  $\psi$  lines, with  $m$  replaced by the nucleon mass  $M$ .

### 3.5 Examples of Scattering Amplitudes

Let's apply the Feynman rules to compute the amplitudes for various processes. We start with something familiar:

#### Nucleon Scattering Revisited

Let's look at how this works for the  $\psi\psi \rightarrow \psi\psi$  scattering at order  $g^2$ . We can write down the two simplest diagrams contributing to this process. They are shown in Figure 9.



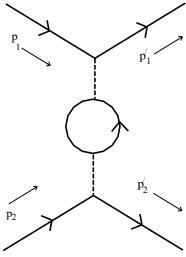
**Figure 9:** The two lowest order Feynman diagrams for nucleon scattering.

Applying the Feynman rules to these diagrams, we get

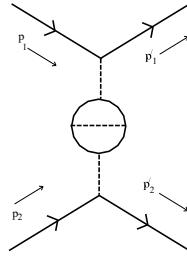
$$i(-ig)^2 \frac{1}{(p_1 - p_1^r)^2 - m^2} + \frac{1}{(p_1 - p_2^r)^2 - m^2} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p_1^r - p_2^r) \quad (3.56)$$

which agrees with the calculation (3.51) that we performed earlier. There is a nice physical interpretation of these diagrams. We talk, rather loosely, of the nucleons exchanging a meson which, in the first diagram, has momentum  $k = (p_1 - p_1^r) = (p_2 - p_2^r)$ . This meson doesn't satisfy the usual energy dispersion relation, because  $k^2 \neq m^2$ : the meson is called a *virtual particle* and is said to be *off-shell* (or, sometimes, off mass-shell). Heuristically, it can't live long enough for its energy to be measured to great accuracy. In contrast, the momentum on the external, nucleon legs all satisfy  $p^2 = M^2$ , the mass of the nucleon. They are *on-shell*. One final note: the addition of the two diagrams above ensures that the particles satisfy Bose statistics.

There are also more complicated diagrams which will contribute to the scattering process at higher orders. For example, we have the two diagrams shown in Figures 10 and 11, and similar diagrams with  $p_1^r$  and  $p_2^r$  exchanged. Using the Feynman rules, each of these diagrams translates into an integral that we will not attempt to calculate here. And so we go on, with increasingly complicated diagrams, all appearing at higher order in the coupling constant  $g$ .



**Figure 10:** A contribution at  $O(g^4)$ .



**Figure 11:** A contribution at  $O(g^6)$

## Amplitudes

Our final result for the nucleon scattering amplitude  $\langle f | S - 1 | i \rangle$  at order  $g^2$  was

$$i(-ig)^2 \frac{1}{(p_1 - p'_1)^2 - m^2} + \frac{1}{(p_1 - p'_2)^2 - m^2} (2\pi)^4 \delta^{(4)}(p_1 + p_2 - p'_1 - p'_2)$$

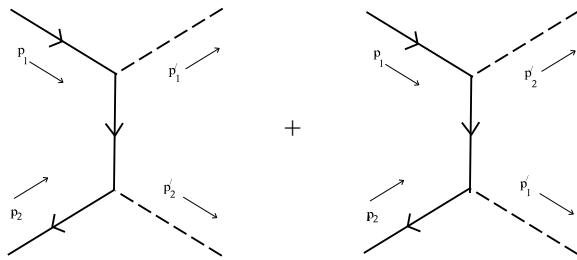
The  $\delta$ -function follows from the conservation of 4-momentum which, in turn, follows from spacetime translational invariance. It is common to all S-matrix elements. We will define the amplitude  $A_{fi}$  by stripping off this momentum-conserving delta-function,

$$\langle f | S - 1 | i \rangle = i A_{fi} (2\pi)^4 \delta^{(4)}(p_F - p_i) \quad (3.57)$$

where  $p_i$  ( $p_F$ ) is the sum of the initial (final) 4-momenta, and the factor of  $i$  out front is a convention which is there to match non-relativistic quantum mechanics. We can now refine our Feynman rules to compute the amplitude  $iA_{fi}$  itself:

- Draw all possible diagrams with appropriate external legs and impose 4-momentum conservation at each vertex.
- Write down a factor of  $(-ig)$  at each vertex.
- For each internal line, write down the propagator
- Integrate over momentum  $k$  flowing through each loop  $d^4k/(2\pi)^4$ .

This last step deserves a short explanation. The diagrams we've computed so far have no loops. They are *tree level* diagrams. It's not hard to convince yourself that in tree diagrams, momentum conservation at each vertex is sufficient to determine the momentum flowing through each internal line. For diagrams with loops, such as those shown in Figures 10 and 11, this is no longer the case.



**Figure 12:** The two lowest order Feynman diagrams for nucleon to meson scattering.

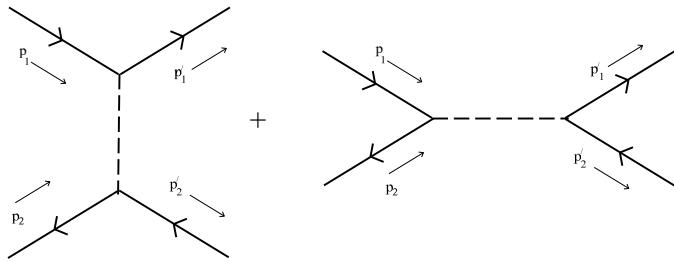
### Nucleon to Meson Scattering

Let's now look at the amplitude for a nucleon-anti-nucleon pair to annihilate into a pair of mesons:  $\psi \bar{\psi} \rightarrow \varphi \varphi$ . The simplest Feynman diagrams for this process are shown in Figure 12 where the virtual particle in these diagrams is now the nucleon  $\psi$  rather than the meson  $\varphi$ . This fact is reflected in the denominator of the amplitudes which are given by

$$iA = (-ig)^2 \frac{i}{(p_1 - q^r)^2 - M^2} + \frac{i}{(p_1 - p_2^r)^2 - M^2} \quad (3.58)$$

As in (3.52), we've dropped the  $i\epsilon$  from the propagators as the denominator never vanishes.

### Nucleon-Anti-Nucleon Scattering



**Figure 13:** The two lowest order Feynman diagrams for nucleon-anti-nucleon scattering.

For the scattering of a nucleon and an anti-nucleon,  $\psi \bar{\psi} \rightarrow \psi \bar{\psi}$ , the Feynman diagrams are a little different. At lowest order, they are given by the diagrams of Figure 13. It is a simple matter to write down the amplitude using the Feynman rules,

$$iA = (-ig)^2 \frac{i}{(p_1 - q^r)^2 - m^2} + \frac{i}{(p_1 + p_2)^2 - m^2 + i\epsilon} \quad (3.59)$$

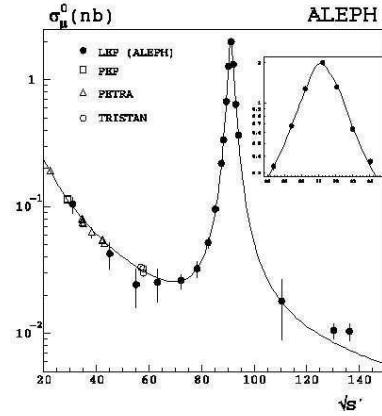
Notice that the momentum dependence in the second term is different from that of nucleon-nucleon scattering (3.56), reflecting the different Feynman diagram that contributes to the process. In the center of mass frame,  $p \rightarrow_1 = -p \rightarrow_2$ , the denominator of the second term is  $4(M^2 + p \rightarrow_1^2) - m^2$ . If  $m < 2M$ , then this term never vanishes and we may drop the  $i\epsilon$ . In contrast, if  $m > 2M$ , then the amplitude corresponding to the second diagram diverges at some value of  $p \rightarrow$ . In this case it turns out that we may also neglect the  $i\epsilon$  term, although for a different reason: the meson is unstable when  $m > 2M$ , a result we derived in (3.30). When correctly treated, this instability adds a finite imaginary piece to the denominator which overwhelms the  $i\epsilon$ . Nonetheless, the increase in the scattering amplitude which we see in the second diagram when  $4(M^2 + p \rightarrow^2) = m^2$  is what allows us to discover new particles: they appear as a resonance in the cross section. For example, the Figure 14 shows the cross-section (roughly the amplitude squared) plotted vertically for  $e^+e^- \rightarrow \mu^+\mu^-$  scattering from the ALEPH experiment in CERN. The horizontal axis shows the center of mass energy. The curve rises sharply around 91 GeV, the mass of the Z-boson.

### Meson Scattering

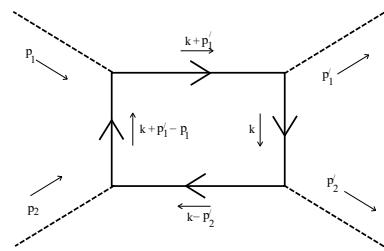
For  $\varphi\varphi \rightarrow \varphi\varphi$ , the simplest diagram we can write down has a single loop, and momentum conservation at each vertex is no longer sufficient to determine every momentum passing through the diagram. We choose to assign the single undetermined momentum  $k$  to the right-hand propagator. All other momenta are then determined. The amplitude corresponding to the diagram shown in the figure is

$$\begin{aligned} & (-ig)^4 \int \frac{d^4k}{(2\pi)^4} \frac{1}{(k^2 - M^2 + i\epsilon)((k + p \rightarrow_1)^2 - M^2 + i\epsilon)} \\ & \times \frac{1}{((k + p \rightarrow_1 - p_1)^2 - M^2 + i\epsilon)((k - p \rightarrow_2)^2 - M^2 + i\epsilon)} \end{aligned}$$

These integrals can be tricky. For large  $k$ , this integral goes as  $d^4 k/k^8$ , which is at least convergent as  $k \rightarrow \infty$ . But this won't always be the case!



**Figure 14:**



**Figure 15:**

### 3.5.1 Mandelstam Variables

We see that in many of the amplitudes above — in particular those that include the exchange of just a single particle — the same combinations of momenta are appearing frequently in the denominators. There are standard names for various sums and differences of momenta: they are known as *Mandelstam variables*. They are

$$\begin{aligned}s &= (p_1 + p_2)^2 = (p_1^r + p_2^r)^2 \\t &= (p_1 - p_1^r)^2 = (p_2 - p_2^r)^2 \\u &= (p_1 - p_2^r)^2 = (p_2 - p_1^r)^2\end{aligned}\tag{3.60}$$

where, as in the examples above,  $p_1$  and  $p_2$  are the momenta of the two initial particles, and  $p_1^r$  and  $p_2^r$  are the momenta of the final two particles. We can define these variables whether the particles involved in the scattering are the same or different. To get a feel for what these variables mean, let's assume all four particles are the same. We sit in the center of mass frame, so that the initial two particles have four-momenta

$$p_1 = (E, 0, 0, p) \quad \text{and} \quad p_2 = (E, 0, 0, -p)\tag{3.61}$$

The particles then scatter at some angle  $\vartheta$  and leave with momenta

$$p_1^r = (E, 0, p \sin \vartheta, p \cos \vartheta) \quad \text{and} \quad p_2^r = (E, 0, -p \sin \vartheta, -p \cos \vartheta)\tag{3.62}$$

Then from the above definitions, we have that

$$s = 4E^2 \quad \text{and} \quad t = -2p^2(1 - \cos \vartheta) \quad \text{and} \quad u = -2p^2(1 + \cos \vartheta)\tag{3.63}$$

The variable  $s$  measures the total center of mass energy of the collision, while the variables  $t$  and  $u$  are measures of the momentum exchanged between particles. (They are basically equivalent, just with the outgoing particles swapped around). Now the amplitudes that involve exchange of a single particle can be written simply in terms of the Mandelstam variables. For example, for nucleon-nucleon scattering, the amplitude (3.56) is schematically  $A \sim (t - m^2)^{-1} + (u - m^2)^{-1}$ . For the nucleon-anti-nucleon scattering, the amplitude (3.59) is  $A \sim (t - m^2)^{-1} + (s - m^2)^{-1}$ . We say that the first case involves “t-channel” and “u-channel” diagrams. Meanwhile the nucleon-anti-nucleon scattering is said to involve “t-channel” and “s-channel” diagrams. (The first diagram indeed includes a vertex that looks like the letter “T”).

Note that there is a relationship between the Mandelstam variables. When all the masses are the same we have  $s + t + u = 4M^2$ . When the masses of all 4 particles differ,

$$\sum_i M_i^2$$

this becomes  $s + t + u = \sum_i M_i^2$ .

### 3.5.2 The Yukawa Potential

So far we've computed the quantum amplitudes for various scattering processes. But these quantities are a little abstract. In Section 3.6 below (and again in next term's "Standard Model" course) we'll see how to turn amplitudes into measurable quantities such as cross-sections, or the lifetimes of unstable particles. Here we'll instead show how to translate the amplitude (3.52) for nucleon scattering into something familiar from Newtonian mechanics: a potential, or force, between the particles.

Let's start by asking a simple question in classical field theory that will turn out to be relevant. Suppose that we have a fixed  $\delta$ -function source for a real scalar field  $\varphi$ , that persists for all time. What is the profile of  $\varphi(\mathbf{x})$ ? To answer this, we must solve the static Klein-Gordon equation,

$$-\nabla^2 \varphi + m^2 \varphi = \delta^{(3)}(\mathbf{x}) \quad (3.64)$$

We can solve this using the Fourier transform,

$$\varphi(\mathbf{x}) = \int \frac{d^3 k}{(2\pi)^3} e^{i \mathbf{k} \cdot \mathbf{x}} \tilde{\varphi}(\mathbf{k}) \quad (3.65)$$

Plugging this into (3.64) tells us that  $(\nabla^2 + m^2) \tilde{\varphi}(\mathbf{k}) = 1$ , giving us the solution

$$\varphi(\mathbf{x}) = \int \frac{d^3 k}{(2\pi)^3} \frac{e^{i \mathbf{k} \cdot \mathbf{x}}}{\mathbf{k}^2 + m^2} \quad (3.66)$$

Let's now do this integral. Changing to polar coordinates, and writing  $\mathbf{k} \cdot \mathbf{x} = kr \cos \vartheta$ , we have

$$\begin{aligned} \varphi(\mathbf{x}) &= \frac{1}{(2\pi)^2} \int_0^\infty dk \frac{k^2}{k^2 + m^2} \frac{2 \sin kr}{kr} \\ &= \frac{1}{(2\pi)^2 r} \int_0^{+\infty} dk \frac{k \sin kr}{k^2 + m^2} \\ &= \frac{1}{2\pi r} \operatorname{Re} \int_{-\infty}^{+\infty} \frac{dk}{2\pi i} \frac{ke^{ikr}}{k^2 + m^2} \end{aligned} \quad (3.67)$$

We compute this last integral by closing the contour in the upper half plane  $k \rightarrow +i\infty$ , picking up the pole at  $k = +im$ . This gives

$$\varphi(\mathbf{x}) = \frac{1}{4\pi r} e^{-mr} \quad (3.68)$$

The field dies off exponentially quickly at distances  $1/m$ , the Compton wavelength of the meson.

Now we understand the profile of the  $\varphi$  field, what does this have to do with the force between  $\psi$  particles? We do very similar calculations to that above in electrostatics where a charged particle acts as a  $\delta$ -function source for the gauge potential:  $-\nabla^2 A_0 = \delta^{(3)}(\mathbf{x})$ , which is solved by  $A_0 = 1/4\pi r$ . The profile for  $A_0$  then acts as the potential energy for another charged (test) particle moving in this background. Can we give the same interpretation to our scalar field? In other words, is there a classical limit of the scalar Yukawa theory where the  $\psi$  particles act as  $\delta$ -function sources for  $\varphi$ , creating the profile (3.68)? And, if so, is this profile then felt as a static potential? The answer is essentially yes, at least in the limit  $M \gg m$ . But the correct way to describe the potential felt by the  $\psi$  particles is not to talk about classical fields at all, but instead work directly with the quantum amplitudes.

Our strategy is to compare the nucleon scattering amplitude (3.52) to the corresponding amplitude in non-relativistic quantum mechanics for two particles interacting through a potential. To make this comparison, we should first take the non-relativistic limit of (3.52). Let's work in the center of mass frame, with  $\not{p} \equiv p \rightarrow_1 = -p \rightarrow_2$  and  $\not{p}' \equiv \not{p}_1' = -\not{p}_2'$ . The non-relativistic limit means  $|p| \ll M$  which, by momentum conservation, ensures that  $|p'| \ll M$ . In fact one can check that, for this particular example, this limit doesn't change the scattering amplitude (3.52): it's given by

$$iA = +ig^2 \frac{1}{(p - p')^2 + m^2} + \frac{1}{(\not{p} + \not{p}')^2 + m^2} \quad (3.69)$$

How do we compare this to scattering in quantum mechanics? Consider two particles, separated by a distance  $\rightarrow r$ , interacting through a potential  $U(\rightarrow r)$ . In non-relativistic quantum mechanics, the amplitude for the particles to scatter from momentum states  $\pm p \rightarrow$  into momentum states  $\pm p' \rightarrow$  can be computed in perturbation theory, using the techniques described in Section 3.1. To leading order, known in this context as the Born approximation, the amplitude is given by

$$\langle p' | U(\rightarrow r) | p \rightarrow \rangle = -i \int d^3r U(\rightarrow r) e^{-i(p - p') \cdot \rightarrow r} \quad (3.70)$$

There's a relative factor of  $(2M)^2$  that arises in comparing the quantum field theory amplitude  $A$  to  $\langle p' | U(\rightarrow r) | p \rightarrow \rangle$ , that can be traced to the relativistic normalization of the states  $|p_1, p_2\rangle$ . (It is also necessary to get the dimensions of the potential to work out correctly). Including this factor, and equating the expressions for the two amplitudes, we get

$$\int d^3r U(\rightarrow r) e^{-i(p - p') \cdot \rightarrow r} = \frac{-\lambda_2}{(p - p')^2 + m^2} \quad (3.71)$$

where we've introduced the dimensionless parameter  $\lambda = g/2M$ . We can trivially invert this to find,

$$U(r) = -\frac{\lambda^2}{(2\pi)^3} \int \frac{d^3 p}{p^2 + m^2} e^{ip \cdot r} \quad (3.72)$$

But this is exactly the integral (3.66) we just did in the classical theory. We have

$$U(r) = \frac{-\lambda^2}{4\pi r} e^{-mr} \quad (3.73)$$

This is the *Yukawa potential*. The force has a range  $1/m$ , the Compton wavelength of the exchanged particle. The minus sign tells us that the potential is attractive.

Notice that quantum field theory has given us an entirely new perspective on the nature of forces between particles. Rather than being a fundamental concept, the force arises from the virtual exchange of other particles, in this case the meson. In Section 6 of these lectures, we will see how the Coulomb force arises from quantum field theory due to the exchange of virtual photons.

We could repeat the calculation for nucleon-anti-nucleon scattering. The amplitude from field theory is given in (3.59). The first term in this expression gives the same result as for nucleon-nucleon scattering *with the same sign*. The second term vanishes in the non-relativistic limit (it is an example of an interaction that doesn't have a simple Newtonian interpretation). There is no longer a factor of  $1/2$  in (3.70), because the incoming/outgoing particles are not identical, so we learn that the potential between a nucleon and anti-nucleon is again given by (3.73). This reveals a key feature of forces arising due to the exchange of scalars: they are universally attractive. Notice that this is different from forces due to the exchange of a spin 1 particle — such as electromagnetism — where the sign flips when we change the charge. However, for forces due to the exchange of a spin 2 particle — i.e. gravity — the force is again universally attractive.

### 3.5.3 $\varphi^4$ Theory

Let's briefly look at the Feynman rules and scattering amplitudes for the interaction Hamiltonian

$$H_{\text{int}} = \frac{\lambda}{4!} \varphi^4 \quad (3.74)$$

The theory now has a single interaction vertex, which comes with a factor of  $(-i\lambda)$ , while the other Feynman rules remain the same. Note that we assign  $(-i\lambda)$  to the

vertex rather than ( $-i\lambda/4!$ ). To see why this is, we can look at  $\varphi\varphi \rightarrow \varphi\varphi$  scattering, which has its lowest contribution at order  $\lambda$ , with the term

$$\frac{-i\lambda}{4!} \langle p_1^r, p_2^r | : \varphi(x)\varphi(x)\varphi(x)\varphi(x) : | p_1, p_2 \rangle \quad (3.75)$$

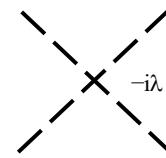
Any one of the fields can do the job of annihilation or creation. This gives  $4!$  different contractions, which cancels the  $1/4!$  sitting out front.

Feynman diagrams in the  $\varphi^4$  theory sometimes come with extra combinatoric factors (typically 2 or 4) which are known as symmetry factors that one must take into account. For more details, see the book by Peskin and Schroeder.

Using the Feynman rules, the scattering amplitude for  $\varphi\varphi \rightarrow \varphi\varphi$  is simply  $iA = -i\lambda$ . Note that it doesn't depend on the angle at which the outgoing particles emerge: in  $\varphi^4$  theory the leading order two-particle scattering occurs with equal probability in all directions. Translating this into a potential between two mesons, we have

$$U(\vec{r}) = \frac{\lambda}{(2m)^2} \int \frac{d^3 p}{(2\pi)^3} e^{+ip \cdot \vec{r}} \Rightarrow \frac{\lambda}{(2m)^2} \delta^{(3)}(\vec{r}) \quad (3.76)$$

So scattering in  $\varphi^4$  theory is due to a  $\delta$ -function potential. The particles don't know what hit them until it's over.



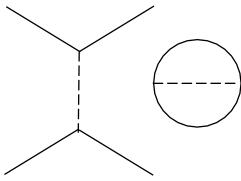
**Figure 16:**

### 3.5.4 Connected Diagrams and Amputated Diagrams

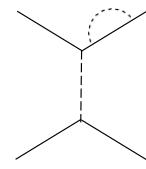
We've seen how one can compute scattering amplitudes by writing down all Feynman diagrams and assigning integrals to them using the Feynman rules. In fact, there are a couple of caveats about what Feynman diagrams you should write down. Both of these caveats are related to the assumption we made earlier that "initial and final states are eigenstates of the free theory" which, as we mentioned at the time, is not strictly accurate. The two caveats which go some way towards ameliorating the problem are the following

- We consider only connected Feynman diagrams, where every part of the diagram is connected to at least one external line. As we shall see shortly, this will be related to the fact that the vacuum  $|0\rangle$  of the free theory is not the true vacuum  $|\Omega\rangle$  of the interacting theory. An example of a diagram that is not connected is shown in Figure 17.

- We do not consider diagrams with loops on external lines, for example the diagram shown in the Figure 18. We will not explain how to take these into account in this course, but you will discuss them next term. They are related to the fact that the one-particle states of the free theory are not the same as the one-particle states of the interacting theory. In particular, correctly dealing with these diagrams will account for the fact that particles in interacting quantum field theories are never alone, but surrounded by a cloud of virtual particles. We will refer to diagrams in which all loops on external legs have been cut-off as “amputated”.



**Figure 17:** A disconnected diagram.



**Figure 18:** An un-amputated diagram

### 3.6 What We Measure: Cross Sections and Decay Rates

So far we've learnt to compute the quantum amplitudes for particles decaying or scattering. As usual in quantum theory, the probabilities for things to happen are the (modulus) square of the quantum amplitudes. In this section we will compute these probabilities, known as decay rates and cross sections. One small subtlety here is that the S-matrix elements  $\langle f | S - 1 | i \rangle$  all come with a factor of  $(2\pi)^4 \delta^{(4)}(p_F - p_i)$ , so we end up with the square of a delta-function. As we will now see, this comes from the fact that we're working in an infinite space.

#### 3.6.1 Fermi's Golden Rule

Let's start with something familiar and recall how to derive Fermi's golden rule from Dyson's formula. For two energy eigenstates  $|m\rangle$  and  $|n\rangle$ , with  $E_m \neq E_n$ , we have to leading order in the interaction,

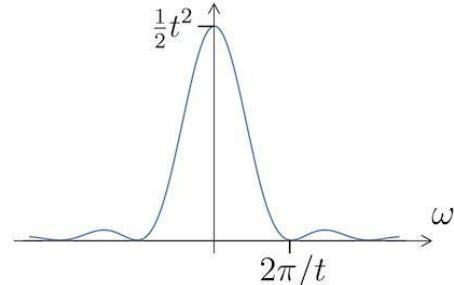
$$\begin{aligned}
 \langle m | U(t) | n \rangle &= -i \langle m | \int_0^t dt' H_I(t') | n \rangle \\
 &= -i \langle m | H_{\text{int}} | n \rangle \int_0^t dt' e^{i\omega t'} \\
 &= -\langle m | H_{\text{int}} | n \rangle \frac{e^{i\omega t} - 1}{\omega}
 \end{aligned} \tag{3.77}$$

where  $\omega = E_m - E_n$ . This gives us the probability for the transition from  $|n\rangle$  to  $|m\rangle$  in time  $t$ , as

$$P_{n \rightarrow m}(t) = |\langle m | U(t) | n \rangle|^2 = 2|\langle m | H_{\text{int}} | n \rangle|^2 \frac{1 - \cos \omega t}{\omega^2} \quad (3.78)$$

The function in brackets is plotted in Figure 19 for fixed  $t$ . We see that in time  $t$ , most transitions happen in a region between energy eigenstates separated by  $\Delta E = 2\pi/t$ . As  $t \rightarrow \infty$ , the function in the figure starts to approach a delta-function. To find the normalization, we can calculate

$$\begin{aligned} & \int_{-\infty}^{+\infty} d\omega \frac{1 - \cos \omega t}{\omega^2} = \pi t \\ \Rightarrow & \frac{1 - \cos \omega t}{\omega^2} \rightarrow \pi t \delta(\omega) \quad \text{as } t \rightarrow \infty \end{aligned}$$



**Figure 19:**

Consider now a transition to a cluster of states with density  $\rho(E)$ . In the limit  $t \rightarrow \infty$ , we get the transition probability

$$\begin{aligned} P_{n \rightarrow m} &= \int dE_m \rho(E_m) 2|\langle m | H_{\text{int}} | n \rangle|^2 \frac{1 - \cos \omega t}{\omega^2} \\ &\rightarrow 2\pi |\langle m | H_{\text{int}} | n \rangle|^2 \rho(E_n) \end{aligned} \quad (3.79)$$

which gives a constant probability for the transition per unit time for states around the same energy  $E_n \sim E_m = E$ .

$$\dot{P}_{n \rightarrow m} = 2\pi |\langle m | H_{\text{int}} | n \rangle|^2 \rho(E) \quad (3.80)$$

This is Fermi's Golden Rule.

In the above derivation, we were fairly careful with taking the limit as  $t \rightarrow \infty$ . Suppose we were a little sloppier, and first chose to compute the amplitude for the state  $|n\rangle$  at  $t \rightarrow -\infty$  to transition to the state  $|m\rangle$  at  $t \rightarrow +\infty$ . Then we get

$$-i \langle m | \int_{t=-\infty}^{t=+\infty} H_l(t) | n \rangle = -i \langle m | H_{\text{int}} | n \rangle 2\pi \delta(\omega) \quad (3.81)$$

Now when squaring the amplitude to get the probability, we run into the problem of the square of the delta-function:  $P_{n \rightarrow m} = |\langle m | H_{\text{int}} | n \rangle|^2 (2\pi)^2 \delta(\omega)^2$ . Tracking through the previous computations, we realize that the extra infinity is coming because  $P_{m \rightarrow n}$

is the probability for the transition to happen in infinite time  $t \rightarrow \infty$ . We can write the delta-functions as

$$(2\pi)^2 \delta(\omega)^2 = (2\pi) \delta(\omega) T \quad (3.82)$$

where  $T$  is shorthand for  $t \rightarrow \infty$  (we used a very similar trick when looking at the vacuum energy in (2.25)). We now divide out by this power of  $T$  to get the transition probability per unit time,

$$P_{n \rightarrow m} = 2\pi | \langle m | H_{\text{int}} | n \rangle |^2 \delta(\omega) \quad (3.83)$$

which, after integrating over the density of final states, gives us back Fermi's Golden rule. The reason that we've stressed this point is because, in our field theory calculations, we've computed the amplitudes in the same way as (3.81), and the square of the  $\delta^{(4)}$ -functions will just be re-interpreted as spacetime volume factors.

### 3.6.2 Decay Rates

Let's now look at the probability for a single particle  $|i\rangle$  of momentum  $p_i$  (I=initial) to decay into some number of particles  $|f\rangle$  with momentum  $p_f$  and total momentum  $p_F = \sum_i p_i$ . This is given by

$$P = \frac{|\langle f | S | i \rangle|^2}{\langle f | f \rangle \langle i | i \rangle} \quad (3.84)$$

Our states obey the relativistic normalization formula (2.65),

$$\langle i | i \rangle = (2\pi)^3 2E_{p \rightarrow} \delta^{(3)}(0) = 2E_{p \rightarrow} V \quad (3.85)$$

where we have replaced  $\delta^{(3)}(0)$  by the volume of 3-space. Similarly,

$$\langle f | f \rangle = \sum_{\text{final states}} 2E_{p \rightarrow} V \quad (3.86)$$

If we place our initial particle at rest, so  $p_{i \rightarrow} = 0$  and  $E_{p \rightarrow} = m$ , we get the probability for decay

$$P = \frac{|A_{fi}|^2}{2mV} (2\pi)^4 \delta^{(4)}(p_i - p_F) VT \sum_{\text{final states}} \frac{1}{2E_{p \rightarrow} V} \quad (3.87)$$

where, as in the second derivation of Fermi's Golden Rule, we've exchanged one of the delta-functions for the volume of spacetime:  $(2\pi)^4 \delta^{(4)}(0) = V T$ . The amplitudes  $A_{fi}$  are, of course, exactly what we've been computing. (For example, in (3.30), we saw

that  $A = -g$  for a single meson decaying into two nucleons). We can now divide out by  $T$  to get the transition function per unit time. But we still have to worry about summing over all final states. There are two steps: the first is to integrate over all possible momenta of the final particles:  $\int d^3 p_i / (2\pi)^3$ . The factors of spatial volume  $V$  in this measure cancel those in (3.87), while the factors of  $1/2E_{p\rightarrow}$ , in (3.87) conspire to produce the Lorentz invariant measure for 3-momentum integrals. The result is an expression for the density of final states given by the Lorentz invariant measure

$$d\Pi = (2\pi)^4 \delta^{(4)}(p_F - p_i) \sum_{\text{final states}} \frac{d^3 p_i}{(2\pi)^3} \frac{1}{2E_{p\rightarrow}} \quad (3.88)$$

The second step is to sum over all final states with different numbers (and possibly types) of particles. This gives us our final expression for the decay probability per unit time,  $\Gamma = P$ .

$$\Gamma = \frac{1}{2m} \sum_{\text{final states}} |A_{fi}|^2 d\Pi \quad (3.89)$$

$\Gamma$  is called the width of the particle. It is equal to the reciprocal of the half-life  $\tau = 1/\Gamma$ .

### 3.6.3 Cross Sections

Collide two beams of particles. Sometimes the particles will hit and bounce off each other; sometimes they will pass right through. The fraction of the time that they collide is called the *cross section* and is denoted by  $\sigma$ . If the incoming flux  $F$  is defined to be the number of incoming particles per area per unit time, then the total number of scattering events  $N$  per unit time is given by,

$$N = F\sigma \quad (3.90)$$

We would like to calculate  $\sigma$  from quantum field theory. In fact, we can calculate a more sensitive quantity  $d\sigma$  known as the *differential cross section* which is the probability for a given scattering process to occur in the solid angle  $(\vartheta, \phi)$ . More precisely

$$d\sigma = \frac{\text{Differential Probability}}{\text{Unit Time} \times \text{Unit Flux}} = \frac{1}{4E_1 E_2 V F} |A_{fi}|^2 d\Pi \quad (3.91)$$

where we've used the expression for probability per unit time that we computed in the previous subsection.  $E_1$  and  $E_2$  are the energies of the incoming particles. We now need an expression for the unit flux. For simplicity, let's sit in the center of mass frame of the collision. We've been considering just a single particle per spatial volume  $V$ ,

meaning that the flux is given in terms of the 3-velocities  $\rightarrow v_i$  as  $F = |\rightarrow v_1 - \rightarrow v_2|/V$ . This then gives,

$$d\sigma = \frac{1}{4E_1 E_2} \frac{1}{|\rightarrow v_1 - \rightarrow v_2|} |A_{fi}|^2 d\Pi \quad (3.92)$$

If you want to write this in terms of momentum, then recall from your course on special relativity that the 3-velocities  $\rightarrow v_i$  are related to the momenta by  $\rightarrow v = \vec{p}/m\sqrt{1 - v^2} = \vec{p}/p^0$ .

Equation (3.92) is our final expression relating the S-matrix to the differential cross section. You may now take your favorite scattering amplitude, and compute the probability for particles to fly out at your favorite angles. This will involve doing the integral over the phase space of final states, with measure  $d\Pi$ . Notice that different scattering amplitudes have different momentum dependence and will result in different angular dependence in scattering amplitudes. For example, in  $\varphi^4$  theory the amplitude for tree level scattering was simply  $A = -\lambda$ . This results in isotropic scattering. In contrast, for nucleon-nucleon scattering we have schematically  $A \sim (t - m^2)^{-1} + (u - m^2)^{-1}$ . This gives rise to angular dependence in the differential cross-section, which follows from the fact that, for example,  $t = -2|\vec{p}|^2(1 - \cos\vartheta)$ , where  $\vartheta$  is the angle between the incoming and outgoing particles.

### 3.7 Green's Functions

So far we've learnt to compute scattering amplitudes. These are nice and physical (well – they're directly related to cross-sections and decay rates which are physical) but there are many questions we want to ask in quantum field theory that aren't directly related to scattering experiments. For example, we might want to compute the viscosity of the quark gluon plasma, or the optical conductivity in a tentative model of strange metals, or figure out the non-Gaussianity of density perturbations arising in the CMB from novel models of inflation. All of these questions are answered in the framework of quantum field theory by computing elementary objects known as *correlation functions*. In this section we will briefly define correlation functions, explain how to compute them using Feynman diagrams, and then relate them back to scattering amplitudes. We'll leave the relationship to other physical phenomena to other courses.

We'll denote the true vacuum of the interacting theory as  $|\Omega\rangle$ . We'll normalize  $H$  such that

$$H |\Omega\rangle = 0 \quad (3.93)$$

and  $\langle \Omega | \Omega \rangle = 1$ . Note that this is different from the state we've called  $|0\rangle$  which is the vacuum of the free theory and satisfies  $H_0|0\rangle = 0$ . Define

$$G^{(n)}(x_1, \dots, x_n) = \langle \Omega | T \varphi_H(x_1) \dots \varphi_H(x_n) | \Omega \rangle \quad (3.94)$$

where  $\varphi_H$  is  $\varphi$  in the Heisenberg picture of the full theory, rather than the interaction picture that we've been dealing with so far. The  $G^{(n)}$  are called correlation functions, or *Green's functions*. There are a number of different ways of looking at these objects which tie together nicely. Let's start by asking how to compute  $G^{(n)}$  using Feynman diagrams. We prove the following result

**Claim:** We use the notation  $\varphi_1 = \varphi(x_1)$ , and write  $\varphi_{1H}$  to denote the field in the Heisenberg picture, and  $\varphi_{1I}$  to denote the field in the interaction picture. Then

$$G^{(n)}(x_1, \dots, x_n) = \langle \Omega | T \varphi_{1H} \dots \varphi_{nH} | \Omega \rangle = \frac{\langle 0 | T \varphi_{1I} \dots \varphi_{nI} S | 0 \rangle}{\langle 0 | S | 0 \rangle} \quad (3.95)$$

where the operators on the right-hand side are evaluated on  $|0\rangle$ , the vacuum of the free theory.

**Proof:** Take  $t_1 > t_2 > \dots > t_n$ . Then we can drop the  $T$  and write the numerator of the right-hand side as

$$\langle 0 | U_I(+\infty, t_1) \varphi_{1I} U(t_1, t_2) \varphi_{2I} \dots \varphi_{nI} U_I(t_n, -\infty) | 0 \rangle$$

We'll use the factors of  $U_I(t_k, t_{k+1}) = T \exp(-i \int_{t_k}^{t_{k+1}} H)$  to convert each of the  $\varphi_{1I}$  into  $\varphi_H$  and we choose operators in the two pictures to be equal at some arbitrary time  $t_0$ . Then we can write

$$\begin{aligned} \langle 0 | U_I(+\infty, t_1) \varphi_{1I} U(t_1, t_2) \varphi_{2I} \dots \varphi_{nI} U_I(t_n, -\infty) | 0 \rangle \\ = \langle 0 | U_I(+\infty, t_0) \varphi_{1H} \dots \varphi_{nH} U_I(t_0, -\infty) | 0 \rangle \end{aligned}$$

Now let's deal with the two remaining  $U(t_0, \pm\infty)$  at either end of the string of operators. Consider an arbitrary state  $|\Psi\rangle$  and look at

$$\langle \Psi | U(t, -\infty) | 0 \rangle = \langle \Psi | U(t, -\infty) | 0 \rangle \quad (3.96)$$

where  $U(t, -\infty)$  is the Schrödinger evolution operator, and the equality above follows because  $H_0|0\rangle = 0$ . Now insert a complete set of states, which we take to be energy eigenstates of  $H = H_0 + H_{\text{int}}$ ,

$$\begin{aligned} \langle \Psi | U(t, -\infty) | 0 \rangle &= \langle \Psi | U(t, -\infty) | \Omega \rangle \langle \Omega | + \sum_{n \neq 0}^{\#} | n \rangle \langle n | | 0 \rangle \\ &= \langle \Psi | \Omega \rangle \langle \Omega | 0 \rangle + \lim_{t' \rightarrow -\infty} \sum_{n \neq 0}^{\#} e^{iE_n(t' - t)} \langle \Psi | n \rangle \langle n | 0 \rangle \quad (3.97) \end{aligned}$$

But the last term vanishes. This follows from the Riemann-Lebesgue lemma which says that for any well-behaved function

$$\lim_{\mu \rightarrow \infty} \int_a^b dx f(x) e^{i\mu x} = 0 \quad (3.98)$$

Why is this relevant? The point is that the  $\sum_n$  in (3.97) is really an integral  $\int dn$ , because all states are part of a continuum due to the momentum. (There is a caveat here: we want the vacuum  $|\Omega\rangle$  to be special, so that it sits on its own, away from the continuum of the integral. This means that we must be working in a theory with a mass gap – i.e. with no massless particles). So the Riemann-Lebesgue lemma gives us

$$\lim_{t' \rightarrow -\infty} \langle \Psi | U(t, t') | 0 \rangle = \langle \Psi | \Omega \rangle \langle \Omega | 0 \rangle \quad (3.99)$$

(Notice that to derive this result, Peskin and Schroeder instead send  $t \rightarrow -\infty$  in a slightly imaginary direction, which also does the job). We now apply the formula (3.99), to the top and bottom of the right-hand side of (3.95) to find

$$\frac{\langle 0 | \Omega \rangle \langle \Omega | T\varphi_1 \dots \varphi_n | \Omega \rangle \langle \Omega | 0 \rangle}{\langle 0 | \Omega \rangle \langle \Omega | \Omega \rangle \langle \Omega | 0 \rangle} \quad (3.100)$$

which, using the normalization  $\langle \Omega | \Omega \rangle = 1$ , gives us the left-hand side, completing the proof.

### 3.7.1 Connected Diagrams and Vacuum Bubbles

We're getting closer to our goal of computing the Green's functions  $G^{(n)}$  since we can compute both  $\langle 0 | T\varphi_1(x_1) \dots \varphi_n(x_n) S | 0 \rangle$  and  $\langle 0 | S | 0 \rangle$  using the same methods we developed for S-matrix elements; namely Dyson's formula and Wick's theorem or, alternatively, Feynman diagrams. But what about dividing one by the other? What's that all about? In fact, it has a simple interpretation. For the following discussion, we will work in  $\varphi^4$  theory. Since there is no ambiguity in the different types of lines in Feynman diagrams, we will represent the  $\varphi$  particles as solid lines, rather than the dashed lines that we used previously. Then we have the diagrammatic expansion for  $\langle 0 | S | 0 \rangle$ .

$$\langle 0 | S | 0 \rangle = 1 + \text{---} + \left( \text{---} + \text{---} + \text{---} \right) + \dots \quad (3.101)$$

These diagrams are called vacuum bubbles. The combinatoric factors (as well as the symmetry factors) associated with each diagram are such that the whole series sums

to an exponential,

$$\langle 0 | S | 0 \rangle = \exp \left( \text{ } \textcircled{8} \text{ } + \text{ } \textcircled{\textcircled{8}} \text{ } + \text{ } \textcircled{\textcircled{\textcircled{8}}} \text{ } + \text{ } \dots \right) \quad (3.102)$$

So the amplitude for the vacuum of the free theory to evolve into itself is  $\langle 0 | S | 0 \rangle = \exp(\text{all distinct vacuum bubbles})$ . A similar combinatoric simplification occurs for generic correlation functions. Remarkably, the vacuum diagrams all add up to give the same exponential. With a little thought one can show that

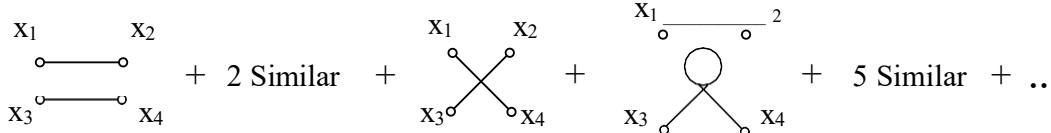
$$\langle 0 | T\varphi_1 \dots \varphi_n S | 0 \rangle = \sum_{\text{connected diagrams}} \langle 0 | S | 0 \rangle \quad (3.103)$$

where “connected” means that every part of the diagram is connected to at least one of the external legs. The upshot of all this is that dividing by  $\langle 0 | S | 0 \rangle$  has a very nice interpretation in terms of Feynman diagrams: we need only consider the connected Feynman diagrams, and don’t have to worry about the vacuum bubbles. Combining this with (3.95), we learn that the Green’s functions  $G^{(n)}(x_1 \dots, x_n)$  can be calculated by summing over all connected Feynman diagrams,

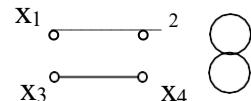
$$\langle \Omega | T \varphi_H(x_1) \dots \varphi_H(x_n) | \Omega \rangle = \sum \text{Connected Feynman Graphs} \quad (3.104)$$

An Example: The Four-Point Correlator:  $\langle \Omega | T\varphi_H(x_1) \dots \varphi_H(x_4) | \Omega \rangle$

As a simple example, let's look at the four-point correlation function in  $\phi^4$  theory. The sum of connected Feynman diagrams is given by,



All of these are connected diagrams, even though they don't look that connected! The point is that a connected diagram is defined by the requirement that every line is joined to an external leg. An example of a diagram that is not connected is shown in the figure. As we have seen, such diagrams are taken care of in shifting the vacuum from  $|0\rangle$  to  $|\Omega\rangle$ .



**Figure 20:**

## Feynman Rules

The Feynman diagrams that we need to calculate for the Green's functions depend on  $x_1, \dots, x_n$ . This is rather different than the Feynman diagrams that we calculated for

the S-matrix elements, where we were working primarily with momentum eigenstates, and ended up integrating over all of space. However, it's rather simple to adapt the Feynman rules that we had earlier in momentum space to compute  $G^{(n)}(x_1, \dots, x_n)$ . For  $\varphi^4$  theory, we have

- Draw  $n$  external points  $x_1, \dots, x_n$ , connected by the usual propagators and vertices. Assign a spacetime position  $y$  to the end of each line.
- For each line  $x \xrightarrow{ } y$  from  $x$  to  $y$  write down a factor of the Feynman propagator  $\Delta_F(x - y)$ .
- For each vertex  at position  $y$ , write down a factor of  $-i\lambda \int d^4y$ .

### 3.7.2 From Green's Functions to S-Matrices

Having described how to compute correlation functions using Feynman diagrams, let's now relate them back to the S-matrix elements that we already calculated. The first step is to perform the Fourier transform,

$$\tilde{G}^{(n)}(p_1, \dots, p_n) = \int_{i=1}^n d^4x_i e^{-ip_i \cdot x_i} G^{(n)}(x_1, \dots, x_n) \quad (3.105)$$

These are very closely related to the S-matrix elements that we've computed above. The difference is that the Feynman rules for  $G^{(n)}(x_1, \dots, x_n)$ , effectively include propagators  $\Delta_F$  for the external legs, as well as the internal legs. A related fact is that the 4-momenta assigned to the external legs is arbitrary: they are not on-shell. Both of these problems are easily remedied to allow us to return to the S-matrix elements: we need to simply cancel off the propagators on the external legs, and place their momentum back on shell. We have

$$\langle p_{r_1}, \dots, p_{r_n} | S - 1 | p_1, \dots, p_n \rangle = (-i)^{n+n'} \sum_{i=1}^{n'} (p_{r_i} - m_i) \sum_{j=1}^{n-n'} (p_{r_j} - m_j) \times \tilde{G}^{(n+n')}( -p_{r_1}, \dots, -p_{r_n}, p_1, \dots, p_n ) \quad (3.106)$$

Each of the factors  $(p^2 - m^2)$  vanishes once the momenta are placed on-shell. This means that we only get a non-zero answer for diagrams contributing to  $G^{(n)}(x_1, \dots, x_n)$  which have propagators for each external leg.

So what's the point of all of this? We've understood that ignoring the unconnected diagrams is related to shifting to the true vacuum  $|\Omega\rangle$ . But other than that, introducing the Green's functions seems like a lot of bother for little reward. The important point

is that this provides a framework in which to deal with the true particle states in the interacting theory through renormalization. Indeed, the formula (3.106), suitably interpreted, remains true even in the interacting theory, taking into account the swarm of virtual particles surrounding asymptotic states. This is the correct way to consider scattering. In this context, (3.106) is known as the LSZ reduction formula. You will derive it properly next term.

## 4. The Dirac Equation

"A great deal more was hidden in the Dirac equation than the author had expected when he wrote it down in 1928. Dirac himself remarked in one of his talks that his equation was more intelligent than its author. It should be added, however, that it was Dirac who found most of the additional insights."

*Weisskopf on Dirac*

So far we've only discussed scalar fields such that under a Lorentz transformation  $x^\mu \rightarrow (x^r)^\mu = \Lambda^\mu_v x^v$ , the field transforms as

$$\varphi(x) \rightarrow \varphi^r(x) = \varphi(\Lambda^{-1}x) \quad (4.1)$$

We have seen that quantization of such fields gives rise to spin 0 particles. But most particles in Nature have an intrinsic angular momentum, or spin. These arise naturally in field theory by considering fields which themselves transform non-trivially under the Lorentz group. In this section we will describe the Dirac equation, whose quantization gives rise to fermionic spin 1/2 particles. To motivate the Dirac equation, we will start by studying the appropriate representation of the Lorentz group.

A familiar example of a field which transforms non-trivially under the Lorentz group is the vector field  $A_\mu(x)$  of electromagnetism,

$$A^\mu(x) \rightarrow \Lambda^\mu_v A^v(\Lambda^{-1}x) \quad (4.2)$$

We'll deal with this in Section 6. (It comes with its own problems!). In general, a field can transform as

$$\varphi^a(x) \rightarrow D[\Lambda]^a_b \varphi^b(\Lambda^{-1}x) \quad (4.3)$$

where the matrices  $D[\Lambda]$  form a *representation* of the Lorentz group, meaning that

$$D[\Lambda_1]D[\Lambda_2] = D[\Lambda_1\Lambda_2] \quad (4.4)$$

and  $D[\Lambda^{-1}] = D[\Lambda]^{-1}$  and  $D[1] = 1$ . How do we find the different representations? Typically, we look at infinitesimal transformations of the Lorentz group and study the resulting Lie algebra. If we write,

$$\Lambda^\mu_v = \delta^\mu_v + \omega^\mu_v \quad (4.5)$$

for infinitesimal  $\omega$ , then the condition for a Lorentz transformation  $\Lambda^\mu_\sigma \Lambda^\nu_\rho \eta^{\sigma\rho} = \eta^{\mu\nu}$

becomes the requirement that  $\omega$  is anti-symmetric:

$$\omega^{\mu\nu} + \omega^{\nu\mu} = 0 \quad (4.6)$$

Note that an antisymmetric  $4 \times 4$  matrix has  $4 \times 3/2 = 6$  independent components, which agrees with the 6 transformations of the Lorentz group: 3 rotations and 3 boosts. It's going to be useful to introduce a basis of these six  $4 \times 4$  anti-symmetric matrices. We could call them  $(M^A)^{\mu\nu}$ , with  $A = 1, \dots, 6$ . But in fact it's better for us (although initially a little confusing) to replace the single index  $A$  with a pair of antisymmetric indices  $[\rho\sigma]$ , where  $\rho, \sigma = 0, \dots, 3$ , so we call our matrices  $(M^{\rho\sigma})^\mu_v$ . The antisymmetry on the  $\rho$  and  $\sigma$  indices means that, for example,  $M^{01} = -M^{10}$ , etc, so that  $\rho$  and  $\sigma$  again label six different matrices. Of course, the matrices are also antisymmetric on the  $\mu\nu$  indices because they are, after all, antisymmetric matrices. With this notation in place, we can write a basis of six  $4 \times 4$  antisymmetric matrices as

$$(M^{\rho\sigma})^{\mu\nu} = \eta^{\rho\mu} \eta^{\sigma\nu} - \eta^{\sigma\mu} \eta^{\rho\nu} \quad (4.7)$$

where the indices  $\mu$  and  $\nu$  are those of the  $4 \times 4$  matrix, while  $\rho$  and  $\sigma$  denote which basis element we're dealing with. If we use these matrices for anything practical (for example, if we want to multiply them together, or act on some field) we will typically need to lower one index, so we have

$$(M^{\rho\sigma})^\mu_v = \eta^{\rho\mu} \delta^\sigma_v - \eta^{\sigma\mu} \delta^\rho_v \quad (4.8)$$

Since we lowered the index with the Minkowski metric, we pick up various minus signs which means that when written in this form, the matrices are no longer necessarily antisymmetric. Two examples of these basis matrices are,

$$(M^{01})^\mu_v = \begin{matrix} & & & \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \quad \text{and} \quad (M^{12})^\mu_v = \begin{matrix} & & & \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \quad (4.9)$$

The first,  $M^{01}$ , generates boosts in the  $x^1$  direction. It is real and symmetric. The second,  $M^{12}$ , generates rotations in the  $(x^1, x^2)$ -plane. It is real and antisymmetric. We can now write any  $\omega^\mu_v$  as a linear combination of the  $M^{\rho\sigma}$ ,

$$\omega^\mu_v = \frac{1}{2} \Omega_{\rho\sigma} (M^{\rho\sigma})^\mu_v \quad (4.10)$$

where  $\Omega_{\rho\sigma}$  are just six numbers (again antisymmetric in the indices) that tell us what Lorentz transformation we're doing. The six basis matrices  $M^{\rho\sigma}$  are called the *generators* of the Lorentz transformations. The generators obey the Lorentz Lie algebra relations,

$$[M^{\rho\sigma}, M^{\tau\nu}] = \eta^{\sigma\tau} M^{\rho\nu} - \eta^{\rho\tau} M^{\sigma\nu} + \eta^{\rho\nu} M^{\sigma\tau} - \eta^{\sigma\nu} M^{\rho\tau} \quad (4.11)$$

where we have suppressed the matrix indices. A finite Lorentz transformation can then be expressed as the exponential

$$\Lambda = \exp \frac{1}{2} \Omega_{\rho\sigma} M^{\rho\sigma} \quad (4.12)$$

Let me stress again what each of these objects are: the  $M^{\rho\sigma}$  are six  $4 \times 4$  basis elements of the Lorentz Lie algebra; the  $\Omega_{\rho\sigma}$  are six numbers telling us what kind of Lorentz transformation we're doing (for example, they say things like rotate by  $\vartheta = \pi/7$  about the  $x^3$ -direction and run at speed  $v = 0.2$  in the  $x^1$  direction).

#### 4.1 The Spinor Representation

We're interested in finding other matrices which satisfy the Lorentz algebra commutation relations (4.11). We will construct the spinor representation. To do this, we start by defining something which, at first sight, has nothing to do with the Lorentz group. It is the *Clifford algebra*,

$$\{\gamma^\mu, \gamma^\nu\} \equiv \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2\eta^{\mu\nu}1 \quad (4.13)$$

where  $\gamma^\mu$ , with  $\mu = 0, 1, 2, 3$ , are a set of four matrices and the 1 on the right-hand side denotes the unit matrix. This means that we must find four matrices such that

$$\gamma^\mu \gamma^\nu = -\gamma^\nu \gamma^\mu \quad \text{when } \mu \neq \nu \quad (4.14)$$

and

$$(\gamma^0)^2 = 1 , \quad (\gamma^i)^2 = -1 \quad i = 1, 2, 3 \quad (4.15)$$

It's not hard to convince yourself that there are no representations of the Clifford algebra using  $2 \times 2$  or  $3 \times 3$  matrices. The simplest representation of the Clifford algebra is in terms of  $4 \times 4$  matrices. There are many such examples of  $4 \times 4$  matrices which obey (4.13). For example, we may take

$$\gamma^0 = \begin{matrix} & & 1 & \\ 0 & 1 & & \\ & & & 1 \\ 1 & 0 & & \end{matrix}, \quad \gamma^i = \begin{matrix} & & & 1 \\ & & 0 & \sigma^i \\ & -\sigma^i & 0 & \\ & & & 1 \end{matrix} \quad (4.16)$$

where each element is itself a  $2 \times 2$  matrix, with the  $\sigma^i$  the Pauli matrices

$$\sigma^1 = \begin{matrix} & & 1 & \\ 0 & 1 & & \\ & & & 1 \\ 1 & 0 & & \end{matrix}, \quad \sigma^2 = \begin{matrix} & & & 1 \\ & & 0 & -i \\ & i & 0 & \\ & & & 1 \end{matrix}, \quad \sigma^3 = \begin{matrix} & & & 1 \\ & & 0 & \\ & -1 & 0 & \\ & & & 1 \end{matrix} \quad (4.17)$$

which themselves satisfy  $\{\sigma^i, \sigma^j\} = 2\delta^{ij}$ .

One can construct many other representations of the Clifford algebra by taking  $V \gamma^\mu V^{-1}$  for any invertible matrix  $V$ . However, up to this equivalence, it turns out that there is a unique irreducible representation of the Clifford algebra. The matrices (4.16) provide one example, known as the *Weyl* or *chiral representation* (for reasons that will soon become clear). We will soon restrict ourselves further, and consider only representations of the Clifford algebra that are related to the chiral representation by a unitary transformation  $V$ .

So what does the Clifford algebra have to do with the Lorentz group? Consider the commutator of two  $\gamma^\mu$ ,

$$S^{\rho\sigma} = \frac{1}{4} [\gamma^\rho, \gamma^\sigma] = \begin{array}{ccccc} & 0 & & \rho = \sigma \\ 4 & & & & \\ & \frac{1}{2} \gamma^\rho \gamma^\sigma & & \rho \neq \sigma & 2 \\ & & & & 2 \end{array} = \frac{1}{2} \gamma^\rho \gamma^\sigma - \frac{1}{2} \eta^{\rho\sigma} \quad (4.18)$$

Let's see what properties these matrices have:

**Claim 4.1:**  $[S^{\mu\nu}, \gamma^\rho] = \gamma^\mu \eta^{\nu\rho} - \gamma^\nu \eta^{\rho\mu}$

**Proof:** When  $\mu \neq \nu$  we have

$$\begin{aligned} [S^{\mu\nu}, \gamma^\rho] &= \frac{1}{2} [\gamma^\mu \gamma^\nu, \gamma^\rho] \\ &= \frac{1}{2} \gamma^\mu \gamma^\nu \gamma^\rho - \frac{1}{2} \gamma^\rho \gamma^\mu \gamma^\nu \\ &= \frac{1}{2} \gamma^\mu \{ \gamma^\nu, \gamma^\rho \} - \frac{1}{2} \gamma^\mu \gamma^\rho \gamma^\nu - \frac{1}{2} \{ \gamma^\rho, \gamma^\mu \} \gamma^\nu + \frac{1}{2} \gamma^\mu \gamma^\rho \gamma^\nu \\ &= \gamma^\mu \eta^{\nu\rho} - \gamma^\nu \eta^{\rho\mu} \end{aligned}$$

**Claim 4.2:** The matrices  $S^{\mu\nu}$  form a representation of the Lorentz algebra (4.11), meaning

$$[S^{\mu\nu}, S^{\rho\sigma}] = \eta^{\nu\rho} S^{\mu\sigma} - \eta^{\mu\rho} S^{\nu\sigma} + \eta^{\mu\sigma} S^{\nu\rho} - \eta^{\nu\sigma} S^{\mu\rho} \quad (4.19)$$

**Proof:** Taking  $\rho \neq \sigma$ , and using Claim 4.1 above, we have

$$\begin{aligned} [S^{\mu\nu}, S^{\rho\sigma}] &= \frac{1}{2} [S^{\mu\nu}, \gamma^\rho \gamma^\sigma] \\ &= \frac{1}{2} [S^{\mu\nu}, \gamma^\rho] \gamma^\sigma + \frac{1}{2} \gamma^\rho [S^{\mu\nu}, \gamma^\sigma] \\ &= \frac{1}{2} \gamma^\mu \gamma^\nu \eta^{\nu\rho} - \frac{1}{2} \gamma^\nu \gamma^\mu \eta^{\rho\mu} + \frac{1}{2} \gamma^\rho \gamma^\mu \eta^{\nu\sigma} - \frac{1}{2} \gamma^\mu \gamma^\nu \eta^{\sigma\mu} \end{aligned} \quad (4.20)$$

Now using the expression (4.18) to write  $\gamma^\mu \gamma^\sigma = 2S^{\mu\sigma} + \eta^{\mu\sigma}$ , we have

$$[S^{\mu\nu}, S^{\rho\sigma}] = S^{\mu\sigma} \eta^{\nu\rho} - S^{\nu\sigma} \eta^{\rho\mu} + S^{\rho\mu} \eta^{\nu\sigma} - S^{\rho\nu} \eta^{\sigma\mu} \quad (4.21)$$

which is our desired expression.

### 4.1.1 Spinors

The  $S^{\mu\nu}$  are  $4 \times 4$  matrices, because the  $\gamma^\mu$  are  $4 \times 4$  matrices. So far we haven't given an index name to the rows and columns of these matrices: we're going to call them  $\alpha, \beta = 1, 2, 3, 4$ .

We need a field for the matrices  $(S^{\mu\nu})^\alpha{}_\beta$  to act upon. We introduce the Dirac *spinor* field  $\psi^\alpha(x)$ , an object with four complex components labelled by  $\alpha = 1, 2, 3, 4$ . Under Lorentz transformations, we have

$$\psi^\alpha(x) \rightarrow S[\Lambda]^\alpha{}_\beta \psi^\beta(\Lambda^{-1}x) \quad (4.22)$$

where

$$\Lambda = \exp \frac{1}{2} \Omega_{\rho\sigma} M^{\rho\sigma} \quad (4.23)$$

$$S[\Lambda] = \exp \frac{1}{2} \Omega_{\rho\sigma} S^{\rho\sigma} \quad (4.24)$$

Although the basis of generators  $M^{\rho\sigma}$  and  $S^{\rho\sigma}$  are different, we use the same six numbers  $\Omega_{\rho\sigma}$  in both  $\Lambda$  and  $S[\Lambda]$ : this ensures that we're doing the same Lorentz transformation on  $x$  and  $\psi$ . Note that we denote both the generator  $S^{\rho\sigma}$  and the full Lorentz transformation  $S[\Lambda]$  as "S". To avoid confusion, the latter will always come with the square brackets  $[\Lambda]$ .

Both  $\Lambda$  and  $S[\Lambda]$  are  $4 \times 4$  matrices. So how can we be sure that the spinor representation is something new, and isn't equivalent to the familiar representation  $\Lambda^\mu{}_\nu$ ? To see that the two representations are truly different, let's look at some specific transformations.

### Rotations

$$S^{ij} = \frac{1}{2} \begin{matrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{matrix} \begin{matrix} 0 & \sigma^j \\ -\sigma^j & 0 \end{matrix} = \frac{i}{2} \epsilon^{ijk} \begin{matrix} \sigma^k & 0 \\ 0 & \sigma^k \end{matrix} \quad (\text{for } i \neq j) \quad (4.25)$$

If we write the rotation parameters as  $\Omega_{ij} = -\epsilon_{ijk}\phi^k$  (meaning  $\Omega_{12} = -\phi^3$ , etc) then the rotation matrix becomes

$$S[\Lambda] = \exp \frac{1}{2} \Omega_{\rho\sigma} S^{\rho\sigma} = \begin{matrix} e^{i\phi^3 \cdot \sigma/2} & 0 \\ 0 & e^{i\phi^3 \cdot \sigma/2} \end{matrix} \quad (4.26)$$

where we need to remember that  $\Omega_{12} = -\Omega_{21} = -\phi^3$  when following factors of 2. Consider now a rotation by  $2\pi$  about, say, the  $x^3$ -axis. This is achieved by  $\phi = (0, 0, 2\pi)$ ,

and the spinor rotation matrix becomes,

$$S[\Lambda] = \begin{matrix} e^{+i\pi\sigma^3} & 0 \\ 0 & e^{-i\pi\sigma^3} \end{matrix} = -1 \quad (4.27)$$

Therefore under a  $2\pi$  rotation

$$\psi^\alpha(x) \rightarrow -\psi^\alpha(x) \quad (4.28)$$

which is definitely not what happens to a vector! To check that we haven't been cheating with factors of 2, let's see how a vector would transform under a rotation by  $\phi = (0, 0, \phi^3)$ . We have

$$\Lambda = \exp \frac{1}{2} \Omega_{\rho\sigma} M^{\rho\sigma} = \exp \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \phi^3 & 0 \\ 0 & -\phi^3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} \quad (4.29)$$

So when we rotate a vector by  $\phi^3 = 2\pi$ , we learn that  $\Lambda = 1$  as you would expect. So  $S[\Lambda]$  is definitely a different representation from the familiar vector representation  $\Lambda^\mu_v$ .

### Boosts

$$S^{0i} = \begin{matrix} 1 & 0 & 1 \\ 2 & 1 & 0 \end{matrix} \begin{matrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{matrix} = \begin{matrix} 1 & -\sigma^i & 0 \\ 2 & 0 & \sigma^i \end{matrix} \quad (4.30)$$

Writing the boost parameter as  $\Omega_{i0} = -\Omega_{0i} = \chi_i$ , we have

$$S[\Lambda] = \begin{matrix} e^{i\chi^i \cdot \sigma/2} & 0 \\ 0 & e^{-i\chi^i \cdot \sigma/2} \end{matrix} \quad (4.31)$$

### **Representations of the Lorentz Group are not Unitary**

Note that for rotations given in (4.26),  $S[\Lambda]$  is unitary, satisfying  $S[\Lambda]^\dagger S[\Lambda] = 1$ . But for boosts given in (4.31),  $S[\Lambda]$  is not unitary. In fact, there are *no* finite dimensional unitary representations of the Lorentz group. We have demonstrated this explicitly for the spinor representation using the chiral representation (4.16) of the Clifford algebra. We can get a feel for why it is true for a spinor representation constructed from any representation of the Clifford algebra. Recall that

$$S[\Lambda] = \exp \frac{i}{2} \Omega_{\rho\sigma} S^{\rho\sigma} \quad (4.32)$$

so the representation is unitary if  $S^{\mu\nu}$  are anti-hermitian, i.e.  $(S^{\mu\nu})^\dagger = -S^{\mu\nu}$ . But we have

$$(S^{\mu\nu})^\dagger = -\frac{1}{4} [(\gamma^\mu)^\dagger, (\gamma^\nu)^\dagger] \quad (4.33)$$

which can be anti-hermitian if all  $\gamma^\mu$  are hermitian or all are anti-hermitian. However, we can never arrange for this to happen since

$$\begin{aligned} (\gamma^0)^2 &= 1 \Rightarrow \text{Real Eigenvalues} \\ (\gamma^i)^2 &= -1 \Rightarrow \text{Imaginary Eigenvalues} \end{aligned} \quad (4.34)$$

So we could pick  $\gamma^0$  to be hermitian, but we can only pick  $\gamma^i$  to be anti-hermitian. Indeed, in the chiral representation (4.16), the matrices have this property:  $(\gamma^0)^\dagger = \gamma^0$  and  $(\gamma^i)^\dagger = -\gamma^i$ . In general there is no way to pick  $\gamma^\mu$  such that  $S^{\mu\nu}$  are anti-hermitian.

## 4.2 Constructing an Action

We now have a new field to work with, the Dirac spinor  $\psi$ . We would like to construct a Lorentz invariant equation of motion. We do this by constructing a Lorentz invariant action.

We will start in a naive way which won't work, but will give us a clue how to proceed. Define

$$\psi^\dagger(x) = (\psi^\lambda)^T(x) \quad (4.35)$$

which is the usual adjoint of a multi-component object. We could then try to form a Lorentz scalar by taking the product  $\psi^\dagger \psi$ , with the spinor indices summed over. Let's see how this transforms under Lorentz transformations,

$$\begin{aligned} \psi(x) &\rightarrow S[\Lambda] \psi(\Lambda^{-1}x) \\ \psi^\dagger(x) &\rightarrow \psi^\dagger(\Lambda^{-1}x) S[\Lambda]^\dagger \end{aligned} \quad (4.36)$$

So  $\psi^\dagger(x)\psi(x) \rightarrow \psi^\dagger(\Lambda^{-1}x)S[\Lambda]^\dagger S[\Lambda]\psi(\Lambda^{-1}x)$ . But, as we have seen, for some Lorentz transformation  $S[\Lambda]^\dagger S[\Lambda] \neq 1$  since the representation is not unitary. This means that  $\psi^\dagger \psi$  isn't going to do it for us: it doesn't have any nice transformation under the Lorentz group, and certainly isn't a scalar. But now we see why it fails, we can also see how to proceed. Let's pick a representation of the Clifford algebra which, like the chiral representation (4.16), satisfies  $(\gamma^0)^\dagger = \gamma^0$  and  $(\gamma^i)^\dagger = -\gamma^i$ . Then for all  $\mu = 0, 1, 2, 3$  we have

$$\gamma^0 \gamma^\mu \gamma^0 = (\gamma^\mu)^\dagger \quad (4.37)$$

which, in turn, means that

$$(S^{\mu\nu})^\dagger = \frac{1}{4} [(\gamma^\nu)^\dagger, (\gamma^\mu)^\dagger] = -\gamma^0 S^{\mu\nu} \gamma^0 \quad (4.38)$$

so that

$$S[\Lambda]^\dagger = \exp \frac{1}{2} \Omega_{\rho\sigma} (S^{\rho\sigma})^\dagger = \gamma^0 S[\Lambda]^{-1} \gamma^0 \quad (4.39)$$

With this in mind, we now define the *Dirac adjoint*

$$\bar{\psi}(x) = \psi^\dagger(x) \gamma^0 \quad (4.40)$$

Let's now see what Lorentz covariant objects we can form out of a Dirac spinor  $\psi$  and its adjoint  $\bar{\psi}$ .

**Claim 4.3:**  $\bar{\psi} \bar{\psi}$  is a Lorentz scalar.

**Proof:** Under a Lorentz transformation,

$$\begin{aligned} \bar{\psi}(x) \bar{\psi}(x) &= \psi^\dagger(x) \gamma^0 \psi(x) \\ &\rightarrow \psi^\dagger(\Lambda^{-1}x) \gamma^0 S[\Lambda] \gamma^0 S[\Lambda] \psi(\Lambda^{-1}x) \\ &= \psi^\dagger(\Lambda^{-1}x) \gamma^0 \psi(\Lambda^{-1}x) \\ &= \bar{\psi}(\Lambda^{-1}x) \psi(\Lambda^{-1}x) \end{aligned} \quad (4.41)$$

which is indeed the transformation law for a Lorentz scalar.

**Claim 4.4:**  $\bar{\psi} \gamma^\mu \psi$  is a Lorentz vector, which means that

$$\bar{\psi}(x) \gamma^\mu \psi(x) \rightarrow \Lambda^\mu \bar{\psi}(\Lambda^{-1}x) \gamma^\nu \psi(\Lambda^{-1}x) \quad (4.42)$$

This equation means that we can treat the  $\mu = 0, 1, 2, 3$  index on the  $\gamma^\mu$  matrices as a true vector index. In particular we can form Lorentz scalars by contracting it with other Lorentz indices.

**Proof:** Suppressing the  $x$  argument, under a Lorentz transformation we have,

$$\bar{\psi} \gamma^\mu \psi \rightarrow \bar{\psi} S[\Lambda]^{-1} \gamma^\mu S[\Lambda] \psi \quad (4.43)$$

If  $\bar{\psi} \gamma^\mu \psi$  is to transform as a vector, we must have

$$S[\Lambda]^{-1} \gamma^\mu S[\Lambda] = \Lambda^\mu_\nu \gamma^\nu \quad (4.44)$$

We'll now show this. We work infinitesimally, so that

$$\Lambda = \exp \frac{1}{2} \Omega_{\rho\sigma} M^{\rho\sigma} \approx 1 + \frac{1}{2} \Omega_{\rho\sigma} M^{\rho\sigma} + \dots \quad (4.45)$$

$$S[\Lambda] = \exp \frac{1}{2} \Omega_{\rho\sigma} S^{\rho\sigma} \approx 1 + \frac{1}{2} \Omega_{\rho\sigma} S^{\rho\sigma} + \dots \quad (4.46)$$

so the requirement (4.44) becomes

$$-[S^{\rho\sigma}, \gamma^\mu] = (M^{\rho\sigma})^\mu_\nu \gamma^\nu \quad (4.47)$$

where we've suppressed the  $\alpha, \beta$  indices on  $\gamma^\mu$  and  $S^{\mu\nu}$ , but otherwise left all other indices explicit. In fact equation (4.47) follows from Claim 4.1 where we showed that  $[S^{\rho\sigma}, \gamma^\mu] = \gamma^\rho \eta^{\sigma\mu} - \gamma^\sigma \eta^{\mu\rho}$ . To see this, we write the right-hand side of (4.47) by expanding out  $M$ ,

$$\begin{aligned} (M^{\rho\sigma})^\mu_\nu \gamma^\nu &= (\eta^{\rho\mu} \delta^\sigma_\nu - \eta^{\sigma\mu} \delta^\rho_\nu) \gamma^\nu \\ &= \eta^{\rho\mu} \gamma^\sigma - \eta^{\sigma\mu} \gamma^\rho \end{aligned} \quad (4.48)$$

which means that the proof follows if we can show

$$-[S^{\rho\sigma}, \gamma^\mu] = \eta^{\rho\mu} \gamma^\sigma - \eta^{\sigma\mu} \gamma^\rho \quad (4.49)$$

which is exactly what we proved in Claim 4.1.

**Claim 4.5:**  $\bar{\psi} \gamma^\mu \gamma^\nu \psi$  transforms as a Lorentz tensor. More precisely, the symmetric part is a Lorentz scalar, proportional to  $\eta^{\mu\nu} \bar{\psi} \psi$ , while the antisymmetric part is a Lorentz tensor, proportional to  $\bar{\psi} S^{\mu\nu} \psi$ .

**Proof:** As above.

We are now armed with three bilinears of the Dirac field,  $\bar{\psi} \psi$ ,  $\bar{\psi} \gamma^\mu \psi$  and  $\bar{\psi} \gamma^\mu \gamma^\nu \psi$ , each of which transforms covariantly under the Lorentz group. We can try to build a Lorentz invariant action from these. In fact, we need only the first two. We choose

$$S = \int d^4x \bar{\psi}(x) (i\gamma^\mu \partial_\mu - m) \psi(x) \quad (4.50)$$

This is the Dirac action. The factor of "i" is there to make the action real; upon complex conjugation, it cancels a minus sign that comes from integration by parts. (Said another way, it's there for the same reason that the Hermitian momentum operator  $-i\nabla$  in quantum mechanics has a factor  $i$ ). As we will see in the next section, after quantization this theory describes particles and anti-particles of mass  $|m|$  and spin 1/2. Notice that the Lagrangian is first order, rather than the second order Lagrangians we were working with for scalar fields. Also, the mass appears in the Lagrangian as  $m$ , which can be positive or negative.

### 4.3 The Dirac Equation

The equation of motion follows from the action (4.50) by varying with respect to  $\psi$  and  $\bar{\psi}$  independently. Varying with respect to  $\bar{\psi}$ , we have

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad (4.51)$$

This is the *Dirac equation*. It's completely gorgeous. Varying with respect to  $\psi$  gives the conjugate equation

$$i\partial_\mu \bar{\psi} \gamma^\mu + m \bar{\psi} = 0 \quad (4.52)$$

The Dirac equation is first order in derivatives, yet miraculously Lorentz invariant. If we tried to write down a first order equation of motion for a scalar field, it would look like  $v^\mu \partial_\mu \varphi = \dots$ , which necessarily includes a privileged vector in spacetime  $v^\mu$  and is not Lorentz invariant. However, for spinor fields, the magic of the  $\gamma^\mu$  matrices means that the Dirac Lagrangian is Lorentz invariant.

The Dirac equation mixes up different components of  $\psi$  through the matrices  $\gamma^\mu$ . However, each individual component itself solves the Klein-Gordon equation. To see this, write

$$(i\gamma^\nu \partial_\nu + m)(i\gamma^\mu \partial_\mu - m)\psi = -\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + m^2 \psi = 0 \quad (4.53)$$

But  $\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu = \frac{1}{2}\{\gamma^\mu, \gamma^\nu\} \partial_\mu \partial_\nu = \partial_\mu \partial^\mu$ , so we get

$$-(\partial_\mu \partial^\mu + m^2)\psi = 0 \quad (4.54)$$

where this last equation has no  $\gamma^\mu$  matrices, and so applies to each component  $\psi^\alpha$ , with  $\alpha = 1, 2, 3, 4$ .

### The Slash

Let's introduce some useful notation. We will often come across 4-vectors contracted with  $\gamma^\mu$  matrices. We write

$$A_\mu \gamma^\mu \equiv A \quad (4.55)$$

so the Dirac equation reads

$$(i\partial/ - m)\psi = 0 \quad (4.56)$$

## 4.4 Chiral Spinors

When we've needed an explicit form of the  $\gamma^\mu$  matrices, we've used the chiral representation

$$\gamma^0 = \begin{matrix} & 1 \\ 0 & 1 \\ & 0 \end{matrix}, \quad \gamma^i = \begin{matrix} & 1 \\ 0 & \sigma^i \\ -\sigma^i & 0 \end{matrix} \quad (4.57)$$

In this representation, the spinor rotation transformation  $S[\Lambda_{\text{rot}}]$  and boost transformation  $S[\Lambda_{\text{boost}}]$  were computed in (4.26) and (4.31). Both are block diagonal,

$$S[\Lambda_{\text{rot}}] = \begin{matrix} e^{i\phi \cdot \vec{\sigma}/2} & 0 \\ 0 & e^{i\phi \cdot \vec{\sigma}/2} \end{matrix} \quad \text{and} \quad S[\Lambda_{\text{boost}}] = \begin{matrix} e^{i\chi \cdot \vec{\sigma}/2} & 0 \\ 0 & e^{-i\chi \cdot \vec{\sigma}/2} \end{matrix} \quad (4.58)$$

This means that the Dirac spinor representation of the Lorentz group is *reducible*. It decomposes into two irreducible representations, acting only on two-component spinors  $u_\pm$  which, in the chiral representation, are defined by

$$\psi = \begin{matrix} u_+ \\ u_- \end{matrix} \quad (4.59)$$

The two-component objects  $u_\pm$  are called *Weyl spinors* or *chiral spinors*. They transform in the same way under rotations,

$$u_\pm \rightarrow e^{i\phi \cdot \vec{\sigma}/2} u_\pm \quad (4.60)$$

but oppositely under boosts,

$$u_\pm \rightarrow e^{\pm i\chi \cdot \vec{\sigma}/2} u_\pm \quad (4.61)$$

In group theory language,  $u_+$  is in the  $(\frac{1}{2}, 0)$  representation of the Lorentz group, while  $u_-$  is in the  $(0, \frac{1}{2})$  representation. The Dirac spinor  $\psi$  lies in the  $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$  representation. (Strictly speaking, the spinor is a representation of the double cover of the Lorentz group  $SL(2, \mathbb{C})$ ).

### 4.4.1 The Weyl Equation

Let's see what becomes of the Dirac Lagrangian under the decomposition (4.59) into Weyl spinors. We have

$$L = \bar{\psi} (i \not{d} - m) \psi = i u_+^\dagger \not{\sigma}^\mu \partial_\mu u_- + i u_-^\dagger \not{\sigma}^\mu \partial_\mu u_+ - m(u_+^\dagger u_- + u_-^\dagger u_+) = 0 \quad (4.62)$$

where we have introduced some new notation for the Pauli matrices with a  $\mu = 0, 1, 2, 3$  index,

$$\sigma^\mu = (1, \sigma^i) \quad \text{and} \quad \sigma^{-\mu} = (1, -\sigma^i) \quad (4.63)$$

From (4.62), we see that a massive fermion requires both  $u_+$  and  $u_-$ , since they couple through the mass term. However, a massless fermion can be described by  $u_+$  (or  $u_-$ ) alone, with the equation of motion

$$\begin{aligned} i\bar{\sigma}^\mu \partial_\mu u_+ &= 0 \\ \text{or} \quad i\sigma^\mu \partial_\mu u_- &= 0 \end{aligned} \quad (4.64)$$

These are the *Weyl equations*.

### Degrees of Freedom

Let me comment here on the degrees of freedom in a spinor. The Dirac fermion has 4 complex components = 8 real components. How do we count degrees of freedom? In classical mechanics, the number of degrees of freedom of a system is equal to the dimension of the configuration space or, equivalently, half the dimension of the phase space. In field theory we have an infinite number of degrees of freedom, but it makes sense to count the number of degrees of freedom per spatial point: this should at least be finite. For example, in this sense a real scalar field  $\varphi$  has a single degree of freedom. At the quantum level, this translates to the fact that it gives rise to a single type of particle. A classical complex scalar field has two degrees of freedom, corresponding to the particle and the anti-particle in the quantum theory.

But what about a Dirac spinor? One might think that there are 8 degrees of freedom. But this isn't right. Crucially, and in contrast to the scalar field, the equation of motion is first order rather than second order. In particular, for the Dirac Lagrangian, the momentum conjugate to the spinor  $\psi$  is given by

$$\pi_\psi = \partial L / \partial \dot{\psi} = i\psi^\dagger \quad (4.65)$$

It is not proportional to the time derivative of  $\psi$ . This means that the phase space for a spinor is therefore parameterized by  $\psi$  and  $\psi^\dagger$ , while for a scalar it is parameterized by  $\varphi$  and  $\pi = \dot{\varphi}$ . So the *phase space* of the Dirac spinor  $\psi$  has 8 real dimensions and correspondingly the number of real degrees of freedom is 4. We will see in the next section that, in the quantum theory, this counting manifests itself as two degrees of freedom (spin up and down) for the particle, and a further two for the anti-particle.

A similar counting for the Weyl fermion tells us that it has two degrees of freedom.

### 4.4.2 $\gamma^5$

The Lorentz group matrices  $S[\Lambda]$  came out to be block diagonal in (4.58) because we chose the specific representation (4.57). In fact, this is why the representation (4.57) is called the chiral representation: it's because the decomposition of the Dirac spinor  $\psi$  is simply given by (4.59). But what happens if we choose a different representation  $\gamma^\mu$  of the Clifford algebra, so that

$$\gamma^\mu \rightarrow U\gamma^\mu U^{-1} \quad \text{and} \quad \psi \rightarrow U\psi \quad ? \quad (4.66)$$

Now  $S[\Lambda]$  will not be block diagonal. Is there an invariant way to define chiral spinors? We can do this by introducing the “fifth” gamma-matrix

$$\gamma^5 = -i\gamma^0\gamma^1\gamma^2\gamma^3 \quad (4.67)$$

You can check that this matrix satisfies

$$\{\gamma^5, \gamma^\mu\} = 0 \quad \text{and} \quad (\gamma^5)^2 = +1 \quad (4.68)$$

The reason that this is called  $\gamma^5$  is because the set of matrices  $\tilde{\gamma}^A = (\gamma^\mu, i\gamma^5)$ , with  $A = 0, 1, 2, 3, 4$  satisfy the  $d = 4 + 1$  Clifford algebra  $\{\tilde{\gamma}^A, \tilde{\gamma}^B\} = 2\eta^{AB}$ . (You might think that  $\gamma^4$  would be a better name! But  $\gamma^5$  is the one everyone chooses - it's a more sensible name in Euclidean space, where  $A = 1, 2, 3, 4, 5$ ). You can also check that  $[S_{\mu\nu}, \gamma^5] = 0$ , which means that  $\gamma^5$  is a scalar under rotations and boosts. Since  $(\gamma^5)^2 = 1$ , this means we may form the Lorentz invariant projection operators

$$P_\pm = \frac{1}{2} (1 \pm \gamma^5) \quad (4.69)$$

such that  $P_+^2 = P_+$  and  $P_-^2 = P_-$  and  $P_+P_- = 0$ . One can check that for the chiral representation (4.57),

$$\gamma^5 = \begin{matrix} & & & & 1 & 0 \\ & & & & 0 & -1 \end{matrix} \quad ! \quad (4.70)$$

from which we see that the operators  $P_\pm$  project onto the Weyl spinors  $u_\pm$ . However, for an arbitrary representation of the Clifford algebra, we may use  $\gamma^5$  to define the chiral spinors,

$$\psi_\pm = P_\pm \psi \quad (4.71)$$

which form the irreducible representations of the Lorentz group.  $\psi_+$  is often called a “left-handed” spinor, while  $\psi_-$  is “right-handed”. The name comes from the way the spin precesses as a massless fermion moves: we'll see this in Section 4.7.2.

### 4.4.3 Parity

The spinors  $\psi_{\pm}$  are related to each other by *parity*. Let's pause to define this concept. The Lorentz group is defined by  $x^{\mu} \rightarrow \Lambda_v^{\mu} x^v$  such that

$$\Lambda_v^{\mu} \Lambda_v^{\rho} \eta^{\nu\sigma} = \eta^{\mu\rho} \quad (4.72)$$

So far we have only considered transformations  $\Lambda$  which are continuously connected to the identity; these are the ones which have an infinitesimal form. However there are also two discrete symmetries which are part of the Lorentz group. They are

$$\begin{aligned} \text{Time Reversal } T : x^0 &\rightarrow -x^0 ; x^i \rightarrow x^i \\ \text{Parity } P : x^0 &\rightarrow x^0 ; x^i \rightarrow -x^i \end{aligned} \quad (4.73)$$

We won't discuss time reversal too much in this course. (It turns out to be represented by an anti-unitary transformation on states. See, for example the book by Peskin and Schroeder). But parity has an important role to play in the standard model and, in particular, the theory of the weak interaction.

Under parity, the left and right-handed spinors are exchanged. This follows from the transformation of the spinors under the Lorentz group. In the chiral representation, we saw that the rotation (4.60) and boost (4.61) transformations for the Weyl spinors  $u_{\pm}$  are

$$u_{\pm} \xrightarrow{\text{rot}} e^{i\phi_{\pm}/2} u_{\pm} \quad \text{and} \quad u_{\pm} \xrightarrow{\text{boost}} e^{\pm \mathbf{x} \cdot \mathbf{\hat{p}}/2} u_{\pm} \quad (4.74)$$

Under parity, rotations don't change sign. But boosts do flip sign. This confirms that parity exchanges right-handed and left-handed spinors,  $P : u_{\pm} \rightarrow u_{\mp}$ , or in the notation  $\psi_{\pm} = \frac{1}{2}(1 \pm \gamma^5)\psi$ , we have

$$P : \psi_{\pm}(\mathbf{x}, t) \rightarrow \psi_{\mp}(-\mathbf{x}, t) \quad (4.75)$$

Using this knowledge of how chiral spinors transform, and the fact that  $P^2 = 1$ , we see that the action of parity on the Dirac spinor itself can be written as

$$P : \psi(\mathbf{x}, t) \rightarrow \gamma^0 \psi(-\mathbf{x}, t) \quad (4.76)$$

Notice that if  $\psi(\mathbf{x}, t)$  satisfies the Dirac equation, then the parity transformed spinor  $\gamma^0 \psi(-\mathbf{x}, t)$  also satisfies the Dirac equation, meaning

$$(i\gamma^0 \partial_t + i\gamma^i \partial_i - m) \gamma^0 \psi(-\mathbf{x}, t) = \gamma^0 (i\gamma^0 \partial_t - i\gamma^i \partial_i - m) \psi(-\mathbf{x}, t) = 0 \quad (4.77)$$

where the extra minus sign from passing  $\gamma^0$  through  $\gamma^i$  is compensated by the derivative acting on  $-\mathbf{x}$  instead of  $+\mathbf{x}$ .

#### 4.4.4 Chiral Interactions

Let's now look at how our interaction terms change under parity. We can look at each of our spinor bilinears from which we built the action,

$$P: \bar{\psi} \psi (\mathbf{x}, t) \rightarrow \bar{\psi} \psi (-\mathbf{x}, t) \quad (4.78)$$

which is the transformation of a scalar. For the vector  $\bar{\psi} \gamma^\mu \psi$ , we can look at the temporal and spatial components separately,

$$\begin{aligned} P: \bar{\psi} \gamma^0 \psi (\mathbf{x}, t) &\rightarrow \bar{\psi} \gamma^0 \psi (-\mathbf{x}, t) \\ P: \bar{\psi} \gamma^i \psi (\mathbf{x}, t) &\rightarrow \bar{\psi} \gamma^0 \gamma^i \gamma^0 \psi (-\mathbf{x}, t) = -\bar{\psi} \gamma^i \psi (-\mathbf{x}, t) \end{aligned} \quad (4.79)$$

which tells us that  $\bar{\psi} \gamma^\mu \psi$  transforms as a vector, with the spatial part changing sign. You can also check that  $\bar{\psi} S^{\mu\nu} \psi$  transforms as a suitable tensor.

However, now we've discovered the existence of  $\gamma^5$ , we can form another Lorentz scalar and another Lorentz vector,

$$\bar{\psi} \gamma^5 \psi \text{ and } \bar{\psi} \gamma^5 \gamma^\mu \psi \quad (4.80)$$

How do these transform under parity? We can check:

$$\begin{aligned} P: \bar{\psi} \gamma^5 \psi (\mathbf{x}, t) &\rightarrow \bar{\psi} \gamma^0 \gamma^5 \gamma^0 \psi (-\mathbf{x}, t) = -\bar{\psi} \gamma^5 \psi (-\mathbf{x}, t) \quad (4.81) \\ P: \bar{\psi} \gamma^5 \psi (\mathbf{x}, t) &\rightarrow \bar{\psi} \gamma^0 \gamma^5 \gamma^0 \psi (-\mathbf{x}, t) = -\bar{\psi} \gamma^5 \gamma^0 \psi (-\mathbf{x}, t) \quad \mu = 0 \\ &\quad + \bar{\psi} \gamma^5 \gamma^i \psi (-\mathbf{x}, t) \quad \mu = i \end{aligned}$$

which means that  $\bar{\psi} \gamma^5 \psi$  transforms as a *pseudoscalar*, while  $\bar{\psi} \gamma^5 \gamma^\mu \psi$  transforms as an *axial vector*. To summarize, we have the following spinor bilinears,

$$\begin{aligned} \bar{\psi} \psi &: \text{scalar} \\ \bar{\psi} \gamma^\mu \psi &: \text{vector} \\ \bar{\psi} S^{\mu\nu} \psi &: \text{tensor} \\ \bar{\psi} \gamma^5 \psi &: \text{pseudoscalar} \\ \bar{\psi} \gamma^5 \gamma^\mu \psi &: \text{axial vector} \end{aligned} \quad (4.82)$$

The total number of bilinears is  $1 + 4 + (4 \times 3/2) + 4 + 1 = 16$  which is all we could hope for from a 4-component object.

We're now armed with new terms involving  $\gamma^5$  that we can start to add to our Lagrangian to construct new theories. Typically such terms will break parity invariance of the theory, although this is not always true. (For example, the term  $\varphi\bar{\psi}\gamma^5\psi$  doesn't break parity if  $\varphi$  is itself a pseudoscalar). Nature makes use of these parity violating interactions by using  $\gamma^5$  in the weak force. A theory which treats  $\psi_{\pm}$  on an equal footing is called a *vector-like theory*. A theory in which  $\psi_+$  and  $\psi_-$  appear differently is called a *chiral theory*.

## 4.5 Majorana Fermions

Our spinor  $\psi^\alpha$  is a complex object. It has to be because the representation  $S[\Lambda]$  is typically also complex. This means that if we were to try to make  $\psi$  real, for example by imposing  $\psi = \psi^\wedge$ , then it wouldn't stay that way once we make a Lorentz transformation. However, there is a way to impose a reality condition on the Dirac spinor  $\psi$ . To motivate this possibility, it's simplest to look at a novel basis for the Clifford algebra, known as the *Majorana basis*.

$$\gamma^0 = \begin{matrix} & & & & & & & \\ & 0 & \sigma^2 & & & & & \\ & \sigma^2 & 0 & & & & & \end{matrix}, \quad \gamma^1 = \begin{matrix} & & & & & & & \\ & i\sigma^3 & 0 & & & & & \\ & 0 & i\sigma^3 & & & & & \end{matrix}, \quad \gamma^2 = \begin{matrix} & & & & & & & \\ & 0 & -\sigma^2 & & & & & \\ & \sigma^2 & 0 & & & & & \end{matrix}, \quad \gamma^3 = \begin{matrix} & & & & & & & \\ & -i\sigma^1 & 0 & & & & & \\ & 0 & -i\sigma^1 & & & & & \end{matrix}$$

These matrices satisfy the Clifford algebra. What is special about them is that they are all pure imaginary  $(\gamma^\mu)^\wedge = -\gamma^\mu$ . This means that the generators of the Lorentz group  $S^{\mu\nu} = \frac{1}{4} [\gamma^\mu, \gamma^\nu]$ , and hence the matrices  $S[\Lambda]$  are real. So with this basis of the Clifford algebra, we can work with a real spinor simply by imposing the condition,

$$\psi = \psi^\wedge \tag{4.83}$$

which is preserved under Lorentz transformation. Such spinors are called *Majorana spinors*.

So what's the story if we use a general basis for the Clifford algebra? We'll ask only that the basis satisfies  $(\gamma^0)^\dagger = \gamma^0$  and  $(\gamma^i)^\dagger = -\gamma^i$ . We then define the *charge conjugate* of a Dirac spinor  $\psi$  as

$$\psi^{(c)} = C\psi^\wedge \tag{4.84}$$

Here  $C$  is a  $4 \times 4$  matrix satisfying

$$C^\dagger C = 1 \quad \text{and} \quad C^\dagger \gamma^\mu C = -(\gamma^\mu)^\wedge \tag{4.85}$$

Let's firstly check that (4.84) is a good definition, meaning that  $\psi^{(c)}$  transforms nicely under a Lorentz transformation. We have

$$\psi^{(c)} \rightarrow CS[\Lambda]^\wedge \psi^\wedge = S[\Lambda]C\psi^\wedge = S[\Lambda]\psi^{(c)} \tag{4.86}$$

where we've made use of the properties (4.85) in taking the matrix  $C$  through  $S[\Lambda]^\wedge$ . In fact, not only does  $\psi^{(c)}$  transform nicely under the Lorentz group, but if  $\psi$  satisfies the Dirac equation, then  $\psi^{(c)}$  does too. This follows from,

$$\begin{aligned}(i\cancel{d} - m)\psi &= 0 \Rightarrow (-i\partial^\wedge - m)\psi^\wedge = 0 \\ \Rightarrow C(-i\partial^\wedge - m)\psi^\wedge &= (+i\cancel{d} - m)\psi^{(c)} = 0\end{aligned}$$

Finally, we can now impose the Lorentz invariant reality condition on the Dirac spinor, to yield a Majorana spinor,

$$\psi^{(c)} = \psi \tag{4.87}$$

After quantization, the Majorana spinor gives rise to a fermion that is its own anti-particle. This is exactly the same as in the case of scalar fields, where we've seen that a real scalar field gives rise to a spin 0 boson that is its own anti-particle. (Be aware: In many texts an extra factor of  $\gamma^0$  is absorbed into the definition of  $C$ ).

So what is this matrix  $C$ ? Well, for a given representation of the Clifford algebra, it is something that we can find fairly easily. In the Majorana basis, where the gamma matrices are pure imaginary, we have simply  $C_{\text{Maj}} = 1$  and the Majorana condition  $\psi = \psi^{(c)}$  becomes  $\psi = \psi^\wedge$ . In the chiral basis (4.16), only  $\gamma^2$  is imaginary, and we may take  $C_{\text{chiral}} = i\gamma^2 = \begin{pmatrix} 0 & i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}$ . (The matrix  $i\sigma^2$  that appears here is simply the anti-symmetric matrix  $\epsilon^{\alpha\beta}$ ). It is interesting to see how the Majorana condition (4.87) looks in terms of the decomposition into left and right handed Weyl spinors (4.59). Plugging in the various definitions, we find that  $u_+ = i\sigma^2 u^\wedge$  and  $u_- = -i\sigma^2 u^\wedge$ . In other words, a Majorana spinor can be written in terms of Weyl spinors as

$$\psi = \begin{pmatrix} u_+ \\ -i\sigma^2 u_+^\wedge \end{pmatrix} \tag{4.88}$$

Notice that it's not possible to impose the Majorana condition  $\psi = \psi^{(c)}$  at the same time as the Weyl condition ( $u_- = 0$  or  $u_+ = 0$ ). Instead the Majorana condition relates  $u_-$  and  $u_+$ .

**An Aside: Spinors in Different Dimensions:** The ability to impose Majorana or Weyl conditions on Dirac spinors depends on both the dimension and the signature of spacetime. One can always impose the Weyl condition on a spinor in even dimensional Minkowski space, basically because you can always build a suitable “ $\gamma^5$ ” projection matrix by multiplying together all the other  $\gamma$ -matrices. The pattern for when the Majorana condition can be imposed is a little more sporadic. Interestingly, although the Majorana condition and Weyl condition cannot be imposed simultaneously in four dimensions, you can do this in Minowski spacetimes of dimension 2, 10, 18, . . . .

## 4.6 Symmetries and Conserved Currents

The Dirac Lagrangian enjoys a number of symmetries. Here we list them and compute the associated conserved currents.

### Spacetime Translations

Under spacetime translations the spinor transforms as

$$\delta\psi = \epsilon^\mu \partial_\mu \psi \quad (4.89)$$

The Lagrangian depends on  $\partial_\mu \psi$ , but not  $\partial_\mu \bar{\psi}$ , so the standard formula (1.41) gives us the energy-momentum tensor

$$T^{\mu\nu} = i\bar{\psi} \gamma^\mu \partial^\nu \psi - \eta^{\mu\nu} L \quad (4.90)$$

Since a current is conserved only when the equations of motion are obeyed, we don't lose anything by imposing the equations of motion already on  $T^{\mu\nu}$ . In the case of a scalar field this didn't really buy us anything because the equations of motion are second order in derivatives, while the energy-momentum is typically first order. However, for a spinor field the equations of motion are first order:  $(i\partial/\partial t - m)\psi = 0$ . This means we can set  $L = 0$  in  $T^{\mu\nu}$ , leaving

$$T^{\mu\nu} = i\bar{\psi} \gamma^\mu \partial^\nu \psi \quad (4.91)$$

In particular, we have the total energy

$$E = \int d^3x T^{00} = \int d^3x i\bar{\psi} \gamma^0 \psi = \int d^3x \psi^\dagger \gamma^0 (-i\gamma^i \partial_i + m) \psi \quad (4.92)$$

where, in the last equality, we have again used the equations of motion.

### Lorentz Transformations

Under an infinitesimal Lorentz transformation, the Dirac spinor transforms as (4.22) which, in infinitesimal form, reads

$$\delta\psi^\alpha = -\omega_v^\mu x^\nu \partial_\mu \psi^\alpha + \frac{1}{2} \Omega_{\rho\sigma} (S^{\rho\sigma})^\alpha_\beta \psi^\beta \quad (4.93)$$

where, following (4.10), we have  $\omega^\mu_v = \frac{1}{2} \Omega_{\rho\sigma} (M^{\rho\sigma})^\mu_v$ , and  $M^{\rho\sigma}$  are the generators of the Lorentz algebra given by (4.8)

$$(M^{\rho\sigma})^\mu_v = \eta^{\rho\mu} \delta^\sigma_v - \eta^{\sigma\mu} \delta^\rho_v \quad (4.94)$$

which, after direct substitution, tells us that  $\omega^{\mu\nu} = \Omega^{\mu\nu}$ . So we get

$$\delta\psi^\alpha = -\omega^{\mu\nu} x_\nu \partial_\mu \psi^\alpha - \frac{1}{2} (S_{\mu\nu})^\alpha_\beta \psi^\beta \quad (4.95)$$

The conserved current arising from Lorentz transformations now follows from the same calculation we saw for the scalar field (1.54) with two differences: firstly, as we saw above, the spinor equations of motion set  $L = 0$ ; secondly, we pick up an extra piece in the current from the second term in (4.95). We have

$$(J^\mu)^{\rho\sigma} = x^\rho T^{\mu\sigma} - x^\sigma T^{\mu\rho} + i\bar{\psi} \gamma^\mu S^{\rho\sigma} \psi \quad (4.96)$$

After quantization, when  $(J^\mu)^{\rho\sigma}$  is turned into an operator, this extra term will be responsible for providing the single particle states with internal angular momentum, telling us that the quantization of a Dirac spinor gives rise to a particle carrying spin 1/2.

### Internal Vector Symmetry

The Dirac Lagrangian is invariant under rotating the phase of the spinor,  $\psi \rightarrow e^{-i\alpha}\psi$ . This gives rise to the current

$$j_V^\mu = \bar{\psi} \gamma^\mu \psi \quad (4.97)$$

where “V” stands for *vector*, reflecting the fact that the left and right-handed components  $\psi_\pm$  transform in the same way under this symmetry. We can easily check that  $j_V^\mu$  is conserved under the equations of motion,

$$\partial_\mu j_V^\mu = (\partial_\mu \bar{\psi}) \gamma^\mu \psi + \bar{\psi} \gamma^\mu (\partial_\mu \psi) = i m \bar{\psi} \psi - i m \bar{\psi} \psi = 0 \quad (4.98)$$

where, in the last equality, we have used the equations of motion  $i\partial/\psi = m\psi$  and  $i\partial_\mu \bar{\psi} \gamma^\mu = -m\bar{\psi}$ . The conserved quantity arising from this symmetry is

$$Q = \int d^3x \bar{\psi} \gamma^0 \psi = \int d^3x \bar{\psi} \gamma^0 \psi \quad (4.99)$$

We will see shortly that this has the interpretation of electric charge, or particle number, for fermions.

### Axial Symmetry

When  $m = 0$ , the Dirac Lagrangian admits an extra internal symmetry which rotates left and right-handed fermions in opposite directions,

$$\psi \rightarrow e^{i\alpha\gamma^5} \psi \quad \text{and} \quad \bar{\psi} \rightarrow \bar{\psi} e^{i\alpha\gamma^5} \quad (4.100)$$

Here the second transformation follows from the first after noting that  $e^{-ia\gamma^5} \gamma^0 = \gamma^0 e^{+ia\gamma^5}$ . This gives the conserved current,

$$j_A^\mu = \bar{\psi} \gamma^\mu \gamma^5 \psi \quad (4.101)$$

where  $A$  is for “axial” since  $j_A^\mu$  is an axial vector. This is conserved only when  $m = 0$ . Indeed, with the full Dirac Lagrangian we may compute

$$\partial_\mu j_A^\mu = (\partial_\mu \bar{\psi}) \gamma^\mu \gamma^5 \psi + \bar{\psi} \gamma^\mu \gamma^5 \partial_\mu \psi = 2im \bar{\psi} \gamma^5 \psi \quad (4.102)$$

which vanishes only for  $m = 0$ . However, in the quantum theory things become more interesting for the axial current. When the theory is coupled to gauge fields (in a manner we will discuss in Section 6), the axial transformation remains a symmetry of the classical Lagrangian. But it doesn’t survive the quantization process. It is the archetypal example of an *anomaly*: a symmetry of the classical theory that is not preserved in the quantum theory.

## 4.7 Plane Wave Solutions

Let’s now study the solutions to the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0 \quad (4.103)$$

We start by making a simple ansatz:

$$\psi = u(\not{p}) e^{-ip \cdot x} \quad (4.104)$$

where  $u(p \rightarrow)$  is a four-component spinor, independent of spacetime  $x$  which, as the notation suggests, can depend on the 3-momentum  $p \rightarrow$ . The Dirac equation then becomes

$$(i\gamma^\mu p_\mu - m) u(\not{p}) = \begin{pmatrix} -m & p_\mu \sigma^\mu \\ p_\mu \bar{\sigma}^\mu & -m \end{pmatrix} u(\not{p}) = 0 \quad (4.105)$$

where we’re again using the definition,

$$\sigma^\mu = (1, \sigma^i) \quad \text{and} \quad \sigma^- \mu = (1, -\sigma^i) \quad (4.106)$$

**Claim:** The solution to (4.105) is

$$u(\not{p}) = \frac{\sqrt{\frac{p \cdot \sigma \xi}{p \cdot \bar{\sigma} \xi}}}{\sqrt{p \cdot \bar{\sigma} \xi}} \quad (4.107)$$

for any 2-component spinor  $\xi$  which we will normalize to  $\xi^\dagger \xi = 1$ .

**Proof:** Let's write  $u(\not{p})^T = (u_1, u_2)$ . Then equation (4.105) reads

$$(p \cdot \sigma) u_2 = m u_1 \quad \text{and} \quad (p \cdot \bar{\sigma}) u_1 = m u_2 \quad (4.108)$$

Either one of these equations implies the other, a fact which follows from the identity  $(p \cdot \sigma)(p \cdot \sigma^\dagger) = p^2 - p_i p_j \sigma^i \sigma^j = p_0^2 - p_i p_j \delta^{ij} = p_\mu p^\mu = m^2$ . To start with, let's try the ansatz  $u_1 = (p \cdot \sigma) \xi^r$  for some spinor  $\xi^r$ . Then the second equation in (4.108) immediately tells us that  $u_2 = m \xi^r$ . So we learn that any spinor of the form

$$\frac{u(\not{p})}{m \xi^r} = A \frac{(p \cdot \sigma) \xi^r}{m \xi^r} \quad (4.109)$$

with constant  $A$  is a solution to (4.105). To make this more symmetric, we choose  $A = 1/m$  and  $\xi^r = \sqrt{-p \cdot \bar{\sigma}} \xi$  with constant  $\xi$ . Then  $u_1 = (p \cdot \sigma) \sqrt{-p \cdot \bar{\sigma}} \xi = m \sqrt{-p \cdot \sigma} \xi$ . So we get the promised result (4.107)

### Negative Frequency Solutions

We get further solutions to the Dirac equation from the ansatz

$$\psi = v(\not{p}) e^{+i\not{p} \cdot x} \quad (4.110)$$

Solutions of the form (4.104), which oscillate in time as  $\psi \sim e^{-iEt}$ , are called positive frequency solutions. If we compute the energy of these solutions using (4.92), we find that it is positive. Those of the form (4.110), which oscillate as  $\psi \sim e^{+iEt}$ , are negative frequency solutions. Now if we compute the energy using (4.92), it is negative.

The Dirac equation requires that the 4-component spinor  $v(\not{p})$  satisfies

$$(\gamma^\mu p_\mu + m) v(\not{p}) = \frac{m}{p_\mu \bar{\sigma}^\mu} \frac{p_\mu \sigma^\mu}{m} v(\not{p}) = 0 \quad (4.111)$$

which is solved by

$$v(\not{p}) = \frac{\sqrt{\frac{\not{p} \cdot \sigma \eta}{\not{p} \cdot \bar{\sigma} \eta}}}{-\not{p} \cdot \bar{\sigma} \eta} \quad (4.112)$$

for some 2-component spinor  $\eta$  which we take to be constant and normalized to  $\eta^\dagger \eta = 1$ .

### 4.7.1 Some Examples

Consider the positive frequency solution with mass  $m$  and 3-momentum  $\not{p} = 0$ ,

$$u(\not{p}) = \frac{\sqrt{m}}{\xi} \begin{pmatrix} \xi^+ \\ \xi^- \end{pmatrix} \quad (4.113)$$

where  $\xi$  is any 2-component spinor. Spatial rotations of the field act on  $\xi$  by (4.26),

$$\xi \rightarrow e^{i\phi \not{\sigma} \cdot \not{\sigma}/2} \xi \quad (4.114)$$

The 2-component spinor  $\xi$  defines the *spin* of the field. This should be familiar from quantum mechanics. A field with spin up (down) along a given direction is described by the eigenvector of the corresponding Pauli matrix with eigenvalue +1 (-1 respectively). For example,  $\xi^T = (1, 0)$  describes a field with spin up along the z-axis. After quantization, this will become the spin of the associated particle. In the rest of this section, we'll indulge in an abuse of terminology and refer to the classical solutions to the Dirac equations as "particles", even though they have no such interpretation before quantization.

Consider now boosting the particle with spin  $\xi^T = (1, 0)$  along the  $x^3$  direction, with  $p^\mu = (E, 0, 0, p^3)$ . The solution to the Dirac equation becomes

$$u(\not{p}) = \frac{\sqrt{p \cdot \sigma}}{\sqrt{p \cdot \bar{\sigma}}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \sqrt{\frac{E - p^3}{E + p^3}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (4.115)$$

In fact, this expression also makes sense for a massless field, for which  $E = p^3$ . (We picked the normalization (4.107) for the solutions so that this would be the case). For a massless particle we have

$$u(\not{p}) = \frac{\sqrt{2E}}{\sqrt{2E}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (4.116)$$

Similarly, for a boosted solution of the spin down  $\xi^T = (0, 1)$  field, we have

$$u(p \rightarrow) = \frac{\sqrt{p \cdot \sigma}}{\sqrt{p \cdot \bar{\sigma}}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sqrt{\frac{E + p^3}{E - p^3}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \xrightarrow{m \rightarrow 0} \sqrt{\frac{2E}{2E}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (4.117)$$

### 4.7.2 Helicity

The helicity operator is the projection of the angular momentum along the direction of momentum,

$$h = \frac{i}{2} \epsilon_{ij} k p^j S^i = \frac{1}{2} \begin{matrix} p^i & \sigma^i \\ 0 & \sigma^i \end{matrix} \quad (4.118)$$

where  $S^{ij}$  is the rotation generator given in (4.25). The massless field with spin  $\xi^T = (1, 0)$  in (4.116) has helicity  $h = 1/2$ : we say that it is *right-handed*. Meanwhile, the field (4.117) has helicity  $h = -1/2$ : it is *left-handed*.

### 4.7.3 Some Useful Formulae: Inner and Outer Products

There are a number of identities that will be very useful in the following section, regarding the inner (and outer) products of the spinors  $u(p)$  and  $v(p)$ . It's firstly convenient to introduce a basis  $\xi^s$  and  $\eta^s$ ,  $s = 1, 2$  for the two-component spinors such that

$$\xi^r \xi^s = \delta^{rs} \quad \text{and} \quad \eta^r \eta^s = \delta^{rs} \quad (4.119)$$

for example,

$$\xi^1 = \begin{matrix} 1 & ! \\ 0 & \end{matrix} \quad \text{and} \quad \xi^2 = \begin{matrix} 0 & ! \\ 1 & \end{matrix} \quad (4.120)$$

and similarly for  $\eta^s$ . Let's deal first with the positive frequency plane waves. The two independent solutions are now written as

$$u^s(p) = \sqrt{\frac{p \cdot \sigma}{p \cdot \bar{\sigma}}} \xi^s \quad (4.121)$$

We can take the inner product of four-component spinors in two different ways: either as  $u^\dagger \cdot u$ , or as  $\bar{u} \cdot u$ . Of course, only the latter will be Lorentz invariant, but it turns out that the former is needed when we come to quantize the theory. Here we state both:

$$\begin{aligned} u^r(p) \cdot u^s(p) &= \xi^r \dagger \sqrt{p \cdot \sigma} \xi^s, \xi^r \dagger \sqrt{p \cdot \bar{\sigma}} \xi^s \\ &= \xi^r \dagger p \cdot \sigma \xi^s + \xi^r \dagger p \cdot \bar{\sigma} \xi^s = 2 \xi^r \dagger p_0 \xi^s = 2 p_0 \delta^{rs} \end{aligned} \quad (4.122)$$

while the Lorentz invariant inner product is

$$u^{-r}(p) \cdot u^s(p) = \xi^r \dagger \sqrt{p \cdot \sigma} \xi^s, \xi^r \dagger \sqrt{p \cdot \bar{\sigma}} \xi^s = 2m \delta^{rs} \quad (4.123)$$

We have analogous results for the negative frequency solutions, which we may write as

$$v^s(\not{p}) = \frac{\sqrt{\frac{p \cdot \sigma}{\not{p} \cdot \bar{\sigma}} \eta^s}}{\sqrt{p \cdot \bar{\sigma} \eta^s}} \quad \text{with} \quad v^{r\dagger}(\not{p}) \cdot v^s(\not{p}) = 2p_0 \delta^{rs} \quad (4.124)$$

$$\text{and} \quad \bar{v}^{r\dagger}(\not{p}) \cdot v^s(\not{p}) = -2m \delta^{rs}$$

We can also compute the inner product between  $u$  and  $v$ . We have

$$\begin{aligned} \bar{u}^r(\not{p}) \cdot v^s(\not{p}) &= \xi^{r\dagger} \frac{\sqrt{p \cdot \sigma}}{\sqrt{p \cdot \bar{\sigma}}} \gamma^0 \frac{\sqrt{\frac{p \cdot \sigma}{\not{p} \cdot \bar{\sigma}} \eta^s}}{\sqrt{p \cdot \bar{\sigma} \eta^s}} \\ &= \xi^{r\dagger} \frac{\sqrt{(p \cdot \bar{\sigma})(p \cdot \sigma) \eta^s}}{\sqrt{(p \cdot \bar{\sigma})(p \cdot \sigma) \eta^s}} = \xi^{r\dagger} (p \cdot \bar{\sigma})(p \cdot \sigma) \eta^s = 0 \end{aligned} \quad (4.125)$$

and similarly,  $v^{-r}(\not{p}) \cdot u^s(\not{p}) = 0$ . However, when we come to  $u^\dagger \cdot v$ , it is a slightly different combination that has nice properties (and this same combination appears when we quantize the theory). We look at  $u^{r\dagger}(\not{p}) \cdot v^s(-\not{p})$ , with the 3-momentum in the spinor  $v$  taking the opposite sign. Defining the 4-momentum  $(p^r)^\mu = (p^0, -\not{p})$ , we have

$$\begin{aligned} u^{r\dagger}(\not{p}) \cdot v^s(-\not{p}) &= \xi^{r\dagger} \frac{\sqrt{p^r \cdot \sigma}}{\sqrt{p^r \cdot \bar{\sigma}}} \eta^s \frac{\sqrt{p^r \cdot \sigma \eta^s}}{\sqrt{p^r \cdot \bar{\sigma} \eta^s}} \\ &= \xi^{r\dagger} (p \cdot \sigma)(p^r \cdot \sigma) \eta^s - \xi^{r\dagger} (p \cdot \bar{\sigma})(p^r \cdot \bar{\sigma}) \eta^s \end{aligned} \quad (4.126)$$

Now the terms under the square-root are given by  $(p \cdot \sigma)(p^r \cdot \sigma) = (p_0 + p_i \sigma^i)(p_0 - p_i \sigma^i) = p_0^2 - \not{p}^2 = m^2$ . The same expression holds for  $(p \cdot \bar{\sigma})(p^r \cdot \bar{\sigma})$ , and the two terms cancel. We learn

$$u^{r\dagger}(\not{p}) \cdot v^s(-\not{p}) = v^{r\dagger}(\not{p}) \cdot u^s(-\not{p}) = 0 \quad (4.127)$$

## Outer Products

There's one last spinor identity that we need before we turn to the quantum theory. It is:

### Claim:

$$\sum_{s=1} u^s(\not{p}) \bar{u}^s(\not{p}) = p/+m \quad (4.128)$$

where the two spinors are not now contracted, but instead placed back to back to give a  $4 \times 4$  matrix. Also,

$$\sum_{s=1} v^s(\not{p}) \bar{v}^s(\not{p}) = p/-m \quad (4.129)$$

**Proof:**

$$\sum_{s=1}^{\Sigma} u^s(\not{p}) \bar{u}^s(\not{p}) = \sum_{s=1}^{\Sigma} \frac{\sqrt{p \cdot \sigma} \xi^s}{\sqrt{p \cdot \bar{\sigma}} \xi^s} ! \quad \xi^{s\dagger} \sqrt{\frac{p \cdot \bar{\sigma}}{p \cdot \sigma}}, \xi^{s\dagger} \sqrt{\frac{p \cdot \sigma}{p \cdot \bar{\sigma}}} \quad (4.130)$$

But  $\sum_s \xi^s \xi^{s\dagger} = \mathbf{1}$ , the  $2 \times 2$  unit matrix, which then gives us

$$\sum_{s=1}^{\Sigma} u^s(\not{p}) \bar{u}^s(\not{p}) = \frac{m}{p \cdot \bar{\sigma}} \frac{p \cdot \sigma}{m} ! \quad (4.131)$$

which is the desired result. A similar proof works for  $\sum_s v^s(\not{p}) \bar{v}^s(\not{p})$ .

## 5. Quantizing the Dirac Field

We would now like to quantize the Dirac Lagrangian,

$$L = \bar{\psi}(x) i\cancel{d} - m \psi(x) \quad (5.1)$$

We will proceed naively and treat  $\psi$  as we did the scalar field. But we'll see that things go wrong and we will have to reconsider how to quantize this theory.

### 5.1 A Glimpse at the Spin-Statistics Theorem

We start in the usual way and define the momentum,

$$\pi = \frac{\partial L}{\partial \dot{\psi}} = i\bar{\psi} \gamma^0 = i\psi^\dagger \quad (5.2)$$

For the Dirac Lagrangian, the momentum conjugate to  $\psi$  is  $i\psi^\dagger$ . It does not involve the time derivative of  $\psi$ . This is as it should be for an equation of motion that is first order in time, rather than second order. This is because we need only specify  $\psi$  and  $\psi^\dagger$  on an initial time slice to determine the full evolution.

To quantize the theory, we promote the field  $\psi$  and its momentum  $\psi^\dagger$  to operators, satisfying the canonical commutation relations, which read

$$\begin{aligned} [\psi_\alpha(\mathbf{x}), \psi_\beta(\mathbf{y})] &= [\psi_\alpha^\dagger(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})] = 0 \\ [\psi_\alpha(\mathbf{x}), \psi_\beta^\dagger(\mathbf{y})] &= \delta_{\alpha\beta} \delta^{(3)}(\mathbf{x} - \mathbf{y}) \end{aligned} \quad (5.3)$$

It's this step that we'll soon have to reconsider.

Since we're dealing with a free theory, where any classical solution is a sum of plane waves, we may write the quantum operators as

$$\begin{aligned} \psi(\mathbf{x}) &= \sum_s \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{\frac{h}{2E_p}}} b_{p \rightarrow} u(p) e^{i p \cdot x} + c_{p \rightarrow} v(p) e^{-i p \cdot x} \\ \psi^\dagger(\mathbf{x}) &= \sum_{s=1}^2 \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{\frac{h}{2E_p}}} b_{p \rightarrow}^* u(p) e^{i p \cdot x} + c_{p \rightarrow}^* v(p) e^{-i p \cdot x} \end{aligned} \quad (5.4)$$

where the operators  $b_{p \rightarrow}^{s\dagger}$  create particles associated to the spinors  $u^s(p)$ , while  $c_{p \rightarrow}^{s\dagger}$  create particles associated to  $v^s(p)$ . As with the scalars, the commutation relations of the fields imply commutation relations for the annihilation and creation operators.

**Claim:** The field commutation relations (5.3) are equivalent to

$$\begin{aligned}[b^r_{\vec{p}}, b^{s\dagger}_{\vec{q}}] &= (2\pi)^3 \delta^{rs} \delta^{(3)}(\vec{p} \rightarrow \vec{q}) \\ [c^r_{\vec{p}}, c^{s\dagger}_{\vec{q}}] &= -(2\pi)^3 \delta^{rs} \delta^{(3)}(\vec{p} \rightarrow \vec{q})\end{aligned}\quad (5.5)$$

with all other commutators vanishing.

Note the strange minus sign in the  $[c, c^\dagger]$  term. This means that we can't define the ground state  $|0\rangle$  as something annihilated by  $c^r_{\vec{p}}|0\rangle = 0$ , because then the excited states  $c^{s\dagger}_{\vec{p}}|0\rangle$  would have negative norm. To avoid this, we will have to flip the interpretation of  $c$  and  $c^\dagger$ , with the vacuum defined by  $c^{s\dagger}_{\vec{p}}|0\rangle = 0$  and the excited states by  $c^r_{\vec{p}}|0\rangle$ . This, as we will see, will be our undoing.

**Proof:** Let's show that the  $[b, b^\dagger]$  and  $[c, c^\dagger]$  commutators reproduce the field commutators (5.3),

$$\begin{aligned} [\psi(\vec{x}), \psi^\dagger(\vec{y})] &= \sum_{r,s} \frac{d^3 p}{(2\pi)^3} \frac{d^3 q}{2E_{\vec{p}}} \frac{1}{\sqrt{\frac{1}{4E_{\vec{p}} E_{\vec{q}}}}} [b^r_{\vec{p}}, b^{s\dagger}_{\vec{q}}] u^r(\vec{p}) u^{s\dagger}(\vec{q}) e^{i(\vec{x} \cdot \vec{p} - \vec{y} \cdot \vec{q})} \\ &\quad + [c^{r\dagger}_{\vec{p}}, c^s_{\vec{q}}] v^r(\vec{p}) v^s(\vec{q})^\dagger e^{-i(\vec{x} \cdot \vec{p} - \vec{y} \cdot \vec{q})} \\ &= \sum_s \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} u^s(\vec{p}) \bar{u}^s(\vec{p}) \gamma^0 e^{i\vec{p} \cdot (\vec{x} - \vec{y})} + v^s(\vec{p}) \bar{v}^s(\vec{p}) \gamma^0 e^{-i\vec{p} \cdot (\vec{x} - \vec{y})} \end{aligned}\quad (5.6)$$

At this stage we use the outer product formulae (4.128) and (4.129) which tell us  $\sum_s u^s(\vec{p}) \bar{u}^s(\vec{p}) = p/+m$  and  $\sum_s v^s(\vec{p}) \bar{v}^s(\vec{p}) = p/-m$ , so that

$$\begin{aligned} [\psi(\vec{x}), \psi^\dagger(\vec{y})] &= \int \frac{d^3 p}{(2\pi)^3 2E_{\vec{p}}} \frac{1}{2E_{\vec{p}}} (p/+m) \gamma^0 e^{i\vec{p} \cdot (\vec{x} - \vec{y})} + (p/-m) \gamma^0 e^{-i\vec{p} \cdot (\vec{x} - \vec{y})} \\ &= \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_{\vec{p}}} (p_0 \gamma^0 + p_i \gamma^i + m) \gamma^0 + (p_0 \gamma^0 - p_i \gamma^i - m) \gamma^0 e^{i\vec{p} \cdot (\vec{x} - \vec{y})} \end{aligned}$$

where, in the second term, we've changed  $\vec{p} \rightarrow -\vec{p}$  under the integration sign. Now, using  $p_0 = E_{\vec{p}}$  we have

$$[\psi(\vec{x}), \psi^\dagger(\vec{y})] = \int \frac{d^3 p}{(2\pi)^3} e^{i\vec{p} \cdot (\vec{x} - \vec{y})} = \delta^{(3)}(\vec{x} - \vec{y}) \quad (5.7)$$

as promised. Notice that it's a little tricky in the middle there, making sure that the  $p_i \gamma^i$  terms cancel. This was the reason we needed the minus sign in the  $[c, c^\dagger]$  commutator terms in (5.5).

### 5.1.1 The Hamiltonian

To proceed, let's construct the Hamiltonian for the theory. Using the momentum  $\pi = i\psi^\dagger$ , we have

$$H = \pi\psi^\dagger - L = \bar{\psi}(-i\gamma^i\partial_i + m)\psi \quad (5.8)$$

which means that  $H = \int d^3x H$  agrees with the conserved energy computed using Noether's theorem (4.92). We now wish to turn the Hamiltonian into an operator. Let's firstly look at

$$(-i\gamma^i\partial_i + m)\psi = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} b_{p\rightarrow}^s (-i\gamma^i p_i + m) u(s\vec{p}) e^{+ip\cdot\vec{x}} + c_{p\rightarrow}^s (\gamma^i p_i + m) v(s\vec{p}) e^{-ip\cdot\vec{x}} \quad \mathbf{i}$$

where, for once we've left the sum over  $s = 1, 2$  implicit. There's a small subtlety with the minus signs in deriving this equation that arises from the use of the Minkowski metric in contracting indices, so that  $\not{p} \cdot \not{x} \equiv i x^i p_i = -x^i p_i$ . Now we use the defining equations for the spinors  $u^s(\vec{p})$  and  $v^s(\vec{p})$  given in (4.105) and (4.111), to replace

$$(-i\gamma^i p_i + m) u^s(\vec{p}) = \gamma^0 p_0 u^s(\vec{p}) \text{ and } (\gamma^i p_i + m) v^s(\vec{p}) = -\gamma^0 p_0 v^s(\vec{p}) \quad (5.9)$$

so we can write

$$(-i\gamma^i\partial_i + m)\psi = \int \frac{d^3p}{(2\pi)^3} \frac{1}{2E_p} \gamma^0 b_{p\rightarrow}^s u(s\vec{p}) e^{+ip\cdot\vec{x}} - c_{p\rightarrow}^s v(s\vec{p}) e^{-ip\cdot\vec{x}} \quad \mathbf{i} \quad (5.10)$$

We now use this to write the operator Hamiltonian

$$\begin{aligned} H &= \int d^3x \psi^\dagger \gamma^0 (-i\gamma^i\partial_i + m)\psi \\ &= \frac{\int d^3x d^3p d^3q}{(2\pi)^6} \frac{1}{4E_{\vec{p}}} h \\ &= \frac{1}{(2\pi)^3} \frac{1}{2} b_{p\rightarrow}^{r\dagger} b_{p\rightarrow}^s [u(r\vec{p})^\dagger \cdot u(s\vec{p})] - c_{p\rightarrow}^{r\dagger} c_{p\rightarrow}^s [v(r\vec{p})^\dagger \cdot v(s\vec{p})] \\ &\quad - b_{r\rightarrow}^{s\dagger} b_{r\rightarrow}^s [u(r\vec{p}) \cdot v(-r\vec{p})] + c_{r\rightarrow}^{s\dagger} c_{r\rightarrow}^s [v(r\vec{p}) \cdot u(-r\vec{p})] \end{aligned} \quad \mathbf{i}$$

where, in the last two terms we have relabelled  $\not{p} \rightarrow -\not{p}$ . We now use our inner product formulae (4.122), (4.124) and (4.127) which read

$$u^r(\vec{p})^\dagger \cdot u^s(\vec{p}) = v^r(\vec{p})^\dagger \cdot v^s(\vec{p}) = 2p_0 \delta^{rs} \quad \text{and} \quad u^r(\vec{p})^\dagger \cdot v^s(-\vec{p}) = v^r(\vec{p})^\dagger \cdot u^s(-\vec{p}) = 0$$

giving us

$$H = \int \frac{d^3 p}{(2\pi)^3} E_{p \rightarrow} b^{s\dagger} b^s - c^s c^{s\dagger} \quad (5.11)$$

$$= \frac{d^3 p}{(2\pi)^3} E_{p \rightarrow} b^{s\dagger} b^s - c^{s\dagger} c^s + (2\pi)^3 \delta^{(3)}(0) \quad (5.12)$$

The  $\delta^{(3)}$  term is familiar and easily dealt with by normal ordering. The  $b^\dagger b$  term is familiar and we can check that  $b^\dagger$  create positive energy states as expected,

$$[H, b^{s\dagger}] = E_{p \rightarrow} b^{s\dagger}$$

The minus sign in front of the  $c^\dagger c$  term should make us nervous. If we think of  $c^\dagger$  as creation operators then there's no problem since, using the commutation relation (5.5), we still find that  $c^\dagger$  creates positive energy states,

$$[H, c^{s\dagger}] = E_{p \rightarrow} c^{s\dagger}$$

However, as we noted after (5.5), these states have negative norm. To have a sensible Hilbert space, we need to interpret  $c$  as the creation operator. But then the Hamiltonian is not bounded below because

$$[H, c_s] = -E_{p \rightarrow} c_s$$

This is a disaster. Taken seriously it would tell us that we could tumble to states of lower and lower energy by continually producing  $c$  particles. As the English would say, it's all gone a bit Pete Tong. (No relation).

Since the above calculation was a little tricky, you might think that it's possible to rescue the theory to get the minus signs to work out right. You can play around with different things, but you'll always find this minus sign cropping up somewhere. And, in fact, it's telling us something important that we missed.

## 5.2 Fermionic Quantization

The key piece of physics that we missed is that spin 1/2 particles are fermions, meaning that they obey Fermi-Dirac statistics with the quantum state picking up a minus sign upon the interchange of any two particles. This fact is embedded into the structure of relativistic quantum field theory: the spin-statistics theorem says that integer spin fields must be quantized as bosons, while half-integer spin fields must be quantized as fermions. Any attempt to do otherwise will lead to an inconsistency, such as the unbounded Hamiltonian we saw in (5.12).

So how do we go about quantizing a field as a fermion? Recall that when we quantized the scalar field, the resulting particles obeyed bosonic statistics because the creation and annihilation operators satisfied the commutation relations,

$$[a_{p\rightarrow}^\dagger, a_{q\rightarrow}^\dagger] = 0 \Rightarrow a_{p\rightarrow}^\dagger a_{q\rightarrow}^\dagger |0\rangle \equiv |\not{p}, \not{q}\rangle = |\not{q}, p\not{\rightarrow}\rangle \quad (5.13)$$

To have states obeying fermionic statistics, we need anti-commutation relations,  $\{A, B\} \equiv AB + BA$ . Rather than (5.3), we will ask that the spinor fields satisfy

$$\begin{aligned} \{\psi_\alpha(\not{x}), \psi_\beta(\not{y})\} &= \{\psi_\alpha(\not{x}), \psi_\beta(\not{y})^\dagger\} = 0 \\ \{\psi_\alpha(\not{x}), \psi_\beta^\dagger(\not{y})\} &= \delta_{\alpha\beta} \delta^{(3)}(\not{x} - \not{y}) \end{aligned} \quad (5.14)$$

We still have the expansion (5.4) of  $\psi$  and  $\psi^\dagger$  in terms of  $b, b^\dagger, c$  and  $c^\dagger$ . But now the same proof that led us to (5.5) tells us that

$$\begin{aligned} \{b_{p\rightarrow}^r, b_{q\rightarrow}^{s\dagger}\} &= (2\pi)^3 \delta_r^s \delta^{(3)}(\not{p} - \not{q}) \\ \{c_{p\rightarrow}^r, c_{q\rightarrow}^{s\dagger}\} &= (2\pi)^3 \delta_r^s \delta^{(3)}(\not{p} - \not{q}) \end{aligned} \quad (5.15)$$

with all other *anti-commutators* vanishing,

$$\{b_{p\rightarrow}^r, b_{q\rightarrow}^s\} = \{c_{p\rightarrow}^r, c_{q\rightarrow}^s\} = \{b_{p\rightarrow}^r, c_{q\rightarrow}^s\} = \{b_{p\rightarrow}^r, c_{q\rightarrow}^{s\dagger}\} = \dots = 0 \quad (5.16)$$

The calculation of the Hamiltonian proceeds as before, all the way through to the penultimate line (5.11). At that stage, we get

$$\begin{aligned} H &= \int \frac{d^3 p}{(2\pi)^3} E_{p\rightarrow}^h b_{p\rightarrow}^{s\dagger} b_{p\rightarrow}^s - c_{p\rightarrow}^s c_{p\rightarrow}^{s\dagger} \mathbf{i} \\ &= \frac{d^3 p}{(2\pi)^3} E_{p\rightarrow}^h b_{p\rightarrow}^{s\dagger} b_{p\rightarrow}^s + c_{p\rightarrow}^{s\dagger} c_{p\rightarrow}^s - (2\pi)^3 \delta^{(3)}(0) \mathbf{i} \end{aligned} \quad (5.17)$$

The anti-commutators have saved us from the indignity of a Hamiltonian unbounded below. Note that when normal ordering the Hamiltonian we now throw away a negative contribution  $-(2\pi)^3 \delta^{(3)}(0)$ . In principle, this could partially cancel the positive contribution from bosonic fields. Cosmological constant problem anyone?!

### 5.2.1 Fermi-Dirac Statistics

Just as in the bosonic case, we define the vacuum  $|0\rangle$  to satisfy,

$$b_{p\rightarrow}^s |0\rangle = c_{p\rightarrow}^s |0\rangle = 0 \quad (5.18)$$

Although  $b$  and  $c$  obey anti-commutation relations, the Hamiltonian (5.17) has nice commutation relations with them. You can check that

$$\begin{aligned} [H, b^r] &= -E_{\vec{p}} b^r \quad \text{and} \quad [H, b^r\dagger] = E_{\vec{p}} b^r\dagger \\ [H, c^r_{\vec{p}}] &= -E_{\vec{p}} c^r_{\vec{p}} \quad \text{and} \quad [H, c^r\dagger_{\vec{p}}] = E_{\vec{p}} c^r\dagger_{\vec{p}} \end{aligned} \quad (5.19)$$

This means that we can again construct a tower of energy eigenstates by acting on the vacuum by  $b^r\dagger$  and  $c^r\dagger$  to create particles and antiparticles, just as in the bosonic case. For example, we have the one-particle states

$$|\vec{p}, r\rangle = b^r\dagger |\vec{0}\rangle \quad (5.20)$$

The two particle states now satisfy

$$|\vec{p}_1, r_1; \vec{p}_2, r_2\rangle \equiv b^r_{\vec{p}_1} b^r_{\vec{p}_2} |\vec{0}\rangle = -|\vec{p}_2, r_2; \vec{p}_1, r_1\rangle \quad (5.21)$$

confirming that the particles do indeed obey Fermi-Dirac statistics. In particular, we have the Pauli-Exclusion principle  $|\vec{p}, r; \vec{p}, r\rangle = 0$ . Finally, if we wanted to be sure about the spin of the particle, we could act with the angular momentum operator (4.96) to confirm that a stationary particle  $|\vec{p} = 0, r\rangle$  does indeed carry intrinsic angular momentum 1/2 as expected.

### 5.3 Dirac's Hole Interpretation

“In this attempt, the success seems to have been on the side of Dirac rather than logic”

*Pauli on Dirac*

Let’s pause our discussion to make a small historical detour. Dirac originally viewed his equation as a relativistic version of the Schrödinger equation, with  $\psi$  interpreted as the wavefunction for a single particle with spin. To reinforce this interpretation, he wrote  $(i\partial/\partial t - m)\psi = 0$  as

$$i\frac{\partial\psi}{\partial t} = -i\alpha\cdot\vec{\nabla}\psi + m\beta\psi \equiv \hat{H}\psi \quad (5.22)$$

where  $\alpha = -\gamma_0 \vec{\gamma}$  and  $\beta = \gamma_0$ . Here the operator  $\hat{H}$  is interpreted as the one-particle

Hamiltonian. This is a very different viewpoint from the one we now have, where  $\psi$  is a classical field that should be quantized. In Dirac’s view, the Hamiltonian of the system is  $\hat{H}$  defined above, while for us the Hamiltonian is the field operator (5.17). Let’s see where Dirac’s viewpoint leads.

With the interpretation of  $\psi$  as a single-particle wavefunction, the plane-wave solutions (4.104) and (4.110) to the Dirac equation are thought of as energy eigenstates, with

$$\begin{aligned}\psi &= u(\vec{p}) e^{-ip \cdot x} & \Rightarrow i \frac{\partial \psi}{\partial p} &= E_{\vec{p}} \psi \\ \psi &= v(\vec{p}) e^{+ip \cdot x} & \Rightarrow i \frac{\partial \psi}{\partial t} &= -E_{\vec{p}} \psi\end{aligned}\quad (5.23)$$

which look like positive and negative energy solutions. The spectrum is once again unbounded below; there are states  $v(\vec{p})$  with arbitrarily low energy  $-E_{\vec{p}}$ . At first glance this is disastrous, just like the unbounded field theory Hamiltonian (5.12). Dirac postulated an ingenious solution to this problem: since the electrons are fermions (a fact which is put in by hand to Dirac's theory) they obey the Pauli-exclusion principle. So we could simply stipulate that in the true vacuum of the universe, all the negative energy states are filled. Only the positive energy states are accessible. These filled negative energy states are referred to as the *Dirac sea*. Although you might worry about the infinite negative charge of the vacuum, Dirac argued that only charge differences would be observable (a trick reminiscent of the normal ordering prescription we used for field operators).

Having avoided disaster by floating on an infinite sea comprised of occupied negative energy states, Dirac realized that his theory made a shocking prediction. Suppose that a negative energy state is excited to a positive energy state, leaving behind a hole. The hole would have all the properties of the electron, except it would carry positive charge. After flirting with the idea that it may be the proton, Dirac finally concluded that the hole is a new particle: the positron. Moreover, when a positron comes across an electron, the two can annihilate. Dirac had predicted anti-matter, one of the greatest achievements of theoretical physics. It took only a couple of years before the positron was discovered experimentally in 1932.

Although Dirac's physical insight led him to the right answer, we now understand that the interpretation of the Dirac spinor as a single-particle wavefunction is not really correct. For example, Dirac's argument for anti-matter relies crucially on the particles being fermions while, as we have seen already in this course, anti-particles exist for both fermions and bosons. What we really learn from Dirac's analysis is that there is no consistent way to interpret the Dirac equation as describing a single particle. It is instead to be thought of as a classical field which has only positive energy solutions because the Hamiltonian (4.92) is positive definite. Quantization of this field then gives rise to both particle and anti-particle excitations.

This from Julian Schwinger:

*"Until now, everyone thought that the Dirac equation referred directly to physical particles. Now, in field theory, we recognize that the equations refer to a sublevel. Experimentally we are concerned with particles, yet the old equations describe fields.... When you begin with field equations, you operate on a level where the particles are not there from the start. It is when you solve the field equations that you see the emergence of particles."*

## 5.4 Propagators

Let's now move to the Heisenberg picture. We define the spinors  $\psi(x, t)$  at every point in spacetime such that they satisfy the operator equation

$$\frac{\partial \psi}{\partial t} = i[H, \psi] \quad (5.24)$$

We solve this by the expansion

$$\begin{aligned} \psi(x) &= \sum_{s=1}^2 \int d^3 p \frac{1}{(2\pi)^3} \frac{h}{2E_p} b_{p \rightarrow} u(p) e^{-ip \cdot x} + c_{p \rightarrow} v(p) e^{+ip \cdot x} \\ \psi^\dagger(x) &= \sum_{s=1}^2 \int d^3 p \frac{1}{(2\pi)^3} \frac{h}{2E_p} b_{p \rightarrow}^* u(p) e^{+ip \cdot x} + c_{p \rightarrow}^* v(p) e^{-ip \cdot x} \end{aligned} \quad (5.25)$$

Let's now look at the anti-commutators of these fields. We define the fermionic propagator to be

$$iS_{\alpha\beta} = \{\psi_\alpha(x), \bar{\psi}_\beta(y)\} \quad (5.26)$$

In what follows we will often drop the indices and simply write  $iS(x-y) = \{\psi(x), \bar{\psi}(y)\}$ , but you should remember that  $S(x-y)$  is a  $4 \times 4$  matrix. Inserting the expansion (5.25), we have

$$\begin{aligned} iS(x-y) &= \int \frac{d^3 p d^3 q}{(2\pi)^6} \frac{1}{4E_p E_q} h \{b_s^s, b_{q \rightarrow}^{r \dagger}\} u(p) u^r(q) e^{-i(p \cdot x - q \cdot y)} \\ &\quad + \{c_{p \rightarrow}^s, c_{q \rightarrow}^r\} v^s(p) \bar{v}^r(q) e^{+i(p \cdot x - q \cdot y)} \\ &= \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_p} u(p) \bar{u}(p) e^{-ip \cdot (x-y)} + v(p) \bar{v}(p) e^{+ip \cdot (x-y)} \\ &= \frac{(p/+m)e^{-ip \cdot (x-y)} + (p/-m)e^{+ip \cdot (x-y)}}{(2\pi)^3 2E_p} \end{aligned} \quad (5.27)$$

where to reach the final line we have used the outer product formulae (4.128) and (4.129). We can then write

$$iS(x - y) = (i\partial_x + m)(D(x - y) - D(y - x)) \quad (5.28)$$

in terms of the propagator for a real scalar field  $D(x - y)$  which, recall, can be written as (2.90)

$$D(x - y) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_p} e^{-ip \cdot (x-y)} \quad (5.29)$$

Some comments:

- For spacelike separated points  $(x - y)^2 < 0$ , we have already seen that  $D(x - y) - D(y - x) = 0$ . In the bosonic theory, we made a big deal of this since it ensured that

$$[\varphi(x), \varphi(y)] = 0 \quad (x - y)^2 < 0 \quad (5.30)$$

outside the lightcone, which we trumpeted as proof that our theory was causal. However, for fermions we now have

$$\{\psi_\alpha(x), \psi_\beta(y)\} = 0 \quad (x - y)^2 < 0 \quad (5.31)$$

outside the lightcone. What happened to our precious causality? The best that we can say is that all our observables are bilinear in fermions, for example the Hamiltonian (5.17). These still commute outside the lightcone. The theory remains causal as long as fermionic operators are not observable. If you think this is a little weak, remember that no one has ever seen a physical measuring apparatus come back to minus itself when you rotate by 360 degrees!

- At least away from singularities, the propagator satisfies

$$(i\partial_x - m)S(x - y) = 0 \quad (5.32)$$

which follows from the fact that  $(\partial_x^2 + m^2)D(x - y) = 0$  using the mass shell condition  $p^2 = m^2$ .

## 5.5 The Feynman Propagator

By a similar calculation to that above, we can determine the vacuum expectation value,

$$\begin{aligned} \langle 0 | \psi_\alpha(x) \bar{\psi}_\beta(y) | 0 \rangle &= \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_p} (p/+m)_{\alpha\beta} e^{-ip \cdot (x-y)} \\ \langle 0 | \bar{\psi}_\beta(y) \psi_\alpha(x) | 0 \rangle &= \int \frac{d^3 p}{(2\pi)^3} \frac{1}{2E_p} (p/-m)_{\alpha\beta} e^{+ip \cdot (x-y)} \end{aligned} \quad (5.33)$$

We now define the Feynman propagator  $S_F(x - y)$ , which is again a  $4 \times 4$  matrix, as the time ordered product,

$$S_F(x - y) = \langle 0 | T\psi(x)\bar{\psi}(y) | 0 \rangle \equiv \begin{cases} \langle 0 | \psi(x)\bar{\psi}(y) | 0 \rangle & x^0 > y^0 \\ \langle 0 | -\bar{\psi}(y)\psi(x) | 0 \rangle & y^0 > x^0 \end{cases} \quad (5.34)$$

Notice the minus sign! It is necessary for Lorentz invariance. When  $(x-y)^2 < 0$ , there is no invariant way to determine whether  $x^0 > y^0$  or  $y^0 > x^0$ . In this case the minus sign is necessary to make the two definitions agree since  $\{\psi(x), \bar{\psi}(y)\} = 0$  outside the lightcone. We have the 4-momentum integral representation for the Feynman propagator,

$$S_F(x - y) = i \int \frac{d^4 p}{(2\pi)^4} e^{-ip \cdot (x-y)} \frac{\gamma \cdot p + m}{p^2 - m^2 + i\epsilon} \quad (5.35)$$

which satisfies  $(i\partial_x - m)S_F(x - y) = i\delta^{(4)}(x - y)$ , so that  $S_F$  is a Green's function for the Dirac operator.

The minus sign that we see in (5.34) also occurs for any string of operators inside a time ordered product  $T(\dots)$ . While bosonic operators commute inside  $T$ , fermionic operators anti-commute. We have this same behaviour for normal ordered products as well, with fermionic operators obeying  $:\psi_1\psi_2: = - :\psi_2\psi_1::$ . With the understanding that all fermionic operators anti-commute inside  $T$  and ::, Wick's theorem proceeds just as in the bosonic case. We define the contraction

$$\overline{\psi(x)\bar{\psi}(y)} = T(\psi(x)\bar{\psi}(y)) - :\psi(x)\bar{\psi}(y):: = S_F(x - y) \quad (5.36)$$

## 5.6 Yukawa Theory

The interaction between a Dirac fermion of mass  $m$  and a real scalar field of mass  $\mu$  is governed by the Yukawa theory,

$$L = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - \frac{1}{2}\mu^2\varphi^2 + \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \lambda\varphi\bar{\psi}\psi \quad (5.37)$$

which is the proper version of the baby scalar Yukawa theory we looked at in Section 3. Couplings of this type appear in the standard model, between fermions and the Higgs boson. In that context, the fermions can be leptons (such as the electron) or quarks.

Yukawa originally proposed an interaction of this type as an effective theory of nuclear forces. With an eye to this, we will again refer to the  $\varphi$  particles as mesons, and the  $\psi$  particles as nucleons. Except, this time, the nucleons have spin. (This is still not a particularly realistic theory of nucleon interactions, not least because we're omitting isospin. Moreover, in Nature the relevant mesons are pions which are pseudoscalars, so a coupling of the form  $\varphi\bar{\psi}\gamma^5\psi$  would be more appropriate. We'll turn to this briefly in Section 5.7.3).

Note the dimensions of the various fields. We still have  $[\varphi] = 1$ , but the kinetic terms require that  $[\psi] = 3/2$ . Thus, unlike in the case with only scalars, the coupling is dimensionless:  $[\lambda] = 0$ .

We'll proceed as we did in Section 3, firstly computing the amplitude of a particular scattering process then, with that calculation as a guide, writing down the Feynman rules for the theory. We start with:

### 5.6.1 An Example: Putting Spin on Nucleon Scattering

Let's study  $\psi\psi \rightarrow \psi\psi$  scattering. This is the same calculation we performed in Section (3.3.3) except now the fermions have spin. Our initial and final states are

$$|i\rangle = \sqrt{\frac{4E_p E_q}{4E_{p'} E_{q'}}} b_{p\rightarrow}^{s\dagger} b_{q\rightarrow}^{r\dagger} |0\rangle \equiv |p\rightarrow, s; q\rightarrow, r\rangle$$

$$|f\rangle = \sqrt{\frac{4E_{p'} E_{q'}}{4E_p E_q}} b_{p'\rightarrow}^r b_{q'\rightarrow}^s |0\rangle \equiv |p'\rightarrow, r; q'\rightarrow, s\rangle \quad (5.38)$$

We need to be a little cautious about minus signs, because the  $b^\dagger$ 's now anti-commute. In particular, we should be careful when we take the adjoint. We have

$$\langle f | = \sqrt{\frac{4E_p E_q}{4E_{p'} E_{q'}}} \langle 0 | b_{q\rightarrow}^r b_{p\rightarrow}^s \quad (5.39)$$

We want to calculate the order  $\lambda^2$  terms from the S-matrix element  $\langle f | S - 1 | i \rangle$ .

$$\frac{(-i\lambda)^2}{2} \int d^4x_1 d^4x_2 T \bar{\psi}(x_1)\psi(x_1)\varphi(x_1) \bar{\psi}(x_2)\psi(x_2)\varphi(x_2) \quad (5.40)$$

where, as usual, all fields are in the interaction picture. Just as in the bosonic calculation, the contribution to nucleon scattering comes from the contraction

$$:\bar{\psi}(x_1)\psi(x_1)\bar{\psi}(x_2)\psi(x_2):\varphi(x_1)\varphi(x_2) \quad (5.41)$$

We just have to be careful about how the spinor indices are contracted. Let's start by looking at how the fermionic operators act on  $|i\rangle$ . We expand out the  $\psi$  fields, leaving the  $\bar{\psi}$  fields alone for now. We may ignore the  $c^\dagger$  pieces in  $\psi$  since they give no contribution at order  $\lambda^2$ . We have

$$:\bar{\psi}(x_1)\psi(x_1)\bar{\psi}(x_2)\psi(x_2): b_{p\rightarrow}^{s\dagger} b_{q\rightarrow}^{r\dagger} |0\rangle = \int \frac{d^3k_1 d^3k_2}{(2\pi)^6} [\bar{\psi}(x_1) \cdot u^m(k_1)] [\bar{\psi}(x_2) \cdot u^n(k_2)] e^{-ik_1 \cdot x_1 - ik_2 \cdot x_2} \frac{4E_{k_1} E_{k_2}}{4E_{p\rightarrow} E_{q\rightarrow}} b_{k_1}^m b_{k_2}^n b_{p\rightarrow}^{s\dagger} b_{q\rightarrow}^{r\dagger} |0\rangle \quad (5.42)$$

where we've used square brackets  $[\cdot]$  to show how the spinor indices are contracted. The minus sign that sits out front came from moving  $\psi(x_1)$  past  $\bar{\psi}(x_2)$ . Now anti-commuting the  $b$ 's past the  $b^\dagger$ 's, we get

$$= \frac{1}{2 \sqrt{E_p \rightarrow E_{q \rightarrow}}} [\bar{\psi}(x_1) \cdot u^r(\not{q})] [\bar{\psi}(x_2) \cdot u^s(\not{p})] e^{-ip \cdot x_2 - iq \cdot x_1} \\ - [\bar{\psi}(x_1) \cdot u^s(\not{p})] [\bar{\psi}(x_2) \cdot u^r(\not{q})] e^{-ip \cdot x_1 - iq \cdot x_2} |0\rangle \quad (5.43)$$

Note, in particular, the relative minus sign that appears between these two terms. Now let's see what happens when we hit this with  $\langle f |$ . We look at

$$\langle 0 | b_q^{r'} \bar{b}_{p \rightarrow}^{s'} [\bar{\psi}(x_1) \cdot u^r(\not{q})] [\bar{\psi}(x_2) \cdot u^s(\not{p})] |0\rangle = \frac{e^{+ip' \cdot x_1 + iq' \cdot x_2}}{2 \sqrt{E_p \rightarrow E_{q \rightarrow}}} [\bar{u}^{s'}(\not{p}) \cdot u^r(\not{q})] [\bar{u}^{r'}(\not{q}) \cdot u^s(\not{p})] \\ - \frac{e^{+ip' \cdot x_2 + iq' \cdot x_1}}{2 \sqrt{E_p \rightarrow E_{q \rightarrow}}} [\bar{u}^{r'}(\not{q}) \cdot u^s(\not{q})] [\bar{u}^{s'}(\not{p}) \cdot u^r(\not{p})]$$

The  $[\bar{\psi}(x_1) \cdot u^s(\not{p})] [\bar{\psi}(x_2) \cdot u^r(\not{q})]$  term in (5.43) ~~double~~ ~~cancel~~ up with this, cancelling the factor of  $1/2$  in front of (5.40). Meanwhile, the  $1/E$  terms cancel the relativistic state normalization. Putting everything together, we have the following expression for  $\langle f | S - 1 | i \rangle$

$$(-i\lambda)^2 \int \frac{d^4x_1 d^4x_2 d^4k}{(2\pi)^4} \frac{ie^{ik \cdot (x_1 - x_2)}}{k^2 - \mu^2 + i\epsilon} [\bar{u}^{s'}(\not{p}) \cdot u^s(\not{p})] [\bar{u}^{r'}(\not{q}) \cdot u^r(\not{q})] e^{ix_1 \cdot (q' - q) + ix_2 \cdot (p' - p)} \\ - [\bar{u}^{s'}(\not{p}) \cdot u^s(\not{p})] [\bar{u}^{r'}(\not{q}) \cdot u^r(\not{q})] e^{ix_1 \cdot (p' - q) + ix_2 \cdot (q' - p)}$$

where we've put the  $\varphi$  propagator back in. Performing the integrals over  $x_1$  and  $x_2$ , this becomes,

$$\int d^4k \frac{(2\pi)^4 i (-i\lambda)^2}{k^2 - \mu^2 + i\epsilon} [\bar{u}^{s'}(\not{p}) \cdot u^s(\not{p})] [\bar{u}^{r'}(\not{q}) \cdot u^r(\not{q})] \delta^{(4)}(q^r - q + k) \delta^{(4)}(p^r - p - k) \\ - [\bar{u}^{s'}(\not{p}) \cdot u^s(\not{p})] [\bar{u}^{r'}(\not{q}) \cdot u^r(\not{q})] \delta^{(4)}(p_r - q + k) \delta^{(4)}(q_r - p - k)$$

And we're almost there! Finally, writing the S-matrix element in terms of the amplitude in the usual way,  $\langle f | S - 1 | i \rangle = iA(2\pi)^4 \delta^{(4)}(p + q - p^r - q^r)$ , we have

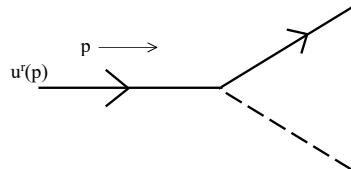
$$A = (-i\lambda)^2 \frac{[\bar{u}^{s'}(\not{p}) \cdot u^s(\not{p})] [\bar{u}^{r'}(\not{q}) \cdot u^r(\not{q})]}{(p^r - p)^2 - \mu^2 + i\epsilon} - \frac{[\bar{u}^{s'}(\not{p}) \cdot u^s(\not{p})] [\bar{u}^{r'}(\not{q}) \cdot u^r(\not{q})]}{(q^r - p)^2 - \mu^2 + i\epsilon}$$

which is our final answer for the amplitude.

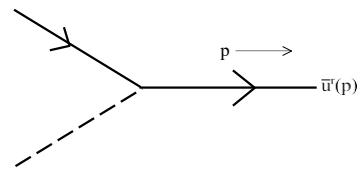
## 5.7 Feynman Rules for Fermions

It's important to bear in mind that the calculation we just did kind of blows. Thankfully the Feynman rules will once again encapsulate the combinatoric complexities and make life easier for us. The rules to compute amplitudes are the following

- To each incoming fermion with momentum  $p$  and spin  $r$ , we associate a spinor  $u^r(p)$ . For outgoing fermions we associate  $u^{-r}(p \rightarrow)$ .

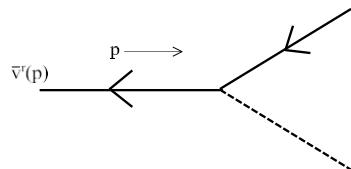


**Figure 21:** An incoming fermion

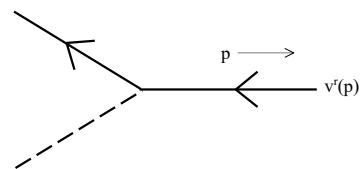


**Figure 22:** An outgoing fermion

- To each incoming anti-fermion with momentum  $p$  and spin  $r$ , we associate a spinor  $\bar{v}^r(p \rightarrow)$ . For outgoing anti-fermions we associate  $v^r(p)$ .



**Figure 23:** An incoming anti-fermion



**Figure 24:** An outgoing anti-fermion

- Each vertex gets a factor of  $-i\lambda$ .
- Each internal line gets a factor of the relevant propagator.

$\overline{\text{---}} \quad p \longrightarrow$	$\frac{i}{p^2 - \mu^2 + i\epsilon} \quad \text{for scalars}$	$\overline{\text{---}} \quad p \longrightarrow$	$\frac{i(p/+ m)}{p^2 - m^2 + i\epsilon} \quad \text{for fermions}$
---	--	---	--

(5.44)

The arrows on the fermion lines must flow consistently through the diagram (this ensures fermion number conservation). Note that the fermionic propagator is a  $4 \times 4$  matrix. The matrix indices are contracted at each vertex, either with further propagators, or with external spinors  $u, u^-, v$  or  $v^-$ .

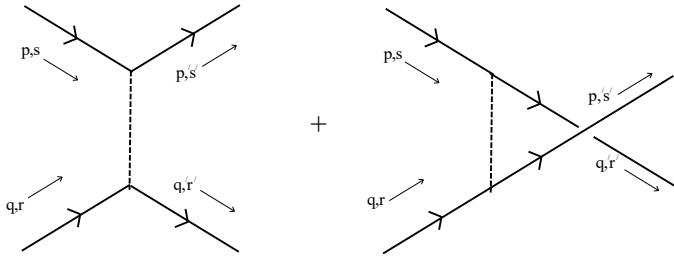
- Impose momentum conservation at each vertex, and integrate over undetermined loop momenta.
- Add extra minus signs for statistics. Some examples will be given below.

### 5.7.1 Examples

Let's run through the same examples we did for the scalar Yukawa theory. Firstly, we have

#### Nucleon Scattering

For the example we worked out previously, the two lowest order Feynman diagrams are shown in Figure 25. We've drawn the second Feynman diagram with the legs crossed



**Figure 25:** The two Feynman diagrams for nucleon scattering

to emphasize the fact that it picks up a minus sign due to statistics. (Note that the way the legs point in the Feynman diagram doesn't tell us the direction in which the particles leave the scattering event: the momentum label does that. The two diagrams above are different because the incoming legs are attached to different outgoing legs). Using the Feynman rules we can read off the amplitude.

$$A = (-i\lambda)^2 \frac{[\bar{u}(p) \cdot u(p)] [\bar{u}(\not{q}) \cdot u(\not{q})]}{(p - p^r)^2 - \mu^2} + \frac{[\bar{u}(p) \cdot u(p)] [\bar{u}(\not{q}) \cdot u(\not{q})]}{(p - q^r)^2 - \mu^2} \quad (5.45)$$

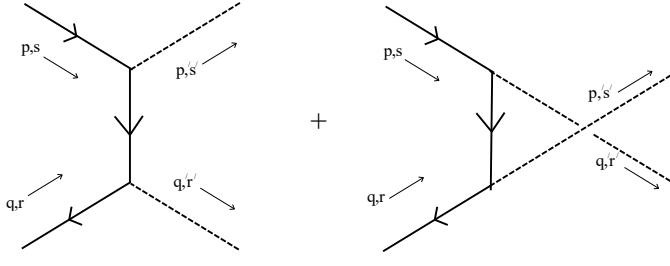
The denominators in each term are due to the meson propagator, with the momentum determined by conservation at each vertex. This agrees with the amplitude we computed earlier using Wick's theorem.

#### Nucleon to Meson Scattering

Let's now look at  $\psi^- \rightarrow \varphi\varphi$ . The two lowest order Feynman diagrams are shown in Figure 26. Applying the Feynman rules, we have

$$A = (-i\lambda)^2 \frac{\bar{v}^l(\not{q}) [\gamma^\mu (p_\mu - p^r) + m] u^s(p)}{(p - p^r)^2 - m^2} + \frac{\bar{v}^r(\not{q}) [\gamma^\mu (p_\mu - q^r) + m] u^s(p)}{(p - q^r)^2 - m^2}$$

Since the internal line is now a fermion, the propagator contains  $\gamma_\mu (p_\mu - p^r) + m$  factors. This is a  $4 \times 4$  matrix which sits on the top, sandwiched between the two external spinors. Now the exchange statistics applies to the final meson states. These are bosons and, correspondingly, there is no relative minus sign between the two diagrams.



**Figure 26:** The two Feynman diagrams for nucleon to meson scattering

### Nucleon-Anti-Nucleon Scattering

For  $\psi\bar{\psi} \rightarrow \psi\bar{\psi}$ , the two lowest order Feynman diagrams are of two distinct types, just like in the bosonic case. They are shown in Figure 27.

The corresponding amplitude is given by,

$$A = (-i\lambda)^2 \frac{[\bar{u}(s', r) \cdot u(s, r)] [\bar{v}(r') \cdot v(r)]}{(p - p^r)^2 - \mu^2} + \frac{[\bar{v}(r) \cdot u(s, r)] [\bar{u}(s', r') \cdot v(r')]}{(p + q)^2 - \mu^2 + i\epsilon} \quad (5.46)$$

As in the bosonic diagrams, there is again the difference in the momentum dependence in the denominator. But now the difference in the diagrams is also reflected in the spinor contractions in the numerator.

More subtle are the minus signs. The fermionic statistics mean that the first diagram has an extra minus sign relative to the  $\psi\psi$  scattering of Figure 25. Since this minus sign will be important when we come to figure out whether the Yukawa force is attractive or repulsive, let's go back to basics and see where it comes from. The initial and final states for this scattering process are

$$|i\rangle = \sqrt{\frac{4E_p E_q}{s' + r' + r}} b_{p\rightarrow}^{s\dagger} c_{q\rightarrow}^{r\dagger} |0\rangle \equiv |p, s; q, r\rangle$$

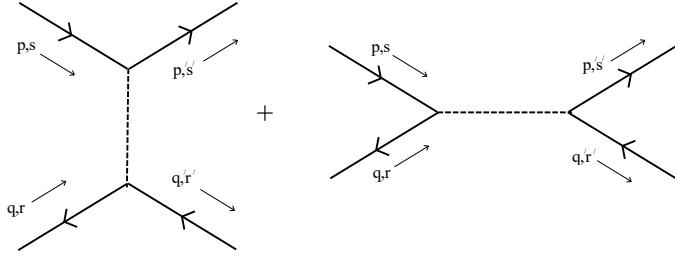
$$|f\rangle = \sqrt{\frac{4E_p E_q}{s' + r' + r}} b_{p\rightarrow}^{r} c_{q\rightarrow}^{r} |0\rangle \equiv |p\rightarrow, s; q, r\rangle \quad (5.47)$$

The ordering of  $b^\dagger$  and  $c^\dagger$  in these states is crucial and reflects the scattering  $\psi\bar{\psi} \rightarrow \psi\bar{\psi}$ , as opposed to  $\psi\bar{\psi} \rightarrow \bar{\psi}\psi$  which would differ by a minus sign. The first diagram in Figure 27 comes from the term in the perturbative expansion,

$$\langle f | : \bar{\psi}(x_1) \psi(x_1) \bar{\psi}(x_2) \psi(x_2) : b_{p\rightarrow}^{s\dagger} c_{q\rightarrow}^{r\dagger} | 0 \rangle \sim \langle f | [\bar{v}^m(\rightarrow k_1) \cdot \psi(x_1)] [\bar{\psi}(x_2) \cdot u^n(k_2)] c_{k\rightarrow}^m b_{k\rightarrow}^{n\dagger} c_{q\rightarrow}^{r\dagger} | 0 \rangle$$

where we've neglected a bunch of objects in this equation like  $d^4 k_i$  and exponential factors because we only want to keep track of the minus signs. Moving the annihilation operators past the creation operators, we have

$$+ \langle f | [\bar{v}^r(\rightarrow q) \cdot \psi(x_1)] [\bar{\psi}(x_2) \cdot u^s(p)] | 0 \rangle \quad (5.48)$$



**Figure 27:** The two Feynman diagrams for nucleon-anti-nucleon scattering

Repeating the process by expanding out the  $\psi(x_1)$  and  $\bar{\psi}(x_2)$  fields and moving them to the left to annihilate  $\langle f |$ , we have

$$\langle 0 | c_{q \rightarrow}^r b_{p \rightarrow}^{s'} c_{\bar{q}_1}^{m \dagger} c_{\bar{q}_2}^{n \dagger} [\bar{v}^r(\not{q})]_1 [\not{v}^m(\not{q})]_2 [\bar{u}^n(\not{p})]_1 [u^s(\not{p})]_2 | 0 \rangle \sim -[\bar{v}^r(\not{q})]_1 [\not{v}^m(\not{q})]_2 [\bar{u}^n(\not{p})]_1 [u^s(\not{p})]_2$$

where the minus sign has appeared from anti-commuting  $c_{l_1}^{m \dagger}$  past  $b_{p \rightarrow}^{s'}$ . This is the overall minus sign found in (5.46). One can also follow similar contractions to compute the second diagram in Figure 27.

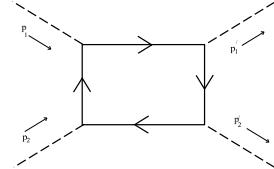
### Meson Scattering

Finally, we can also compute the scattering of  $\varphi\varphi \rightarrow \varphi\varphi$  which, as in the bosonic case, picks up its leading contribution at one-loop. The amplitude for the diagram shown in the figure is

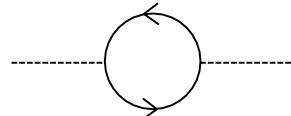
$$iA = -(-i\lambda) \frac{\int d^4k}{(2\pi)^4} \text{Tr} \frac{k/m}{(k^2 - m^2 + i\epsilon)((k+p^r)^2 - m^2 + i\epsilon)} \times \frac{k/p^r + m}{k/p^r - p_1/m} \times \frac{k/p^r + m}{\int ((k+p^r - p_1)^2 - m^2 + i\epsilon)((k-p_2)^2 - m^2 + i\epsilon)}$$

Notice that the high momentum limit of the integral is  $d^4k/k^4$ , which is no longer finite, but diverges logarithmically. You will have to wait until next term to make sense of this integral.

There's an overall minus sign sitting in front of this amplitude. This is a generic feature of diagrams with fermions running in loops: each fermionic loop in a diagram gives rise to an extra minus sign. We can see this rather simply in the diagram



**Figure 28:**



which involves the expression

$$\overline{\psi_\alpha(x)} \overline{\psi_\alpha(x)} \bar{\psi}_\beta(y) \psi_\beta(y) = -\bar{\psi}_\beta(y) \bar{\psi}_\alpha(x) \psi_\alpha(x) \bar{\psi}_\beta(y) \\ = -\text{Tr}(S_F(y-x) S_F(x-y))$$

After passing the fermion fields through each other, a minus sign appears, sitting in front of the two propagators.

### 5.7.2 The Yukawa Potential Revisited

We saw in Section 3.5.2, that the exchange of a real scalar particle gives rise to a universally attractive Yukawa potential between two spin zero particles. Does the same hold for the spin 1/2 particles?

Recall that the strategy to compute the potential is to take the non-relativistic limit of the scattering amplitude, and compare with the analogous result from quantum mechanics. Our new amplitude now also includes the spinor degrees of freedom  $u(p^\rightarrow)$  and  $v(p^\rightarrow)$ . In the non-relativistic limit,  $p \rightarrow (m, p^\rightarrow)$ , and

$$u(p^\rightarrow) = \sqrt{\frac{p \cdot \sigma \xi}{m}} \rightarrow \sqrt{\frac{m}{m}} \xi \\ v(p^\rightarrow) = \frac{\sqrt{\frac{p \cdot \sigma^- \xi}{m}}}{-\sqrt{p \cdot \sigma^- \xi}} \rightarrow \sqrt{\frac{m}{m}} \xi \quad (5.49)$$

In this limit, the spinor contractions in the amplitude for  $\psi\psi \rightarrow \psi\psi$  scattering (5.45) become  $u^{-s'} \cdot u^s = 2m\delta^{ss'}$  and the amplitude is

$$= -i(-i\lambda)^2 (2m) \frac{\delta^{s's} \delta^{r'r}}{(p - p')^2 + \mu^2} - \frac{\delta^{s'r} \delta^{r's}}{(p - q')^2 + \mu^2} \quad (5.50)$$

The  $\delta$  symbols tell us that spin is conserved in the non-relativistic limit, while the momentum dependence is the same as in the bosonic case, telling us that once again the particles feel an attractive Yukawa potential,

$$U(r) = -\frac{\lambda^2 e^{-\mu r}}{4\pi r} \quad (5.51)$$

Repeating the calculation for  $\bar{\psi}\psi \rightarrow \bar{\psi}\psi$ , there are two minus signs which cancel each other. The first is the extra overall minus sign in the scattering amplitude (5.46),

due to the fermionic nature of the particles. The second minus sign comes from the non-relativistic limit of the spinor contraction for anti-particles in (5.46), which is  $v^{-s} \cdot v^s = -2m\delta^{ss}$ . These two signs cancel, giving us once again an attractive Yukawa potential (5.51).

### 5.7.3 Pseudo-Scalar Coupling

Rather than the standard Yukawa coupling, we could instead consider

$$L_{Yuk} = -\lambda \varphi \bar{\psi} \gamma^5 \psi \quad (5.52)$$

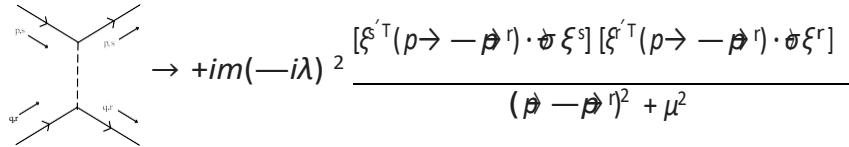
This still preserves parity if  $\varphi$  is a pseudoscalar, i.e.

$$P : \varphi(\mathbf{x}, t) \rightarrow -\varphi(-\mathbf{x}, t) \quad (5.53)$$

We can compute in this theory very simply: the Feynman rule for the interaction vertex is now changed to a factor of  $-i\lambda\gamma^5$ . For example, the Feynman diagrams for  $\psi\psi \rightarrow \psi\psi$  scattering are again given by Figure 25, with the amplitude now

$$A = (-i\lambda)^2 \frac{[\bar{u}^{s'}(p^r)\gamma^5 u^s(p)] [\bar{u}^{r'}(-q^s)\gamma^5 u^r(q)]}{(p - p^r)^2 - \mu^2} - \frac{[\bar{u}^{s'}(p^r)\gamma^5 u^s(p)] [\bar{u}^{r'}(-q^s)\gamma^5 u^r(q)]}{(p - q^r)^2 - \mu^2}$$

We could again try to take the non-relativistic limit for this amplitude. But this time, things work a little differently. Using the expressions for the spinors (5.49), we have  $\bar{u}^s \gamma^5 u^s \rightarrow 0$  in the non-relativistic limit. To find the non-relativistic amplitude, we must go to next to leading order. One can easily check that  $u^{-s}(p \rightarrow r)\gamma^5 u^s(p) \rightarrow m \xi^{sT}(p \rightarrow -p^r) \cdot \sigma \xi^s$ . So, in the non-relativistic limit, the leading order amplitude arising from pseudoscalar exchange is given by a spin-spin coupling,



$$\rightarrow +im(-i\lambda)^2 \frac{[\xi^{sT}(p \rightarrow -p^r) \cdot \sigma \xi^s] [\xi^{rT}(p \rightarrow -p^r) \cdot \sigma \xi^r]}{(p - p^r)^2 + \mu^2} \quad (5.54)$$

## 6. Quantum Electrodynamics

In this section we finally get to quantum electrodynamics (QED), the theory of light interacting with charged matter. Our path to quantization will be as before: we start with the free theory of the electromagnetic field and see how the quantum theory gives rise to a photon with two polarization states. We then describe how to couple the photon to fermions and to bosons.

### 6.1 Maxwell's Equations

The Lagrangian for Maxwell's equations in the absence of any sources is simply

$$L = \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (6.1)$$

where the field strength is defined by

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (6.2)$$

The equations of motion which follow from this Lagrangian are

$$\partial_\mu \frac{\partial L}{\partial (\partial_\mu A_\nu)} = -\partial_\mu F^{\mu\nu} = 0 \quad (6.3)$$

Meanwhile, from the definition of  $F_{\mu\nu}$ , the field strength also satisfies the Bianchi identity

$$\partial_\lambda F_{\mu\nu} + \partial_\mu F_{\nu\lambda} + \partial_\nu F_{\lambda\mu} = 0 \quad (6.4)$$

To make contact with the form of Maxwell's equations you learn about in high school, we need some 3-vector notation. If we define  $A^\mu = (\varphi, \vec{A})$ , then the electric field  $\vec{E}$  and magnetic field  $\vec{B}$  are defined by

$$\vec{E} = -\nabla\varphi - \frac{\partial \vec{A}}{\partial t} \quad \text{and} \quad \vec{B} = \nabla \times \vec{A} \quad (6.5)$$

which, in terms of  $F_{\mu\nu}$ , becomes

$$F_{\mu\nu} = \begin{matrix} & \begin{matrix} 0 & E_x & E_y & E_z \end{matrix} \\ \begin{matrix} 0 & -E_x & -B_z & B_y \\ -E_y & 0 & -B_x & B_z \\ -E_z & B_y & 0 & -B_x \end{matrix} & \end{matrix} \quad ! \quad (6.6)$$

The Bianchi identity (6.4) then gives two of Maxwell's equations,

$$\nabla \cdot \vec{B} = 0 \quad \text{and} \quad \frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} \quad (6.7)$$

These remain true even in the presence of electric sources. Meanwhile, the equations of motion give the remaining two Maxwell equations,

$$\nabla \cdot \vec{E} = 0 \quad \text{and} \quad \frac{\partial \vec{E}}{\partial t} = \nabla \times \vec{B} \quad (6.8)$$

As we will see shortly, in the presence of charged matter these equations pick up extra terms on the right-hand side.

### 6.1.1 Gauge Symmetry

The massless vector field  $A_\mu$  has 4 components, which would naively seem to tell us that the gauge field has 4 degrees of freedom. Yet we know that the photon has only two degrees of freedom which we call its polarization states. How are we going to resolve this discrepancy? There are two related comments which will ensure that quantizing the gauge field  $A_\mu$  gives rise to 2 degrees of freedom, rather than 4.

- The field  $A_0$  has no kinetic term  $\dot{A}_0$  in the Lagrangian: it is not dynamical. This means that if we are given some initial data  $A_i$  and  $\dot{A}_i$  at a time  $t_0$ , then the field  $A_0$  is fully determined by the equation of motion  $\nabla \cdot \vec{E} = 0$  which, expanding out, reads

$$\nabla^2 A_0 + \nabla \cdot \frac{\partial \vec{A}}{\partial t} = 0 \quad (6.9)$$

This has the solution

$$A_0(\vec{x}) = \int d^3x^r \frac{(\nabla \cdot \partial \vec{A}/\partial t)(\vec{x}^r)}{4\pi|\vec{x} - \vec{x}^r|} \quad (6.10)$$

So  $A_0$  is not independent: we don't get to specify  $A_0$  on the initial time slice. It looks like we have only 3 degrees of freedom in  $A_\mu$  rather than 4. But this is still one too many.

- The Lagrangian (6.3) has a *very* large symmetry group, acting on the vector potential as

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \lambda(x) \quad (6.11)$$

for any function  $\lambda(x)$ . We'll ask only that  $\lambda(x)$  dies off suitably quickly at spatial  $\vec{x} \rightarrow \infty$ . We call this a *gauge symmetry*. The field strength is invariant under the gauge symmetry:

$$F_{\mu\nu} \rightarrow \partial_\mu(A_\nu + \partial_\nu \lambda) - \partial_\nu(A_\mu + \partial_\mu \lambda) = F_{\mu\nu} \quad (6.12)$$

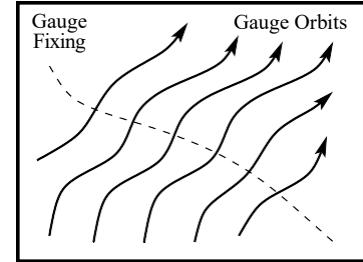
So what are we to make of this? We have a theory with an infinite number of symmetries, one for each function  $\lambda(x)$ . Previously we only encountered symmetries which act the same at all points in spacetime, for example  $\psi \rightarrow e^{i\alpha}\psi$  for a complex scalar field. Noether's theorem told us that these symmetries give rise to conservation laws. Do we now have an infinite number of conservation laws?

The answer is no! Gauge symmetries have a very different interpretation than the global symmetries that we make use of in Noether's theorem. While the latter take a physical state to another physical state with the same properties, the gauge symmetry is to be viewed as a redundancy in our description. That is, two states related by a gauge symmetry are to be identified: they are the same physical state. (There is a small caveat to this statement which is explained in Section 6.3.1). One way to see that this interpretation is necessary is to notice that Maxwell's equations are not sufficient to specify the evolution of  $A_\mu$ . The equations read,

$$[\eta_{\mu\nu}(\partial^\rho\partial_\rho) - \partial_\mu\partial_\nu] A^\nu = 0 \quad (6.13)$$

But the operator  $[\eta_{\mu\nu}(\partial^\rho\partial_\rho) - \partial_\mu\partial_\nu]$  is not invertible: it annihilates any function of the form  $\partial_\mu\lambda$ . This means that given any initial data, we have no way to uniquely determine  $A_\mu$  at a later time since we can't distinguish between  $A_\mu$  and  $A_\mu + \partial_\mu\lambda$ . This would be problematic if we thought that  $A_\mu$  is a physical object. However, if we're happy to identify  $A_\mu$  and  $A_\mu + \partial_\mu\lambda$  as corresponding to the same physical state, then our problems disappear.

Since gauge invariance is a redundancy of the system, we might try to formulate the theory purely in terms of the local, physical, gauge invariant objects  $\mathbf{E}$  and  $\mathbf{B}$ . This is fine for the free classical theory: Maxwell's equations were, after all, first written in terms of  $\mathbf{E}$  and  $\mathbf{B}$ . But it is not possible to describe certain quantum phenomena, such as the Aharonov-Bohm effect, without using the gauge potential  $A_\mu$ . We will see shortly that we also require the gauge potential to describe classically charged fields. To describe Nature, it appears that we have to introduce quantities  $A_\mu$  that we can never measure.



**Figure 29:**

The picture that emerges for the theory of electromagnetism is of an enlarged phase space, foliated by gauge orbits as shown in the figure. All states that lie along a given

line can be reached by a gauge transformation and are identified. To make progress, we pick a representative from each gauge orbit. It doesn't matter which representative we pick — after all, they're all physically equivalent. But we should make sure that we pick a "good" gauge, in which we cut the orbits.

Different representative configurations of a physical state are called different *gauges*. There are many possibilities, some of which will be more useful in different situations. Picking a gauge is rather like picking coordinates that are adapted to a particular problem. Moreover, different gauges often reveal slightly different aspects of a problem. Here we'll look at two different gauges:

- **Lorentz Gauge:**  $\partial_\mu A^\mu = 0$

To see that we can always pick a representative configuration satisfying  $\partial_\mu A^\mu = 0$ , suppose that we're handed a gauge field  $A^\mu$  satisfying  $\partial_\mu (A^\mu)^\mu = f(x)$ . Then we choose  $A_\mu = A^\mu + \partial_\mu \lambda$ , where

$$\partial_\mu \partial^\mu \lambda = -f \quad (6.14)$$

This equation always has a solution. In fact this condition doesn't pick a unique representative from the gauge orbit. We're always free to make further gauge transformations with  $\partial_\mu \partial^\mu \lambda = 0$ , which also has non-trivial solutions. As the name suggests, the Lorentz gauge<sup>3</sup> has the advantage that it is Lorentz invariant.

- **Coulomb Gauge:**  $\nabla \cdot \vec{A} = 0$

We can make use of the residual gauge transformations in Lorentz gauge to pick  $\nabla \cdot \vec{A} = 0$ . (The argument is the same as before). Since  $A_0$  is fixed by (6.10), we have as a consequence

$$A_0 = 0 \quad (6.15)$$

(This equation will no longer hold in Coulomb gauge in the presence of charged matter). Coulomb gauge breaks Lorentz invariance, so may not be ideal for some purposes. However, it is very useful to exhibit the physical degrees of freedom: the 3 components of  $\vec{A}$  satisfy a single constraint:  $\nabla \cdot \vec{A} = 0$ , leaving behind just 2 degrees of freedom. These will be identified with the two polarization states of the photon. Coulomb gauge is sometimes called radiation gauge.

---

<sup>3</sup>Named after Lorenz who had the misfortune to be one letter away from greatness.

## 6.2 The Quantization of the Electromagnetic Field

In the following we shall quantize free Maxwell theory twice: once in Coulomb gauge, and again in Lorentz gauge. We'll ultimately get the same answers and, along the way, see that each method comes with its own subtleties.

The first of these subtleties is common to both methods and comes when computing the momentum  $\pi^\mu$  conjugate to  $A_\mu$ ,

$$\begin{aligned}\pi^0 &= \frac{\partial L}{\partial \dot{A}_0} = 0 \\ \pi^i &= \frac{\partial L}{\partial \dot{A}_i} = -F^{0i} \equiv E^i\end{aligned}\tag{6.16}$$

so the momentum  $\pi^0$  conjugate to  $A_0$  vanishes. This is the mathematical consequence of the statement we made above:  $A_0$  is not a dynamical field. Meanwhile, the momentum conjugate to  $A_i$  is our old friend, the electric field. We can compute the Hamiltonian,

$$\begin{aligned}H &= \int d^3x \pi^i \dot{A}_i - L \\ &= \int d^3x \frac{1}{2} \vec{E} \cdot \vec{E} + \frac{1}{2} \vec{B} \cdot \vec{B} - A_0(\nabla \cdot \vec{E})\end{aligned}\tag{6.17}$$

So  $A_0$  acts as a Lagrange multiplier which imposes Gauss' law

$$\nabla \cdot \vec{E} = 0\tag{6.18}$$

which is now a constraint on the system in which  $\vec{A}$  are the physical degrees of freedom. Let's now see how to treat this system using different gauge fixing conditions.

### 6.2.1 Coulomb Gauge

In Coulomb gauge, the equation of motion for  $\vec{A}$  is

$$\partial_\mu \partial^\mu \vec{A} = 0\tag{6.19}$$

which we can solve in the usual way,

$$\vec{A} = \int \frac{d^3p}{(2\pi)^3} \xi^\rightarrow(p) e^{ipx}\tag{6.20}$$

with  $p_0^2 = |\vec{p}|^2$ . The constraint  $\nabla \cdot \vec{A} = 0$  tells us that  $\xi^\rightarrow$  must satisfy

$$\xi^\rightarrow \cdot \vec{p} = 0\tag{6.21}$$

which means that  $\xi \rightarrow$  is perpendicular to the direction of motion  $p \rightarrow$ . We can pick  $\xi \rightarrow(p \rightarrow)$  to be a linear combination of two orthonormal vectors  $\rightarrow\epsilon_r$ ,  $r = 1, 2$ , each of which satisfies  $\epsilon_r(\hat{p}) \cdot \hat{p} = 0$  and

$$\epsilon_r(\hat{p}) \cdot \epsilon_s(\hat{p}) = \delta_{rs} \quad r, s = 1, 2 \quad (6.22)$$

These two vectors correspond to the two polarization states of the photon. It's worth pointing out that you can't consistently pick a continuous basis of polarization vectors for every value of  $\hat{p}$  because you can't comb the hair on a sphere. But this topological fact doesn't cause any complications in computing QED scattering processes.

To quantize we turn the Poisson brackets into commutators. Naively we would write

$$\begin{aligned} [A_i(\hat{x}), A_j(\hat{y})] &= [E^i(\hat{x}), E^j(\hat{y})] = 0 \\ [A_i(\hat{x}), E^j(\hat{y})] &= i\delta_i^{(3)}(\hat{x} \rightarrow y) \end{aligned} \quad (6.23)$$

But this can't quite be right, because it's not consistent with the constraints. We still want to have  $\nabla \cdot \hat{A} = \nabla \cdot \hat{E} = 0$ , now imposed on the operators. But from the commutator relations above, we see

$$[\nabla \cdot \hat{A}(\hat{x}), \nabla \cdot \hat{E}(\hat{y})] = \nabla^2 \delta^{(3)}(\hat{x} \rightarrow y) \neq 0 \quad (6.24)$$

What's going on? In imposing the commutator relations (6.23) we haven't correctly taken into account the constraints. In fact, this is a problem already in the classical theory, where the Poisson bracket structure is already altered<sup>4</sup>. The correct Poisson bracket structure leads to an alteration of the last commutation relation,

$$[A_i(\hat{x}), E_j(\hat{y})] = i \delta_{ij} \frac{\partial_i \partial_j}{\nabla^2} \delta^{(3)}(\hat{x} \rightarrow y) \quad (6.25)$$

To see that this is now consistent with the constraints, we can rewrite the right-hand side of the commutator in momentum space,

$$[A_i(\hat{x}), E_j(\hat{y})] = i \int \frac{d^3 p}{(2\pi)^3} \delta_{ij} - \frac{p_i p_j}{|\hat{p}|^2} e^{i\hat{p} \cdot (\hat{x} \rightarrow y)} \quad (6.26)$$

which is now consistent with the constraints, for example

$$[\partial_i A_i(\hat{x}), E_j(\hat{y})] = i \int \frac{d^3 p}{(2\pi)^3} \delta_{ij} - \frac{p_i p_j}{|\hat{p}|^2} i p_i e^{i\hat{p} \cdot (\hat{x} \rightarrow y)} = 0 \quad (6.27)$$

---

<sup>4</sup>For a nice discussion of the classical and quantum dynamics of constrained systems, see the small book by Paul Dirac, "Lectures on Quantum Mechanics"

We now write  $\vec{A}$  in the usual mode expansion,

$$\begin{aligned}\vec{A}(\mathbf{x}) &= \frac{\int d^3p}{(2\pi)^3} \frac{1}{2|\vec{p}|} \sum_{r=1}^2 h_r \frac{\epsilon_r(\vec{p})}{\vec{p}} a_{\vec{p}}^r e^{i\vec{p} \cdot \vec{x}} + a_{\vec{p}}^r e^{-i\vec{p} \cdot \vec{x}} \\ \vec{E}(\mathbf{x}) &= \frac{(-i)}{(2\pi)^3} \frac{1}{2} \sum_{r=1}^2 \epsilon_r(\vec{p}) \frac{a_{\vec{p}}^r e^{i\vec{p} \cdot \vec{x}} - a_{\vec{p}}^r e^{-i\vec{p} \cdot \vec{x}}}{\vec{p}}\end{aligned}\quad (6.28)$$

where, as before, the polarization vectors satisfy

$$\rightarrow\epsilon_r(\vec{p}) \cdot \vec{p} = 0 \quad \text{and} \quad \epsilon_r(\vec{p}) \cdot \epsilon_s(\vec{p}) = \delta_{rs} \quad (6.29)$$

It is not hard to show that the commutation relations (6.25) are equivalent to the usual commutation relations for the creation and annihilation operators,

$$\begin{aligned}[a_r^r, a_s^s] &= [a_r^r \dagger, a_s^s \dagger] = 0 \\ [a_{\vec{p}}^r, a_{\vec{q}}^s \dagger] &= (2\pi)^3 \delta^{rs} \delta^{(3)}(\vec{p} \rightarrow \vec{q})\end{aligned}\quad (6.30)$$

where, in deriving this, we need the completeness relation for the polarization vectors,

$$\sum_{r=1}^2 \epsilon_r(\vec{p}) \epsilon_r^*(\vec{p}) = \delta^{ij} \frac{\vec{p} \cdot \vec{p}}{|\vec{p}|^2} \quad (6.31)$$

You can easily check that this equation is true by acting on both sides with a basis of vectors ( $\epsilon_1(\vec{p}), \epsilon_2(\vec{p}), \vec{p}$ ).

We derive the Hamiltonian by substituting (6.28) into (6.17). The last term vanishes in Coulomb gauge. After normal ordering, and playing around with  $\rightarrow\epsilon_r$  polarization vectors, we get the simple expression

$$H = \frac{\int d^3p}{(2\pi)^3} |\vec{p}| \sum_{r=1}^2 a_{\vec{p}}^r \dagger a_{\vec{p}}^r \quad (6.32)$$

The Coulomb gauge has the advantage that the physical degrees of freedom are manifest. However, we've lost all semblance of Lorentz invariance. One place where this manifests itself is in the propagator for the fields  $A_i(x)$  (in the Heisenberg picture). In Coulomb gauge the propagator reads

$$D_{ij}(x - y) \equiv \langle 0 | T A_i(x) A_j(y) | 0 \rangle = \frac{\int d^4p}{(2\pi)^4} \frac{i}{p^2 + i\epsilon} \delta_{ij} \frac{p_i p_j}{|p|^2} e^{-ip \cdot (x-y)} \quad (6.33)$$

The  $tr$  superscript on the propagator refers to the "transverse" part of the photon. When we turn to the interacting theory, we will have to fight to massage this propagator into something a little nicer.

### 6.2.2 Lorentz Gauge

We could try to work in a Lorentz invariant fashion by imposing the Lorentz gauge condition  $\partial_\mu A^\mu = 0$ . The equations of motion that follow from the action are then

$$\partial_\mu \partial^\mu A^\nu = 0 \quad (6.34)$$

Our approach to implementing Lorentz gauge will be a little different from the method we used in Coulomb gauge. We choose to change the theory so that (6.34) arises directly through the equations of motion. We can achieve this by taking the Lagrangian

$$L = \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2 \quad (6.35)$$

The equations of motion coming from this action are

$$\partial_\mu F^{\mu\nu} + \partial^\nu (\partial_\mu A^\mu) = \partial_\mu \partial^\mu A^\nu = 0 \quad (6.36)$$

(In fact, we could be a little more general than this, and consider the Lagrangian

$$L = \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2\alpha} (\partial_\mu A^\mu)^2 \quad (6.37)$$

with arbitrary  $\alpha$  and reach similar conclusions. The quantization of the theory is independent of  $\alpha$  and, rather confusingly, different choices of  $\alpha$  are sometimes also referred to as different “gauges”. We will use  $\alpha = 1$ , which is called “Feynman gauge”. The other common choice,  $\alpha = 0$ , is called “Landau gauge”.)

Our plan will be to quantize the theory (6.36), and only later impose the constraint  $\partial_\mu A^\mu = 0$  in a suitable manner on the Hilbert space of the theory. As we’ll see, we will also have to deal with the residual gauge symmetry of this theory which will prove a little tricky. At first, we can proceed very easily, because both  $\pi^0$  and  $\pi^i$  are dynamical:

$$\begin{aligned} \pi^0 &= \frac{\partial L}{\partial \dot{A}_0} = -\partial_\mu A^\mu \\ \pi^i &= \frac{\partial L}{\partial \dot{A}_i} = \partial^i A^0 - A^i \end{aligned} \quad (6.38)$$

Turning these classical fields into operators, we can simply impose the usual commutation relations,

$$\begin{aligned} [A_\mu(\mathbf{x}), A_\nu(\mathbf{y})] &= [\pi^\mu(\mathbf{x}), \pi^\nu(\mathbf{y})] = 0 \\ [A_\mu(\mathbf{x}), \pi_\nu(\mathbf{y})] &= i\eta_{\mu\nu} \delta^{(3)}(\mathbf{x} \rightarrow \mathbf{y}) \end{aligned} \quad (6.39)$$

and we can make the usual expansion in terms of creation and annihilation operators and 4 polarization vectors  $(\epsilon_\mu)^\lambda$ , with  $\lambda = 0, 1, 2, 3$ .

$$A_\mu(\mathbf{x}) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2|\mathbf{p}|}} \sum_{\lambda=0}^3 \epsilon_\mu(\mathbf{p}) \frac{\mathbf{h}}{a_p^\lambda e^{ip \cdot x} + a_p^\lambda e^{-ip \cdot x}} \quad (6.40)$$

$$\pi^\mu(\mathbf{x}) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2|\mathbf{p}|}} \frac{(+i)}{2} \sum_{\lambda=0}^3 (\epsilon^\mu)^\lambda(\mathbf{p}) \frac{\mathbf{h}}{a^\lambda e^{ip \cdot x} - a^{\lambda\dagger} e^{-ip \cdot x}} \quad (6.40)$$

Note that the momentum  $\pi^\mu$  comes with a factor of  $(+i)$ , rather than the familiar  $(-i)$  that we've seen so far. This can be traced to the fact that the momentum (6.38) for the classical fields takes the form  $\pi^\mu = -A^\mu + \dots$ . In the Heisenberg picture, it becomes clear that this descends to  $(+i)$  in the definition of momentum.

There are now four polarization 4-vectors  $\epsilon^\lambda(\mathbf{p})$ , instead of the two polarization 3-vectors that we met in the Coulomb gauge. Of these four 4-vectors, we pick  $\epsilon^0$  to be timelike, while  $\epsilon^{1,2,3}$  are spacelike. We pick the normalization

$$\epsilon^\lambda \cdot \epsilon^{\lambda'} = \eta^{\lambda\lambda'} \quad (6.41)$$

which also means that

$$(\epsilon_\mu)^\lambda (\epsilon_\nu)^\lambda' \eta_{\lambda\lambda'} = \eta_{\mu\nu} \quad (6.42)$$

The polarization vectors depend on the photon 4-momentum  $p = (|\mathbf{p}|, \mathbf{p})$ . Of the two spacelike polarizations, we will choose  $\epsilon^1$  and  $\epsilon^2$  to lie transverse to the momentum:

$$\epsilon^1 \cdot p = \epsilon^2 \cdot p = 0 \quad (6.43)$$

The third vector  $\epsilon^3$  is the longitudinal polarization. For example, if the momentum lies along the  $x^3$  direction, so  $p \sim (1, 0, 0, 1)$ , then

$$\epsilon^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \epsilon^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \epsilon^3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (6.44)$$

For other 4-momenta, the polarization vectors are the appropriate Lorentz transformations of these vectors, since (6.43) are Lorentz invariant.

We do our usual trick, and translate the field commutation relations (6.39) into those for creation and annihilation operators. We find  $[a_{\mathbf{p}}^\lambda, a_{\mathbf{q}}^{\lambda'}] = [a_{\mathbf{p}}^{\lambda\dagger}, a_{\mathbf{q}}^{\lambda'\dagger}] = 0$  and

$$[a_{\mathbf{p}}^\lambda, a_{\mathbf{q}}^{\lambda'\dagger}] = -\eta^{\lambda\lambda'} (2\pi)^3 \delta^{(3)}(\mathbf{p} - \mathbf{q}) \quad (6.45)$$

The minus signs here are odd to say the least! For spacelike  $\lambda = 1, 2, 3$ , everything looks fine,

$$[a_{p\lambda}^\lambda a_{q\lambda'}^{\lambda'\dagger}] = \delta^{\lambda\lambda'} (2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \quad \lambda, \lambda' = 1, 2, 3 \quad (6.46)$$

But for the timelike annihilation and creation operators, we have

$$[a_{p\lambda}^0 a_{q\lambda}^{0\dagger}] = -(2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \quad (6.47)$$

This is very odd! To see just how strange this is, we take the Lorentz invariant vacuum  $|0\rangle$  defined by

$$a_{p\lambda}^\lambda |0\rangle = 0 \quad (6.48)$$

Then we can create one-particle states in the usual way,

$$|\vec{p}, \lambda\rangle = a_{p\lambda}^{\lambda\dagger} |0\rangle \quad (6.49)$$

For spacelike polarization states,  $\lambda = 1, 2, 3$ , all seems well. But for the timelike polarization  $\lambda = 0$ , the state  $|p\rightarrow, 0\rangle$  has negative norm,

$$\langle \vec{p}, 0 | \rightarrow q, 0 \rangle = \langle 0 | a_{p\rightarrow} a_{q\rightarrow} | 0 \rangle = -(2\pi)^3 \delta^{(3)}(\vec{p} - \vec{q}) \quad (6.50)$$

Wtf? That's very very strange. A Hilbert space with negative norm means negative probabilities which makes no sense at all. We can trace this negative norm back to the wrong sign of the kinetic term for  $A_0$  in our original Lagrangian:  $L = +\frac{1}{2} \vec{A} \cdot \vec{\partial} \vec{A} - \frac{1}{2} \dot{A}_0^2 + \dots$

At this point we should remember our constraint equation,  $\partial_\mu A^\mu = 0$ , which, until now, we've not imposed on our theory. This is going to come to our rescue. We will see that it will remove the timelike, negative norm states, and cut the physical polarizations down to two. We work in the Heisenberg picture, so that

$$\partial_\mu A^\mu = 0 \quad (6.51)$$

makes sense as an operator equation. Then we could try implementing the constraint in the quantum theory in a number of different ways. Let's look at a number of increasingly weak ways to do this

- We could ask that  $\partial_\mu A^\mu = 0$  is imposed as an equation on operators. But this can't possibly work because the commutation relations (6.39) won't be obeyed for  $\pi^0 = -\partial_\mu A^\mu$ . We need some weaker condition.

- We could try to impose the condition on the Hilbert space instead of directly on the operators. After all, that's where the trouble lies! We could imagine that there's some way to split the Hilbert space up into good states  $|\Psi\rangle$  and bad states that somehow decouple from the system. With luck, our bad states will include the weird negative norm states that we're so disgusted by. But how can we define the good states? One idea is to impose

$$\partial_\mu A^\mu |\Psi\rangle = 0 \quad (6.52)$$

on all good, physical states  $|\Psi\rangle$ . But this can't work either! Again, the condition is too strong. For example, suppose we decompose  $A_\mu(x) = A_\mu^+(x) + A_\mu^-(x)$  with

$$A_\mu^+(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2|p|}} \sum_{\lambda=0} \lambda \epsilon_\mu a_p e^{-ip \cdot x}$$

$$A_\mu^-(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2|p|}} \sum_{\lambda=0} \lambda \epsilon_\mu a_p^\dagger e^{+ip \cdot x} \quad (6.53)$$

Then, on the vacuum  $A_\mu^+ |0\rangle = 0$  automatically, but  $\partial^\mu A_\mu^- |0\rangle \neq 0$ . So not even the vacuum is a physical state if we use (6.52) as our constraint

- Our final attempt will be the correct one. In order to keep the vacuum as a good physical state, we can ask that physical states  $|\Psi\rangle$  are defined by

$$\partial^\mu A_\mu^+ |\Psi\rangle = 0 \quad (6.54)$$

This ensures that

$$\langle \Psi | \partial_\mu A^\mu | \Psi \rangle = 0 \quad (6.55)$$

so that the operator  $\partial_\mu A^\mu$  has vanishing matrix elements between physical states. Equation (6.54) is known as the *Gupta-Bleuler* condition. The linearity of the constraint means that the physical states  $|\Psi\rangle$  span a physical Hilbert space  $H_{\text{phys}}$ .

So what does the physical Hilbert space  $H_{\text{phys}}$  look like? And, in particular, have we rid ourselves of those nasty negative norm states so that  $H_{\text{phys}}$  has a positive definite inner product defined on it? The answer is actually no, but almost!

Let's consider a basis of states for the Fock space. We can decompose any element of this basis as  $|\Psi\rangle = |\psi_T\rangle |\varphi\rangle$ , where  $|\psi_T\rangle$  contains only transverse photons, created by

$a_{p\rightarrow}^{1,2\dagger}$ , while  $|\varphi\rangle$  contains the timelike photons created by  $a_{p\rightarrow}^{0\dagger}$  and longitudinal photons created by  $a_{p\rightarrow}^{3\dagger}$ . The Gupta-Bleuler condition (6.54) requires

$$(a_{p\rightarrow}^3 - a_{p\rightarrow}^0) |\varphi\rangle = 0 \quad (6.56)$$

This means that the physical states must contain combinations of timelike and longitudinal photons. Whenever the state contains a timelike photon of momentum  $p\rightarrow$ , it must also contain a longitudinal photon with the same momentum. In general  $|\varphi\rangle$  will be a linear combination of states  $|\varphi_n\rangle$  containing  $n$  pairs of timelike and longitudinal photons, which we can write as

$$|\varphi\rangle = \sum_{n=0}^{\infty} C_n |\varphi_n\rangle \quad (6.57)$$

where  $|\varphi_0\rangle = |0\rangle$  is simply the vacuum. It's not hard to show that although the condition (6.56) does indeed decouple the negative norm states, all the remaining states involving timelike and longitudinal photons have zero norm

$$\langle \varphi_m | \varphi_n \rangle = \delta_{m0}\delta_{n0} \quad (6.58)$$

This means that the inner product on  $H_{\text{phys}}$  is positive semi-definite. Which is an improvement. But we still need to deal with all these zero norm states.

The way we cope with the zero norm states is to treat them as gauge equivalent to the vacuum. Two states that differ only in their timelike and longitudinal photon content,  $|\varphi_n\rangle$  with  $n \geq 1$  are said to be physically equivalent. We can think of the gauge symmetry of the classical theory as descending to the Hilbert space of the quantum theory. Of course, we can't just stipulate that two states are physically identical unless they give the same expectation value for all physical observables. We can check that this is true for the Hamiltonian, which can be easily computed to be

$$H = \int \frac{d^3 p}{(2\pi)^3} |\not{p}| \sum_{i=1}^{\infty} a_i^\dagger a_i^i - a_0^\dagger a_0^0 \quad (6.59)$$

But the condition (6.56) ensures that  $\langle \Psi | a_{p\rightarrow}^{3\dagger} a_{p\rightarrow}^3 | \Psi \rangle = \langle \Psi | a_{p\rightarrow}^{0\dagger} a_{p\rightarrow}^0 | \Psi \rangle$  so that the contributions from the timelike and longitudinal photons cancel amongst themselves in the Hamiltonian. This also renders the Hamiltonian positive definite, leaving us just with the contribution from the transverse photons as we would expect.

In general, one can show that the expectation values of all gauge invariant operators evaluated on physical states are independent of the coefficients  $C_n$  in (6.57).

## Propagators

Finally, it's a simple matter to compute the propagator in Lorentz gauge. It is given by

$$\langle 0 | T A_\mu(x) A_\nu(y) | 0 \rangle = \int \frac{d^4 p}{(2\pi)^4} \frac{-i\eta_{\mu\nu}}{p^2 + i\epsilon} e^{-ip \cdot (x-y)} \quad (6.60)$$

This is a lot nicer than the propagator we found in Coulomb gauge: in particular, it's Lorentz invariant. We could also return to the Lagrangian (6.37). Had we pushed through the calculation with arbitrary coefficient  $\alpha$ , we would find the propagator,

$$\langle 0 | T A_\mu(x) A_\nu(y) | 0 \rangle = \int \frac{d^4 p}{(2\pi)^4} \frac{-i}{p^2 + i\epsilon} \eta_{\mu\nu} + (\alpha - 1) \frac{p_\mu p_\nu}{p^2} e^{-ip \cdot (x-y)} \quad (6.61)$$

## 6.3 Coupling to Matter

Let's now build an interacting theory of light and matter. We want to write down a Lagrangian which couples  $A_\mu$  to some matter fields, either scalars or spinors. For example, we could write something like

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - j^\mu A_\mu \quad (6.62)$$

where  $j^\mu$  is some function of the matter fields. The equations of motion read

$$\partial_\mu F^{\mu\nu} = j^\nu \quad (6.63)$$

so, for consistency, we require

$$\partial_\mu j^\mu = 0 \quad (6.64)$$

In other words,  $j^\mu$  must be a conserved current. But we've got lots of those! Let's look at how we can couple two of them to electromagnetism.

### 6.3.1 Coupling to Fermions

The Dirac Lagrangian

$$L = \bar{\psi} (i \not{D} - m) \psi \quad (6.65)$$

has an internal symmetry  $\psi \rightarrow e^{-i\alpha} \psi$  and  $\bar{\psi} \rightarrow e^{+i\alpha} \bar{\psi}$ , with  $\alpha \in \mathbb{R}$ . This gives rise to the conserved current  $j_\nu^\mu = \bar{\psi} \gamma^\mu \psi$ . So we could look at the theory of electromagnetism coupled to fermions, with the Lagrangian,

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi} (i \not{D} - m) \psi - e \bar{\psi} \gamma^\mu A_\mu \psi \quad (6.66)$$

where we've introduced a coupling constant  $e$ . For the free Maxwell theory, we have seen that the existence of a gauge symmetry was crucial in order to cut down the physical degrees of freedom to the requisite 2. Does our interacting theory above still have a gauge symmetry? The answer is yes. To see this, let's rewrite the Lagrangian as

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\not{D} - m)\psi \quad (6.67)$$

where  $D_\mu\psi = \partial_\mu\psi + ieA_\mu\psi$  is called the *covariant derivative*. This Lagrangian is invariant under gauge transformations which act as

$$A_\mu \rightarrow A_\mu + \partial_\mu\lambda \quad \text{and} \quad \psi \rightarrow e^{-ie\lambda}\psi \quad (6.68)$$

for an arbitrary function  $\lambda(x)$ . The tricky term is the derivative acting on  $\psi$ , since this will also hit the  $e^{-ie\lambda}$  piece after the transformation. To see that all is well, let's look at how the covariant derivative transforms. We have

$$\begin{aligned} D_\mu\psi &= \partial_\mu\psi + ieA_\mu\psi \\ &\rightarrow \partial_\mu(e^{-ie\lambda}\psi) + ie(A_\mu + \partial_\mu\lambda)(e^{-ie\lambda}\psi) \\ &= e^{-ie\lambda}D_\mu\psi \end{aligned} \quad (6.69)$$

so the covariant derivative has the nice property that it merely picks up a phase under the gauge transformation, with the derivative of  $e^{-ie\lambda}$  cancelling the transformation of the gauge field. This ensures that the whole Lagrangian is invariant, since  $\bar{\psi} \rightarrow e^{+ie\lambda(x)}\bar{\psi}$ .

## Electric Charge

The coupling  $e$  has the interpretation of the electric charge of the  $\psi$  particle. This follows from the equations of motion of classical electromagnetism  $\partial_\mu F^{\mu\nu} = j^\nu$ : we know that the  $j^0$  component is the charge density. We therefore have the total charge  $Q$  given by

$$Q = e \int d^3x \bar{\psi}(\not{x})\gamma^0\psi(\not{x}) \quad (6.70)$$

After treating this as a quantum equation, we have

$$Q = e \sum_s \frac{(b^s\dagger b^s - c^s\dagger c^s)}{(2\pi)^3} \quad \begin{matrix} p \rightarrow p \rightarrow & p \rightarrow p \rightarrow \end{matrix} \quad (6.71)$$

which is the number of particles, minus the number of antiparticles. Note that the particle and the anti-particle are required by the formalism to have opposite electric

charge. For QED, the theory of light interacting with electrons, the electric charge is usually written in terms of the dimensionless ratio  $\alpha$ , known as the fine structure constant

$$\alpha = \frac{e^2}{4\pi k c} \approx \frac{1}{137} \quad (6.72)$$

Setting  $k = c = 1$ , we have  $e = \sqrt{\frac{1}{4\pi\alpha}} \approx 0.3$ .

There's a small subtlety here that's worth elaborating on. I stressed that there's a radical difference between the interpretation of a global symmetry and a gauge symmetry. The former takes you from one physical state to another with the same properties and results in a conserved current through Noether's theorem. The latter is a redundancy in our description of the system. Yet in electromagnetism, the gauge symmetry  $\psi \rightarrow e^{+ie\lambda(x)}\psi$  seems to lead to a conservation law, namely the conservation of electric charge. This is because among the infinite number of gauge symmetries parameterized by a function  $\lambda(x)$ , there is also a single global symmetry: that with  $\lambda(x) = \text{constant}$ . This is a true symmetry of the system, meaning that it takes us to another physical state. More generally, the subset of global symmetries from among the gauge symmetries are those for which  $\lambda(x) \rightarrow \alpha = \text{constant}$  as  $x \rightarrow \infty$ . These take us from one physical state to another.

Finally, let's check that the  $4 \times 4$  matrix  $C$  that we introduced in Section 4.5 really deserves the name "charge conjugation matrix". If we take the complex conjugation of the Dirac equation, we have

$$(i\gamma^\mu \partial_\mu - e\gamma^\mu A_\mu - m)\psi = 0 \quad \Rightarrow \quad (-i(\gamma^\mu)^\Lambda \partial_\mu - e(\gamma^\mu)^\Lambda A_\mu - m)\psi^\Lambda = 0$$

Now using the defining equation  $C^\dagger \gamma^\mu C = -(\gamma^\mu)^\Lambda$ , and the definition  $\psi^{(c)} = C\psi^\Lambda$ , we see that the charge conjugate spinor  $\psi^{(c)}$  satisfies

$$(i\gamma^\mu \partial_\mu + e\gamma^\mu A_\mu - m)\psi^{(c)} = 0 \quad (6.73)$$

So we see that the charge conjugate spinor  $\psi^{(c)}$  satisfies the Dirac equation, but with charge  $-e$  instead of  $+e$ .

### 6.3.2 Coupling to Scalars

For a real scalar field, we have no suitable conserved current. This means that we can't couple a real scalar field to a gauge field.

Let's now consider a complex scalar field  $\phi$ . (For this section, I'll depart from our previous notation and call the scalar field  $\phi$  to avoid confusing it with the spinor). We have a symmetry  $\phi \rightarrow e^{-i\alpha}\phi$ . We could try to couple the associated current to the gauge field,

$$L_{int} = -i((\partial_\mu \phi^\lambda) \phi - \phi^\lambda \partial_\mu \phi) A^\mu \quad (6.74)$$

But this doesn't work because

- The theory is no longer gauge invariant
- The current  $j^\mu$  that we coupled to  $A_\mu$  depends on  $\partial_\mu \phi$ . This means that if we try to compute the current associated to the symmetry, it will now pick up a contribution from the  $j^\mu A_\mu$  term. So the whole procedure wasn't consistent.

We solve both of these problems simultaneously by remembering the covariant derivative. In this scalar theory, the combination

$$D_\mu \phi = \partial_\mu \phi + ieA_\mu \phi \quad (6.75)$$

again transforms as  $D_\mu \phi \rightarrow e^{-ie\lambda} D_\mu \phi$  under a gauge transformation  $A_\mu \rightarrow A_\mu + \partial_\mu \lambda$  and  $\phi \rightarrow e^{-ie\lambda} \phi$ . This means that we can construct a gauge invariant action for a charged scalar field coupled to a photon simply by promoting all derivatives to covariant derivatives

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + D_\mu \phi^\lambda D^\mu \phi - m^2 |\phi|^2 \quad (6.76)$$

In general, this trick works for any theory. If we have a  $U(1)$  symmetry that we wish to couple to a gauge field, we may do so by replacing all derivatives by suitable covariant derivatives. This procedure is known as *minimal coupling*.

## 6.4 QED

Let's now work out the Feynman rules for the full theory of quantum electrodynamics (QED) – the theory of electrons interacting with light. The Lagrangian is

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi} (i \not{D} - m) \psi \quad (6.77)$$

where  $D_\mu = \partial_\mu + ieA_\mu$ .

The route we take now depends on the gauge choice. If we worked in Lorentz gauge previously, then we can jump straight to Section 6.5 where the Feynman rules for QED are written down. If, however, we worked in Coulomb gauge, then we still have a bit of work in front of us in order to massage the photon propagator into something Lorentz invariant. We will now do that.

In Coulomb gauge  $\nabla \cdot \vec{A} = 0$ , the equation of motion arising from varying  $A_0$  is now

$$-\nabla^2 A_0 = e\psi^\dagger\psi \equiv ej^0 \quad (6.78)$$

which has the solution

$$A_0(\vec{x}, t) = e \int d^3x' \frac{j^0(\vec{x}', t)}{4\pi|\vec{x} - \vec{x}'|} \quad (6.79)$$

In Coulomb gauge we can rewrite the Maxwell part of the Lagrangian as

$$\begin{aligned} L_{\text{Maxwell}} &= \int d^3x \frac{1}{2}\vec{E}^2 - \frac{1}{2}\vec{B}^2 \\ &= \int d^3x \frac{1}{2}(\dot{\vec{A}} + \nabla A_0)^2 - \frac{1}{2}\vec{B}^2 \\ &= \int d^3x \frac{1}{2}\dot{\vec{A}}^2 + \frac{1}{2}(\nabla A_0)^2 - \frac{1}{2}\vec{B}^2 \end{aligned} \quad (6.80)$$

where the cross-term has vanished using  $\nabla \cdot \vec{A} = 0$ . After integrating the second term by parts and inserting the equation for  $A_0$ , we have

$$L_{\text{Maxwell}} = \int d^3x \left[ \frac{1}{2}\dot{\vec{A}}^2 - \frac{1}{2}\vec{B}^2 + \frac{e^2}{2} \int d^3x' \frac{j_0(\vec{x})j_0(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|} \right] \quad (6.81)$$

We find ourselves with a nonlocal term in the action. This is exactly the type of interaction that we boasted in Section 1.1.4 never arises in Nature! It appears here as an artifact of working in Coulomb gauge: it does not mean that the theory of QED is nonlocal. For example, it wouldn't appear if we worked in Lorentz gauge.

We now compute the Hamiltonian. Changing notation slightly from previous chapters, we have the conjugate momenta,

$$\begin{aligned} \vec{p} &= \frac{\partial L}{\partial \dot{\vec{A}}} = \vec{A} \\ \pi_\psi &= \frac{\partial L}{\partial \dot{\psi}} = i\psi^\dagger \end{aligned} \quad (6.82)$$

which gives us the Hamiltonian

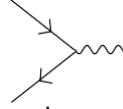
$$H = \int d^3x \left[ \frac{1}{2}\dot{\vec{A}}^2 + \frac{1}{2}\vec{B}^2 - i\psi^\dagger(-i\gamma^\mu \partial_\mu + m)\psi - ej \cdot \vec{A} + \frac{e^2}{2} \int d^3x' \frac{j_0(\vec{x})j_0(\vec{x}')}{4\pi|\vec{x} - \vec{x}'|} \right]$$

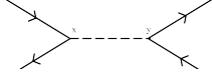
where  $\vec{j} = \bar{\psi} \gamma^\mu \psi$  and  $j^0 = \bar{\psi} \gamma^0 \psi$ .

### 6.4.1 Naive Feynman Rules

We want to determine the Feynman rules for this theory. For fermions, the rules are the same as those given in Section 5. The new pieces are:

- We denote the photon by a wavy line. Each end of the line comes with an  $i, j = 1, 2, 3$  index telling us the component of  $A^\mu$ . We calculated the transverse photon propagator in (6.33): it is  and contributes  $D^{\text{tr}}_{ij} = \frac{i}{p^2 + i\epsilon} \delta_{ij} - \frac{p_i p_j}{|p^\mu|^2}$

- The vertex  contributes  $-ie\gamma^i$ . The index on  $\gamma^i$  contracts with the index on the photon line.

- The non-local interaction which, in position space, is given by  contributes a factor of  $\frac{i(e\gamma^0)^2 \delta(x^0 - y^0)}{4\pi|\mathbf{x} - \mathbf{y}|}$

These Feynman rules are rather messy. This is the price we've paid for working in Coulomb gauge. We'll now show that we can massage these expressions into something much more simple and Lorentz invariant. Let's start with the offending instantaneous interaction. Since it comes from the  $A_0$  component of the gauge field, we could try to redefine the propagator to include a  $D_{00}$  piece which will capture this term. In fact, it fits quite nicely in this form: if we look in momentum space, we have

$$\frac{\delta(x^0 - y^0)}{4\pi|\mathbf{x} - \mathbf{y}|} = \int \frac{d^4 p}{(2\pi)^4} \frac{e^{ip \cdot (x-y)}}{|p^\mu|^2} \quad (6.83)$$

so we can combine the non-local interaction with the transverse photon propagator by defining a new photon propagator

$$D_{\mu\nu}(p) = \begin{cases} \frac{i}{p^2 + i\epsilon} \delta_{ij} - \frac{p_i p_j}{|p^\mu|^2} & \mu, \nu = 0 \\ 0 & \mu = i \neq 0, \nu = j \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (6.84)$$

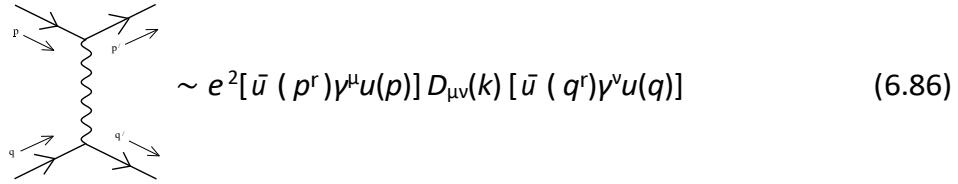
With this propagator, the wavy photon line now carries a  $\mu, \nu = 0, 1, 2, 3$  index, with the extra  $\mu = 0$  component taking care of the instantaneous interaction. We now need to change our vertex slightly: the  $-ie\gamma^i$  above gets replaced by  $-ie\gamma^\mu$  which correctly accounts for the  $(e\gamma^0)^2$  piece in the instantaneous interaction.

The  $D_{00}$  piece of the propagator doesn't look a whole lot different from the transverse photon propagator. But wouldn't it be nice if they were both part of something more symmetric! In fact, they are. We have the following:

**Claim:** We can replace the propagator  $D_{\mu\nu}(p)$  with the simpler, Lorentz invariant propagator

$$D_{\mu\nu}(p) = -\frac{\eta_{\mu\nu}}{p^2} \quad (6.85)$$

**Proof:** There is a general proof using current conservation. Here we'll be more pedestrian and show that we can do this for certain Feynman diagrams. In particular, we focus on a particular tree-level diagram that contributes to  $e^-e^- \rightarrow e^-e^-$  scattering,



where  $k = p - p' = q - q'$ . Recall that  $u(p)$  satisfies the equation

$$(p - m)u(p) = 0 \quad (6.87)$$

Let's define the spinor contractions  $\alpha^\mu = u^-(p \rightarrow r) \gamma^\mu u(p)$  and  $\beta^\nu = u^-(r \rightarrow q) \gamma^\nu u(q)$ . Then since  $k = p - p' = q - q'$ , we have

$$k_\mu \alpha^\mu = u^-(p \rightarrow r)(p - p') u(p) = u^-(p \rightarrow r)(m - m) u(p) = 0 \quad (6.88)$$

and, similarly,  $k_\nu \beta^\nu = 0$ . Using this fact, the diagram can be written as

$$\begin{aligned} \alpha^\mu D_{\mu\nu} \beta^\nu &= i \frac{\alpha \cdot \beta}{k^2} - \frac{(\alpha \cdot k)(\beta \cdot k)}{k^2 |k|^2} + \frac{\alpha^0 \beta^0}{|k|^2} \\ &= i \frac{\alpha \cdot \beta}{k^2} - \frac{0}{k^2 |k|^2} + \frac{\underline{\alpha_0 \beta_0}}{|k|^2} \\ &= i \frac{\alpha \cdot \beta}{k^2} - \frac{1}{k^2 |k|^2} \frac{(k^2 - k^2) \alpha \beta}{0 \ 0} \\ &= -\frac{i}{k^2} \alpha \cdot \beta = \alpha^\mu - \frac{i \eta_{\mu\nu}}{k^2} \beta^\nu \end{aligned} \quad (6.89)$$

which is the claimed result. You can similarly check that the same substitution is legal in the diagram

$$\sim e^2 [\bar{v}(q) \gamma^\mu u(p)] D_{\mu\nu}(k) [\bar{u}(p') \gamma^\nu v(q')] \quad (6.90)$$

In fact, although we won't show it here, it's a general fact that in every Feynman diagram we may use the very nice, Lorentz invariant propagator  $D_{\mu\nu} = -i\eta_{\mu\nu}/p^2$ .

**Note:** This is the propagator we found when quantizing in Lorentz gauge (using the Feynman gauge parameter). In general, quantizing the Lagrangian (6.37) in Lorentz gauge, we have the propagator

$$D_{\mu\nu} = \frac{i}{p^2} \eta_{\mu\nu} + (\alpha - 1) \frac{p_\mu p_\nu}{p^2} \quad (6.91)$$

Using similar arguments to those given above, you can show that the  $p_\mu p_\nu/p^2$  term cancels in all diagrams. For example, in the following diagrams the  $p_\mu p_\nu$  piece of the propagator contributes as

$$\sim \bar{u}(p') \gamma^\mu u(p) k_\mu = \bar{u}(p') (p' - p/\Gamma) u(p) = 0$$

$$\sim \bar{v}(p) \gamma^\mu u(q) k_\mu = \bar{u}(p) (p' + q'/\Gamma) u(q) = 0 \quad (6.92)$$

## 6.5 Feynman Rules

Finally, we have the Feynman rules for QED. For vertices and internal lines, we write

- Vertex:
- Photon Propagator:
- Fermion Propagator:

For external lines in the diagram, we attach

- Photons: We add a polarization vector  $\epsilon_{in}^\mu / \epsilon_{out}^\mu$  for incoming/outgoing photons. In Coulomb gauge,  $\epsilon^0 = 0$  and  $\rightarrow \epsilon \cdot \vec{p} = 0$ .
- Fermions: We add a spinor  $u^r(p) / u^r(p')$  for incoming/outgoing fermions. We add a spinor  $v^r(p) / v^r(p')$  for incoming/outgoing anti-fermions.

### 6.5.1 Charged Scalars

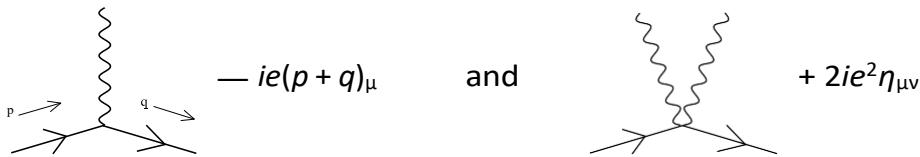
"Pauli asked me to calculate the cross section for pair creation of scalar particles by photons. It was only a short time after Bethe and Heitler had solved the same problem for electrons and positrons. I met Bethe in Copenhagen at a conference and asked him to tell me how he did the calculations. I also inquired how long it would take to perform this task; he answered, "It would take me three days, but you will need about three weeks." He was right, as usual; furthermore, the published cross sections were wrong by a factor of four."

*Viki Weisskopf*

The interaction terms in the Lagrangian for charged scalars come from the covariant derivative terms,

$$L = D_\mu \psi^\dagger D^\mu \psi = \partial_\mu \psi^\dagger \partial^\mu \psi - ie A_\mu (\psi^\dagger \partial^\mu \psi - \psi \partial^\mu \psi^\dagger) + e^2 A_\mu A^\mu \psi^\dagger \psi \quad (6.93)$$

This gives rise to two interaction vertices. But the cubic vertex is something we haven't seen before: it contains kinetic terms. How do these appear in the Feynman rules? After a Fourier transform, the derivative term means that the interaction is stronger for fermions with higher momentum, so we include a momentum factor in the Feynman rule. There is also a second, "seagull" graph. The two Feynman rules are



The factor of two in the seagull diagram arises because of the two identical particles appearing in the vertex. (It's the same reason that the  $1/4!$  didn't appear in the Feynman rules for  $\varphi^4$  theory).

### 6.6 Scattering in QED

Let's now calculate some amplitudes for various processes in quantum electrodynamics, with a photon coupled to a single fermion. We will consider the analogous set of processes that we saw in Section 3 and Section 5. We have

## Electron Scattering

Electron scattering  $e^-e^- \rightarrow e^-e^-$  is described by the two leading order Feynman diagrams, given by

$$= -i(-ie)^2 \frac{[\bar{u}^s(p)\gamma^\mu u^s(p)][\bar{u}^r(q)\gamma_\mu v^r(q)]}{(p^r-p)^2}$$

$$- \frac{[\bar{u}'^s(p)\gamma^\mu u^s(p)][\bar{u}'^r(q)\gamma_\mu v^r(q)]}{(p-q^r)^2}$$

The overall  $-i$  comes from the  $-i\eta_{\mu\nu}$  in the propagator, which contract the indices on the  $\gamma$ -matrices (remember that it's really positive for  $\mu, \nu = 1, 2, 3$ ).

## Electron Positron Annihilation

Let's now look at  $e^-e^+ \rightarrow 2\gamma$ , two gamma rays. The two lowest order Feynman diagrams are,

$$= i(-ie)^2 \frac{\gamma_\mu(p - p^r + m)\gamma_\nu}{(p - p^r)^2 - m^2}$$

$$+ \frac{\gamma_\mu(p - q^r + m)\gamma_\nu}{(p - q^r)^2 - m^2} u^s(p) \epsilon_1^\mu(p) \epsilon_2^\nu(q)$$

## Electron Positron Scattering

For  $e^-e^+ \rightarrow e^-e^+$  scattering (sometimes known as Bhabha scattering) the two lowest order Feynman diagrams are

$$= -i(-ie)^2 \frac{[\bar{u}'^s(p)\gamma^\mu u^s(p)][\bar{v}^r(q)\gamma_\mu v^r(q)]}{(p - p^r)^2}$$

$$+ \frac{[\bar{v}^r(q)\gamma^\mu u^s(p)][\bar{u}'^s(p)\gamma_\mu v^r(q)]}{(p + q)^2}$$

## Compton Scattering

The scattering of photons (in particular x-rays) off electrons  $e^- \gamma \rightarrow e^- \gamma$  is known as Compton scattering. Historically, the change in wavelength of the photon in the

scattering process was one of the conclusive pieces of evidence that light could behave as a particle. The amplitude is given by,

$$\begin{aligned}
 & \text{Diagram 1: } u(p) \xrightarrow{\epsilon_{\text{in}}(q)} u^-(p) \xrightarrow{\epsilon_{\text{out}}(q)} u^-(p) \\
 & \text{Diagram 2: } u(p) \xrightarrow{\epsilon_{\text{in}}(q)} u^-(p) \xrightarrow{\epsilon_{\text{out}}(q)} u^-(p) \\
 & = i(-ie) \bar{u}(p) r'(p) \frac{\gamma_\mu(p/q + q' - m)\gamma_\nu}{(p+q)^2 - m^2} + \frac{\gamma_\nu(p/q - q'^r + m)\gamma_\mu}{(p-q'^r)^2 - m^2} u^s(p) \epsilon_{\text{in}} \epsilon_{\text{out}}
 \end{aligned}$$

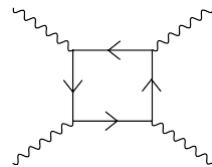
This amplitude vanishes for longitudinal photons. For example, suppose  $\epsilon_{\text{in}} \sim q$ . Then, using momentum conservation  $p + q = p^r + q^r$ , we may write the amplitude as

$$\begin{aligned}
 iA &= i(-ie) \bar{u}(p) \frac{\epsilon_{\text{out}}(p/q + q' - m) q}{(p+q)^2 - m^2} + \frac{q/(p^r - q' - m) \epsilon_{\text{out}}}{(p^r - q'^r)^2 - m^2} u^s(p) \\
 &= i(-ie) \bar{u}(p) \epsilon_{\text{out}} u^s(p) \frac{2p \cdot q}{(p+q)^2 - m^2} + \frac{2p^r \cdot q}{(p^r - q'^r)^2 - m^2} \quad (6.94)
 \end{aligned}$$

where, in going to the second line, we've performed some  $\gamma$ -matrix manipulations, together with the spinor equations  $(p' - m)u(p)$  and  $u^-(p \rightarrow r)(p'^r - m) = 0$ . We now recall the fact that  $q$  is a null vector, while  $p^2 = (p^r)^2 = m^2$  since the external legs are on mass-shell. This means that the two denominators in the amplitude read  $(p+q)^2 - m^2 = 2p \cdot q$  and  $(p^r - q'^r)^2 - m^2 = -2p^r \cdot q$ . This ensures that the combined amplitude vanishes for longitudinal photons as promised. A similar result holds when  $\epsilon_{\text{out}} \sim q^r$ .

## Photon Scattering

In QED, photons no longer pass through each other unimpeded. At one-loop, there is a diagram which leads to photon scattering. Although naively logarithmically divergent, the diagram is actually rendered finite by gauge invariance.



**Figure 30:**

## Adding Muons

Adding a second fermion into the mix, which we could identify as a muon, new processes become possible. For example, we can now have processes such as  $e^- \mu^- \rightarrow e^- \mu^-$  scattering, and  $e^+ e^-$  annihilation into a muon anti-muon pair. Using our standard notation of  $p$  and  $q$  for incoming momenta, and  $p^r$  and  $q^r$  for outgoing

momenta, we have the amplitudes given by

$$\sim \frac{1}{(p - p^r)^2} \quad \text{and} \quad \sim \frac{1}{(p + q)^2} \quad (6.95)$$

### 6.6.1 The Coulomb Potential

We've come a long way. We've understood how to compute quantum amplitudes in a large array of field theories. To end this course, we use our newfound knowledge to rederive a result you learnt in kindergarten: Coulomb's law.

To do this, we repeat our calculation that led us to the Yukawa force in Sections 3.5.2 and 5.7.2. We start by looking at  $e^-e^- \rightarrow e^-e^-$  scattering. We have

$$= -i(-ie)^2 \frac{[\bar{u}(\not{p}^r)\gamma^\mu u(\not{p})][\bar{u}(\not{q}^r)\gamma_\mu u(\not{q})]}{(p^r - p)^2} \quad (6.96)$$

Following (5.49), the non-relativistic limit of the spinor is  $u(p) \rightarrow \sqrt{\frac{m}{\xi}} \xi^{\frac{1}{2}}$ . This means that the  $\gamma^0$  piece of the interaction gives a term  $\bar{u}^s(\not{p}^r)\gamma^0 u^r(\not{q}) \rightarrow 2m\delta^{rs}$ , while the spatial  $\gamma^i$ ,  $i = 1, 2, 3$  pieces vanish in the non-relativistic limit:  $\bar{u}^s(\not{p}^r)\gamma^i u^r(\not{q}) \rightarrow 0$ . Comparing the scattering amplitude in this limit to that of non-relativistic quantum mechanics, we have the effective potential between two electrons given by,

$$U(r) = +e^2 \int \frac{d^3 p}{(2\pi)^3} \frac{e^{ip \cdot r}}{|\not{p}|^2} = +\frac{e^2}{4\pi r} \quad (6.97)$$

We find the familiar repulsive Coulomb potential. We can trace the minus sign that gives a repulsive potential to the fact that only the  $A_0$  component of the intermediate propagator  $\sim -i\eta_{\mu\nu}$  contributes in the non-relativistic limit.

For  $e^-e^+ \rightarrow e^-e^+$  scattering, the amplitude is

$$= +i(-ie)^2 \frac{[\bar{u}(\not{p}^r)\gamma^\mu u(\not{p})][\bar{v}(\not{q})\gamma_\mu v(\not{q}^r)]}{(p^r - p)^2} \quad (6.98)$$

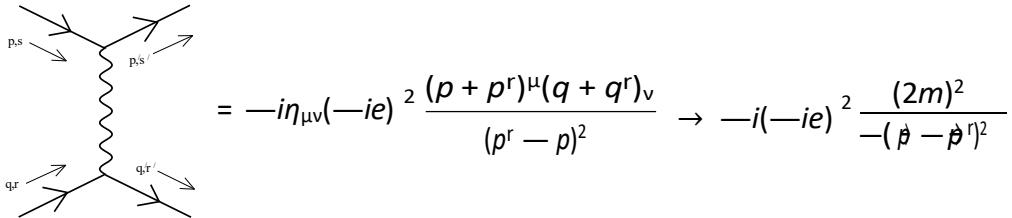
The overall + sign comes from treating the fermions correctly: we saw the same minus sign when studying scattering in Yukawa theory. The difference now comes from looking at the non-relativistic limit. We have  $\bar{v} \gamma^0 v \rightarrow 2m$ , giving us the potential between opposite charges,

$$U(r) = -e^2 \int \frac{d^3 p}{(2\pi)^3} \frac{e^{ip \cdot r}}{|p|^2} = \frac{e^2}{4\pi r} \quad (6.99)$$

Reassuringly, we find an attractive force between an electron and positron. The difference from the calculation of the Yukawa force comes again from the zeroth component of the gauge field, this time in the guise of the  $\gamma^0$  sandwiched between  $\bar{v} \gamma^0 v \rightarrow 2m$ , rather than the  $v^- v \rightarrow -2m$  that we saw in the Yukawa case.

### The Coulomb Potential for Scalars

There are many minus signs in the above calculation which somewhat obscure the crucial one which gives rise to the repulsive force. A careful study reveals the offending sign to be that which sits in front of the  $A_0$  piece of the photon propagator  $-i\eta_{\mu\nu}/p^2$ . Note that with our signature (+— —), the propagating  $A_i$  have the correct sign, while  $A_0$  comes with the wrong sign. This is simpler to see in the case of scalar QED, where we don't have to worry about the gamma matrices. From the Feynman rules of Section 6.5.1, we have the non-relativistic limit of scalar  $e^-e^-$  scattering,



where the non-relativistic limit in the numerator involves  $(p+p^r) \cdot (q+q^r) \approx (p+p^r)^0 (q+q^r)_0 \approx (2m)^2$  and is responsible for selecting the  $A_0$  part of the photon propagator rather than the  $A_i$  piece. This shows that the Coulomb potential for spin 0 particles of the same charge is again repulsive, just as it is for fermions. For  $e^-e^+$  scattering, the amplitude picks up an extra minus sign because the arrows on the legs of the Feynman rules in Section 6.5.1 are correlated with the momentum arrows. Flipping the arrows on one pair of legs in the amplitude introduces the relevant minus sign to ensure that the non-relativistic potential between  $e^-e^+$  is attractive as expected.

## 6.7 Afterword

In this course, we have laid the foundational framework for quantum field theory. Most of the developments that we've seen were already in place by the middle of the 1930s, pioneered by people such as Jordan, Dirac, Heisenberg, Pauli and Weisskopf<sup>5</sup>.

Yet by the end of the 1930s, physicists were ready to give up on quantum field theory. The difficulty lies in the next terms in perturbation theory. These are the terms that correspond to Feynman diagrams with loops in them, which we have scrupulously avoided computing in this course. The reason we've avoided them is because they typically give infinity! And, after ten years of trying, and failing, to make sense of this, the general feeling was that one should do something else. This from Dirac in 1937,

*Because of its extreme complexity, most physicists will be glad to see the end of QED*

But the leading figures of the day gave up too soon. It took a new generation of postwar physicists — people like Schwinger, Feynman, Tomonaga and Dyson — to return to quantum field theory and tame the infinities. The story of how they did that will be told in next term's course.

---

<sup>5</sup>For more details on the history of quantum field theory, see the excellent book “QED and the Men who Made it” by Sam Schweber.